Automatic spinal cord segmentation as organ at risk in planning CT using adaptive template matching and U-Net

João Diniz[1], Jonnison Ferrreira[2], Pedro Diniz[1], Bruno Serejo[1], Neilson Ribeiro[1], Osias Santos[1], Aristófanes Silva[2], Anselmo Paiva[2]

[1] Instituto Federal do Maranhão (IFMA) – Grajaú – MA

[2]Universidade Federal do Maranhão (UFMA) – São Luís – MA

joao.bandeira@ifma.edu.br

Abstract. *Radiotherapy is one of the major option used in cancer management. The treatment involves several steps, one of which is the construction of a computed tomography (CT) model of the patient so that the target tissues and organs at risk (OARs) surrounding that target can be evaluated. With the CT, the responsible physician delimits the OARs slice by slice, as the spinal cord that comprises almost all the tomography becomes more tiring to be segmented and thus susceptible to errors. Thus, this paper presents a method of spinal cord segmentation in planning CT for radiotherapy. The result achieved an accuracy of 99.18%, specificity of 99.30%, sensitivity of 92.64%, and dice index of 80.83%, without any segmentation refinement.*

1. Introduction

One treatment option in about 50% of cancer cases is radiotherapy (RT). Because it is an attractive, effective and common therapy option, RT is a widely used treatment, especially when surgery and chemotherapy introduce great risk to the patient's life. The radiotherapy process consists of delivering radiation to abnormal tissues so that they do not affect the healthy tissues surrounding them. Thus, RT is widely used as cancer treatment and prevention because these tissues are sensitive to radiation while normal tissues better support the incidence of radiation [Evans and Staffurth 2018]. However, although they can withstand better radiation doses, excess radiation can be toxic to organs, so planning radiation therapy is so important.

At the planning step, a computed tomography (CT) of the patient is performed so that the target tissues and healthy tissues are segmented. These tissues of healthy organs that need to be protected are called Organs at Risk (OAR) [Tsang and Hoskin 2017]. The spinal cord is an extremely important OAR, any damage to this organ can cause irreversible damage to the patient, such as total paralysis and partial neurologic loss.

The purpose of this paper is to propose a computational methodology for spinal cord segmentation in planning CT for radiotherapy. By devising an automatic method for spinal cord segmentation, it is believed that we have achieved the following contributions: (a) use of U-Net considered state of the art in segmentation of images for the task of spinal cord segmentation; (b) a useful method of automatically segmenting the spinal cord in CT; and (c) several advances in spinal cord segmentation task compared to previous works. Thus, it is believed that if used in large medical centers, can be a key ally for the expert in the segmentation task of the spinal cord.

2. Related works

Studies can be seen in the literature with spinal cord segmentation, either in CT scans or MRI. These methods are still scarce when we talk about segmentation as an OARs.

A knowledge-based method for spinal cord segmentation is proposed in [Banik et al. 2010]. The technique used for spinal canal segmentation is based on Hough transform, so the authors detect circles that fit the spinal canal and have a ”kickstart”of the spinal cord location and then morphological operations are used to refine this segmentation. The method proposed by [Chen et al. 2013] made the segmentation of the spinal cord into magnetic resonance imaging. They propose spinal cord segmentation using a deformable atlas and topological spinal cord location information. The results achieved by the authors were a Dice index of 85% and an accuracy of 91%.

[De Leener et al. 2015] also presents spinal cord segmentation in magnetic resonance imaging. The authors proposed use of Hough transform to find the spinal cord and from this uses a knowledge-based approach to segmentation refining. The results achieved a Dice index of 91%. A recent work proposed by [Fu et al. 2018] also shows the importance of segmentation of the spinal canal (where the spinal cord is located). The method presents a segmentation algorithm based on extract the spinal canal from the CT images. The Hough transform is also used in this work. The method used 10 patients and achieved an average accuracy of 77.32%.

A proposed work considered state of the art by the techniques employed and by the results achieved is presented in [Diniz et al. 2019]. The presented method is divided into 3 steps: first uses a variation of the template matching to decrease the region; the second uses a superpixel algorithm to group candidate regions of the spinal cord and non-spinal cord; finally, a convolutional neural network (CNN) is used to classify these regions. The method uses a database with 36 patients and the results achieved by this work are 92.55% accuracy and Dice index of 78%.

As shown in the aforementioned works, the importance of spinal cord segmentation is in various imaging modalities, which is summarized to assist the medical specialist. Most of the work uses previous knowledge for the task of segmentation of the spinal cord. Also, some techniques use Hough transforms and other knowledge-based approaches to spinal cord segmentation. Observing the state of the art, we propose a method for spinal cord segmentation as an OAR in computed tomography. For this, we use U-Net based approach that do not require knowledge-based information.

3. Materials and method

3.1. Image acquisition

To validate the methodology, there was a need to use a very specific database which has OARs marking in radiotherapy planning, especially the spinal cord. The database used to build the method is from a challenge called AAPM Thoracic Auto-segmentation[1]. In total there are 36 volumes of CT distributed equally in 3 different institutes, ie 12 patients per institute. The ground truth that accompanies each database volume is available in a Radiotherapy Structure Set (RT-STRUCT) format. In this challenge, each RT-STRUCT is

[1]Available: $http://aapmchallenges.cloudapp.net/competitions/3\#learn\_the\_details$

labeled with 5 organs, namely right and left lung, heart, esophagus and spinal cord. The proposed method will only use the marking of the spinal cord.

3.2. Initial segmentation

Because it is a very large three-dimensional volume with several slices, and because the spinal cord comprises a very small region of the volume, the initial segmentation step has the task of generating a subvolume of interest where the original volume is reduced to only part comprising the spinal cord and its circumscribing structures. For this task, a very common technique was used in the context of image processing and pattern recognition called template matching. Briefly speaking, from one image, the template matching technique finds the area with the most pattern similarity to another, usually smaller, image called the template. In this work, as in the work of [Diniz et al. 2019], the square difference combination method was used as a similarity metric calculation.

3.2.1. Adaptive template matching

First it is necessary to define the standard template. We use the same pipeline describe by [Diniz et al. 2019] to define this standard template. First, a random volume is selected in the database . Then a slice of this volume is also randomly selected. Then the RT-STRUCT file is used and found where there is a ground truth, so the center of mass of this marking is found, and a region corresponding to twice the size of the marking is cropped on all sides of this slide.

The next step is define the initial template. The process to build the initial template is defined as follows: first, the template matching algorithm runs on each slice of the volume, where the template is the standard template; second, the similarity for each matching in each slice is calculated; third, the initial template is defined as the one that the matching has the greatest similarity in relation to the standard template, also the number of the slice that generated the best matching is saved. At the end of this step, there is the initial template, which will be unique for each volume.

The initial template will be adapted to each patient. Template matching will start to execute in slice where the matching occurred, so will result in two artifacts: the new template (how will be adaptable for each slice of each volume, red bounding boxes in Figure 1 represents the matching for each slice and the new template) and the resulting image. The new template will execute in the next slice (as shown in the Figure 1: Slices $n + 1$ and $n + 2$), again, the result will be the new template and the resulting image. It should be noted that the slice where the initial template occurred may be in the middle of the volume, so the same process happens for the previous slices (as shown in the Figure 1: Slices $n - 1$ and $n - 2$). The result must be a region of interest with the matching in each slice of all volume.

We propose to execute the template matching only in the region corresponding to twice the matching region in the following slices (green bounding boxes Figure in 1) and not in the whole slice (as proposed by [Diniz et al. 2019]), in this way we reduce the algorithm computational cost and also that artifacts outside the patient's body do not disturb the algorithm

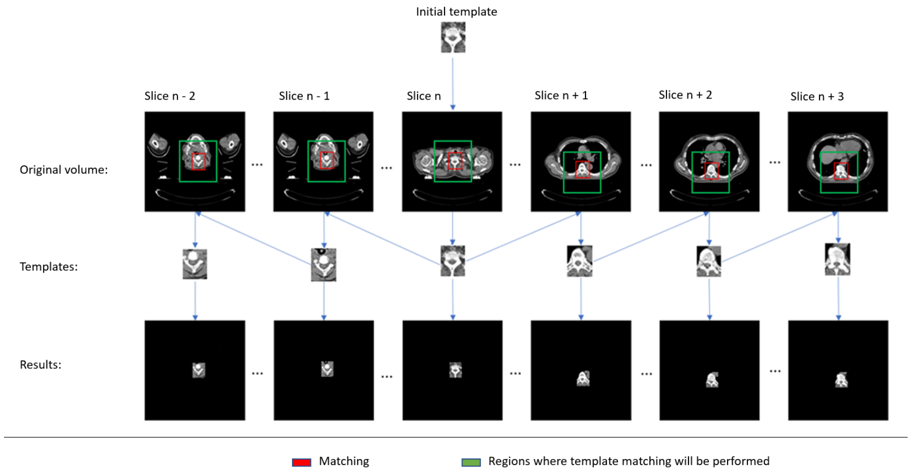

Figura 1. Adaptive template matching.

3.3. Final segmentation

At the end of the initial segmentation, we have a volume of interest (VOI) of each patient. This VOI consists of the spinal cord and the structures around it. Also what is done with the patient, is done with the marking of the specialist, i.e., a VOI with the equal dimension of the patient is generated, but with the marking of the specialist. For the segmentation of the spinal cord the VOI of each patient was passed slice by slice into the network.

The U-Net is a CNN, has the use of same concepts and layers, such as convolutional layers, pooling layers, activation layers, and dropout layers. The first authors to propose the use of this network were [Ronneberger et al. 2015]. U-Net consists of a contraction path to capture the context and asymmetric expansion path that allows precise segmentation. In short, this network can be trained end-to-end, where U-Net simply concatenates the encoder feature maps to map decoder feature maps at all stages to form a ladder structure. This architecture by its concatenation connections allows the decoder at each stage to learn the relevant features that are lost when grouped into the encoder.

The U-Net used in this work is the same proposed by [Ronneberger et al. 2015], differs only in the fact that the input slice has a different size from that used in the architecture proposed by them. Which is passed to the network, are slices of the patient's VOI along with the slice of the specialist's marking. The output of the network will be a slice classified as a spinal cord. The network is trained with the dice loss function, represented by $diceloss = 1 - \frac{2TP}{2TP+FP+FN}$. To validate our method, validation metrics as Dice index, sensitivity (SEN), specificity (ESP), accuracy (ACC), and area under the curve roc (AUC) are used.

4. Results and discussion

To train and validate our method, we divided the database into 2 datasets. The first for training, consisting of 30 patients (10 per institute) and 6 for testing (2 per institute). This division was made randomly, only guaranteeing the proportion by the institute in the two datasets. In the initial segmentation step, both datasets passed for the proposed algorithm using template matching. There was no spinal cord loss in any of the slices of any patient.

In the final segmentation, the result found by the initial segmentation will be presented to U-Net. The images used for the network input were of 128×64, based on the randomness of the standard template creation, this was the size that did not exceed the generated image. After several training sections, and parameter setting, the parameters used in the training of the nets were: number of epochs equal to 50, size of *batch* equal to 1, Adadelta optimizer with initial learning rate equal to 1 and 10% of volumes for validation.

After model generation with 30 volumes of training, they were applied to the 6 test volumes (2 patients per institute) and validation metrics were taken. We made tests with two loss functions, cross-entropy and dice. These results are described in Table 1.

Tabela 1. Results of the step of final segmentation

Loss function	Dice(%)	SEN(%)	ESP(%)	ACC(%)	AUC(%)
Cross-entropy	79.18	83.62	99.47	99.18	91.55
Dice	80.83	92.64	99.30	99.24	95.97

Analyzing the Table 1 with cross-entropy, we can identify that the method was promising in the spinal cord segmentation. Even with ESP, ACC and AUC metrics greater than 90%, the SEN value that says the number of positive class cases were correct was less just 83.92%. By using the dice loss, all metrics were above 90% and an improvement in the dice value was achieved. It is noteworthy that as described in [Diniz et al. 2019], many patient slices do not have markings, which detract from the dice.

To fit our method into the literature, we will compare the related work.

Tabela 2. Comparison of the state of arts methods (K-B means Knowlegde-based)

Works	Technique	Subject	Dice(%)	SEN(%)	ESP(%)	ACC(%)	AUC(%)
[Banik et al. 2010]	K-B	13	60				
[Chen et al. 2013]	K-B	25	85			91	
[De Leener et al. 2015]	K-B	17	91				
[Fu et al. 2018]	CNN	36		72	78	89	
[Diniz et al. 2019]	CNN	36	78	89	93	92	
Proposed Method (U-Net)	U-Net	36	81.69	91.52	99.38	99.24	95.45

As we can see, several methods that do spinal cord segmentation utilize some sort of a priori knowledge. Still, method like [De Leener et al. 2015] use magnetic resonance to make the detection. It is noteworthy that the spinal cord segmentation in planning computed tomography is delimited by the bone limits, and therefore some of these works, make detection of the spinal canal also served as a comparison.

Looking at the Table 2, none of the literature has greater accuracy than our proposed method using U-Net. When comparing with works that use K-B, our work fills the gap of relying on some technique such as atlas or active contour. Moreover, work like [Banik et al. 2010], use Hough transform for segmentation, which is not very usual, since it is not sensitive to the edges of the bone limit. Despite the work of [Chen et al. 2013, De Leener et al. 2015] have a Dice higher than the proposed method, we used a larger amount of planning CT images and does not require prior knowledge. Still, when looking at the other metrics, our method surpasses in metrics of accuracy, sensitivity, specificity, and AUC.

New approaches like [Diniz et al. 2019, Fu et al. 2018] propose the use of CNN for spinal cord segmentation. As already reported, this kind of approach is promising, since using CNN abstracts the extraction and feature selection steps. However, there is always a need to generate an input image, either in patches or with some form of information. Thus, we highlight that the proposed method that uses U-Net achieved satisfactory results in spinal cord segmentation as a method fully automatically, which, besides not using prior knowledge, can also perform a more precise segmentation.

5. Conclusion

The proposed method is composed of two steps. Initial segmentation consists of the use of an adaptive template matching, which decreases the area of interest in the entire patient volume. The second step is to classify this area of interest into spinal cord using U-Net. The result found were dice index of 80.83%, sensitivity of 92.64%, specificity of 99.30%, accuracy of 99.18% and AUC of 95.97%. Thus, the proposed method proved to be promising in the task of spinal cord segmentation, serving as an ally to the specialist in large centers. As future studies, we include others U-Net variations to verify the efficiency in the segmentation. In addition, the use of 3D U-Net aggregating more information. Also, it is suggested to use some bio-inspired algorithm to find the best parameters.

Referências

Banik, S., Rangayyan, R. M., and Boag, G. S. (2010). Automatic segmentation of the ribs, the vertebral column, and the spinal canal in pediatric computed tomographic images. *Journal of digital imaging*, 23(3):301–322.

Chen, M., Carass, A., Oh, J., Nair, G., Pham, D. L., Reich, D. S., and Prince, J. L. (2013). Automatic magnetic resonance spinal cord segmentation with topology constraints for variable fields of view. *Neuroimage*, 83:1051–1062.

De Leener, B., Cohen-Adad, J., and Kadoury, S. (2015). Automatic segmentation of the spinal cord and spinal canal coupled with vertebral labeling. *IEEE transactions on medical imaging*, 34(8):1705–1718.

Diniz, J. O. B., Diniz, P. H. B., Valente, T. L. A., Silva, A. C., and Paiva, A. C. (2019). Spinal cord detection in planning ct for radiotherapy through adaptive template matching, imslic and convolutional neural networks. *Computer methods and programs in biomedicine*, 170(3):53–67.

Evans, E. and Staffurth, J. (2018). Principles of cancer treatment by radiotherapy. *Surgery-Oxford International Edition*, 36(3):111–116.

Fu, G., Lu, H., Tan, J. K., Kim, H., Zhu, X., and Lu, J. (2018). Segmentation of spinal canal region in ct images using 3d region growing technique. In *2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT)*, pages 1–4. IEEE.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.

Tsang, Y. M. and Hoskin, P. (2017). 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*, 3:37–40.