Automated Ventral Cavity Segmentation in Computed Tomography

Rui Castro[1]
ruicastro2000@gmail.com
Inês Sousa[2]
inessousa01@gmail.com
Fábio Nunes[2,3]
fsn@med.up.pt
Jennifer Mancio[2]
jenni@med.up.pt
Ricardo Fontes-Carvalho[2,3]
ricardo@med.up.pt
Carlos Ferreira[1,4]
carlos.a.ferreira@inesctec.pt
João Pedrosa[1,4]
joao.m.pedrosa@inesctec.pt

[1] Faculdade de Engenharia da Universidade do Porto
Porto, Portugal
[2] Faculdade de Medicina da Universidade do Porto
Porto, Portugal
[3] Centro Hospitalar de Vila Nova de Gaia e Espinho
Porto, Portugal
[4] Instituto de Engenharia de Sistemas e Computadores, Tecnologia e Ciência
Porto, Portugal

Abstract

Coronary Artery Disease is one of the leading causes of death worldwide. Computed Tomography and Coronary Computed Tomography Angiography are the gold standard techniques for Coronary Artery Disease diagnosis. Some recent studies have found a connection between Coronary Artery Disease occurrence and the accumulation of visceral adipose tissue in the ventral cavity. By performing ventral cavity segmentations in Computed Tomography scans, it is possible to quantify and analyze important textural characteristics of visceral fat. However the manual delineation of this structures is a time consuming process subject to variability. An automated process would achieve a faster and more precise solution. This paper explores the use of a U-Net architecture to perform ventral cavity segmentations. Experiments with different input image sizes and types of loss functions were employed. The model with the best performance achieved a 0.974 Dice Score Coefficient which is a competitive result when compared to the state of the art methods.

1 Introduction

Coronary artery disease (CAD) is the most common type of heart disease and the third leading cause of death worldwide with 17.8 million deaths annually [2]. Computed Tomography (CT), as well as, Coronary Computed Tomography Angiography (CCTA) are the main techniques used for CAD diagnosis in medical imaging.

Recent studies have indicated that the accumulation of visceral adipose tissue (VAT), rather than subcutaneous adipose tissue (SAT), in the abdominal region, is associated with increased cardiometabolic risk [3]. VAT could thus be studied as an alternative biomarker for CAD. By analyzing VAT and SAT with CT, it is possible to obtain not only a quantification of this adipose tissue but also gather important characteristics such as density as well as textural characteristics (radiomics) that can indicate metabolic changes of the tissue noninvasively, thus providing more robust clinical information. Manually outlining these tissues is a time-consuming task, subject to inconsistency. An automated approach would accomplish a quicker and more accurate segmentation. The use of deep neural network architectures emerge as an obvious solution to this problem as they can obtain automatic VAT/SAT segmentations with high performance that consume less time. In addition, they are able to extract important features that can generate predictions about CAD risk, providing a trustworthy second opinion to help clinical professionals in CAD diagnosis. The majority state of the art approaches for ventral cavity segmentation use either Deep Learning architectures such as the work of and Weston et al. [6] or Threshold-based methods as seen in the work of Nemoto et al. [4].

The objective of this study was the development and validation of an automated deep learning approach for the ventral cavity segmentation in CT scans and validation of its clinical applicability.

2 Methods

2.1 Dataset

The dataset used was the CHVNGE dataset, which is composed of 886 patients randomly selected from inpatients at the Centro Hospitalar de Vila Nova de Gaia (CHVNGE) in Vila Nova de Gaia, Portugal. Ethical Committee approval was obtained prior to data collection and all data was anonymized prior to analysis for the purposes of this study. The dataset includes 886 2D abdominal CT scans (512×512 pixels) acquired on a Siemens Somatom Sensation 64 with a slice thickness of 3 mm. The VAT segmentations were obtained via manual annotation by a medical student using the 3DSlicer software.

2.2 Ventral Cavity Segmentation

In this study, a U-Net model [5] was trained to perform the ventral cavity segmentation. The 886 images and segmentation references were split into 5 folds to perform cross validation. In each fold rotation, 3 folds were used for training, 1 for validation and 1 for testing. Each training iteration was ran by 400 epochs with a 0.00001 learning rate. When training the model, an early stop criteria was imposed to reduce computational resources. Model weights were updated with the ADAM optimizer [1].

Model training was accompanied by experiments that analyzed the influence of different parameters on its performance, such as:

- Image Sizes: 512×512, 256×256 and 256×256 cropped;
- Loss Function: Binary Cross Entropy (BCE) and Dice.

Different image sizes were used to study their dependence in the U-Net model performance. The 256×256 cropped image is obtained through the following steps. Initially, a body mask is created by applying a Hounsfield Unit (HU) threshold, retaining pixel values above 300. This threshold removes image artifacts located under the body. The patient's body is thereby isolated and a bounding box surrounding it is computed. Cropping is performed by resizing the longer side of the bounding box to 256 pixels while maintaining the original aspect ratio. Zero padding is performed in order to complete the 256×256 image.

The first loss function applied was a Binary Cross Entropy loss function. The unreduced loss can be described as:

$$BCELoss = -y\log(x) - (1-y)\log(1-x), \tag{1}$$

where x and y represent the predictions and targets, respectively.

The other loss function used in this study is a Dice loss that measures the dissimilarity between predicted and target ground truth segmentation masks by quantifying the overlap between the two masks. The Dice Loss is computed as 1 minus the DSC (Dice Score Coefficient) which inversely measures the similarity between both masks. The equation describing the Dice Loss is as follows:

$$DiceLoss = 1 - \frac{2\sum_j (x_j \times y_j)}{\sum_j x_j + \sum_j y_j} \tag{2}$$

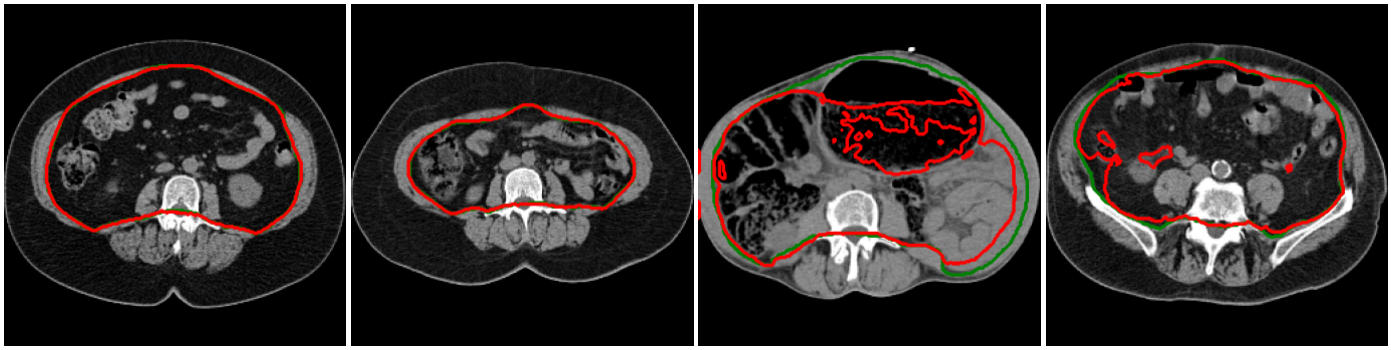

Figure 1: Best (two leftmost) and worst (two rightmost) examples regarding the DSC and HD. The green and red contours represent the ground truth and predicted segmentations, respectively.

Data augmentation techniques were also employed in the training and validation sets in order to create data variation to enhance the model's ability to generalize and perform well on new and unseen data. These techniques consisted of image resizing, rotations up to 35 degrees, horizontal flip transformations with a probability of 0.5, vertical flip transformations with a probability of 0.1 and a [0,255] image normalization. All of these transformations were applied to the training and validating sets, besides the rotation and flip transformations which were only applied to the training sets. Several evaluation metrics were used to study this methods' performance such as the DSC, Hausdorff Distance (HD) and the Mean Absolute Distance (MAD).

3 Results and Discussion

Table 1 shows the DSC, HD and MAD results for each of the trained models.

Table 1: Evaluation metrics results for each trained model.

Input Size	Loss Function	Evaluation Metrics		
		DSC	HD	MAD
512	BCE	0.950±0.023	77.89±83.34	2.88±2.10
	DICE	0.372±0.171	141.07±12.96	27.89±2.21
256	BCE	0.960±0.003	8.84±1.39	1.47±0.10
	DICE	0.902±0.115	33.85±25.85	5.04±5.86
256-Crop	BCE	**0.968±0.003**	**7.60±1.10**	**1.24±0.11**
	DICE	0.966±0.005	9.30±1.88	1.33±0.13

The U-Net model was capable of performing accurate ventral cavity segmentations. One clear trend that can be seen is the fact that one specific set of parameters provide the best results. The model that was trained with 256×256 cropped images and a BCE loss function achieved a DSC of 0.968 ± 0.003 and HD and MAD values of 7.60 ± 1.10 and 1.24 ± 0.11, respectively. These are competitive results when compared with the state of the art results [6][4].

Regarding image size, there is evidence that a smaller image size provide better results when comparing the 512×512 and 256×256 images. Between both 256×256 image sizes (cropped or not), the cropped version leads to a better performance. The U-Net architecture is designed to capture both local and global contextual information. If the image is too large, the network might not be able to capture the necessary global context effectively due to a limited receptive field. This can be one of the reasons why smaller image sizes show a better performance. The cropped version of images provide the same smaller size with bigger resolution, since it zooms in on the body of the patient. This results in images with finer details and a bigger focus on regions of interest that can be the reason for providing better performances.

When comparing the results between the different loss functions, one can conclude that the BCE loss function is better suited for this set of data as it achieved a better performance (higher DSC and lower HD/MAD) for each different image size.

In Figure 1, it is possible to see the best and worst results in terms of DSC and HD, by overlapping the original input image with the ground truth and prediction mask contours. This examples were taken from the predictions generated by the model that achieved the best results.

Looking at the worst example for DSC, it is easy to comprehend that this CT scan came as such a bad result, since it is an abnormal image. The black hole shown in the figure is something unusual in this type of images, and as such, when it enters the model, which is not used to seeing these types of structures, it will generate a prediction that is subject to errors and will be very different from the ground truth.

4 Conclusion

In this study, an investigation into automated ventral cavity segmentation in CT scans, using a U-Net approach, is presented. The employed experiments explore the effects of varying the input image sizes and loss function in the segmentation process, which resulted in a good performance when compared to the state of the art methods. This method has an advantage as it uses CT scans performed with low doses of X-rays, which prevent the contact of patients with contrast agents.

Future work could be later performed in order to continue and evolve this study, such as the employment of post processing techniques like noise removal and filling operations in the predicted masks. These predictions could then be used to perform segmentations on the ventral cavity components, such as VAT. These segmentations can then be used for the quantification of VAT and the extraction of important radiomics features for CAD prediction.

Acknowledgments

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within projects UIDB/50014/2020 (uxCADCT) and LA/P/0063/2020. João Pedrosa and Carlos Ferreira are supported by Fundação para a Ciência e Tecnologia (FCT) through application references 2022.06138.CEECIND and SFRH/BD/146437/2019 respectively.

References

[1] ADAM optimizer. `https://pytorch.org/docs/stable/generated/torch.optim.Adam.html`. Accessed: 2023-08-22.

[2] Jonathan C Brown, Thomas E Gerhardt, and Edward Kwon. Risk factors for coronary artery disease. 2020.

[3] Yuji Matsuzawa, Tadashi Nakamura, Iichiro Shimomura, and Kazuaki Kotani. Visceral fat accumulation and cardiovascular disease. *Obesity research*, 3(S5):645S–647S, 1995.

[4] Yeernuer Nemoto, Mitsutaka et al. Development of automatic visceral fat volume calculation software for ct volume data. *Journal of obesity*, 2014, 2014.

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[6] Alexander D Weston, Panagiotis Korfiatis, Kline, et al. Automated abdominal segmentation of ct scans for body composition analysis using deep learning. *Radiology*, 290(3):669–679, 2019.