Semantic Segmentation with DenseNets for Carotid Artery Ultrasound Interpretation

Anonymous

No Institute Given

Abstract. The measurement of Carotid Intima Media Thickness (CIMT) in ultrasound images can be used to detect the presence of atherosclerotic plaques, which may appear in several territories of the carotid artery. Usually, the CIMT estimation is semi-automatic, since it requires: 1) a manual examination of the ultrasound image for the localization of a Region Of Interest (ROI), and 2) an automatic delineation of the CIM region within the ROI. The existing efforts for automating the process have replicated the same two-step structure, resulting in two consecutive independent approaches. In this work, we explore a single-step approach, based on Densely Connected Convolutional Neural Networks (DenseNets), for semantic segmentation of the whole image, which allows adding contextual information in the image interpretation. The CIMT estimation and plaque detection are validated with a large dataset of more than 8,000 images (more than 2,000 subjects) showing that the proposed method is accurate and objective.

Keywords: Semantic Segmentation of Carotid Artery, Intima Media Thickness, Ultrasound Images, Carotid Plaque Detection

1 Introduction

Cardiovascular Diseases (CVDs) are the leading cause of death in developed countries. Most of at-risk individuals of cardiovascular events suffer atherosclerosis, a chronic inflammatory process characterized morphologically by an asymmetric focal thickening of the innermost layer of the artery. The Ultrasound (US) Carotid Artery (CA) images are used to detect the burden of atherosclerosis, since they provide the possibility to measure the Carotid Intima Media Thickness (CIMT) of the artery and identify the presence of atherosclerotic plaques. The CIM region is defined by the lumen-intima and media-adventitia interfaces, and the CIMT is commonly estimated in the far wall of the CA (see Figure 1). The Mannheim Consensus [15] defines a sufficient criterion for plaque detection: plaques are structures into the arterial lumen showing CIMT$\geq$1.5mm.

Basic techniques for CIM region segmentation presented in the literature include, among others, hough transform [9], active contour [1], edge detection [12] and other solutions as integrated approaches that combine several basic methods [2]. Recently, preliminary efforts using machine learning in CIMT evaluation have been presented [7,16,14]. In [7], a standard multilayer perceptron with an

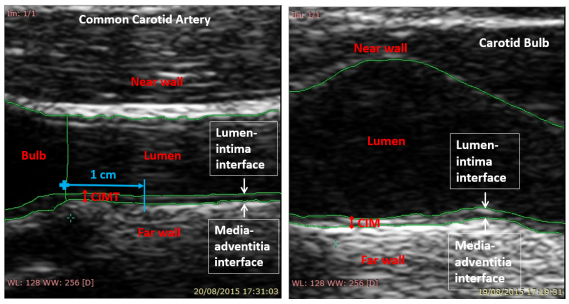

Fig. 1. Common Carotid Artery (CCA) (left) and Bulb (right) ultrasound longitudinal images. The different parts of the CA are delimited with green lines. The CIMT is measured approximately 1 cm distal from the carotid Bulb on the far wall in CCA.

auto-encoder is proposed for CIMT image interpretation, but it did not outperform the snake-based method in [1]. Zhang et al. [16] propose a two-step segmentation method of the CIM region based on patch-based classification and Stacked Sequential Learning. More recently, in [14], patch-based Convolutional Neural Networks (CNNs) are used in the proposed steps for CIMT estimation. Moreover, in [10], a two-step machine learning method for atherosclerotic plaque segmentation is presented. For a more complete survey of methods for automatic CIMT measurements and plaque detection, please refer to the review studies [8] and [6], respectively.

To our knowledge, all the aforementioned methods are two-step approaches which define separate methods to, first, localize the Region Of Interest (ROI) and second, delineate the CIM region within the ROI. Moreover, all the presented works focus on Common Carotid Artery (CCA) images, since the imaging quality of other territories, such as Bulb, is worse than CCA (poorer contrast and more affected by noise) making difficult to segment the CIM region.

In this paper, we present the following contributions. 1) We propose a new approach based on semantic segmentation for CA image interpretation, which represents the first attempt in the literature. This method allows to accurately localize and interpret the different components of the CA (lumen, far wall, near wall, bulb and CIM region, see Figure 1). This information avoids using a separate step for ROI localization, making our segmentation method a single-step approach. Moreover, we define an approach for CIMT estimation. 2) We present an extensive evaluation of the CIMT measurement and plaque detection in a large dataset (8,484 images). Unlike most of previous works that only measure the CIMT on the CCA within a plaque-free region, we broaden the target and build a more general method which is able to accurately estimate the CIMT, even in the presence of plaque. This feature makes this fully automatic method useful for population studies, as the dataset considered in this paper. 3) We demonstrate that the method is easily extensible to different CA territories, after being successfully trained for both CCA and Bulb. 4) We compare the obtained CIM segmentation results with other state-of-the-art approaches to show the

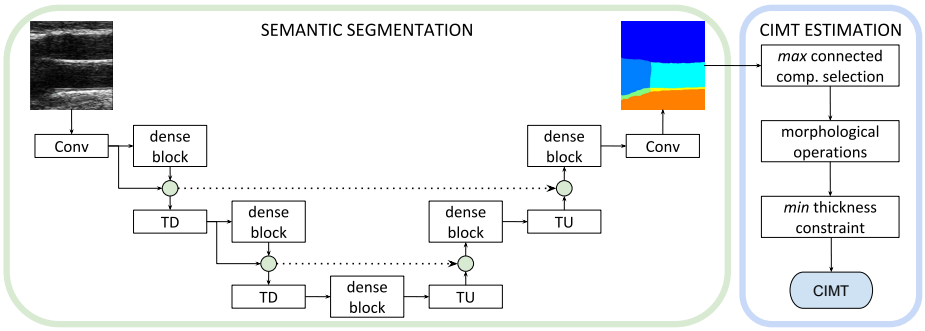

Fig. 2. Workflow of the proposed method for semantic carotid artery segmentation and CIMT estimation. The semantic segmentation model is composed of a down-sampling path with Transition Down (TD) blocks, and an up-sampling path with Transition Up (TU) blocks, both including dense blocks that create the feature maps. Additionally, a Convolution (Conv) is applied at the input of the network as well as at the end, to generate the final segmentation. The small circles represent concatenations, and the dotted arrows are the skip connections.

improvement of the proposed method. 5)We measure the Inter-Observer Variability (IOV) of the manual segmentation showing the degree of difficulty of the problem at hand, especially in the case of Bulb images.

2 Method

Fully Convolutional Networks (FCN) [13], commonly used in semantic segmentation problems, are a particular case of Convolutional Neural Networks (CNN). Among CNN models extended as FCNs, Densely Connected Convolutional Networks (DenseNets) should be highlighted.

DenseNets [4] can be seen as an extension of Residual Networks (ResNets), which have been designed to ease the training of very deep networks. In particular, DenseNets present some characteristics that make them very appropriate for semantic segmentation: parameter efficiency, implicit deep supervision, and feature reuse. For these reasons, we propose the use of the so-called Tiramisu model [5], an extension of DenseNets as FCNs.

The Tiramisu architecture (see Figure 2) is composed of a down-sampling path that extracts coarse semantic features, and a up-sampling path that recovers the input image resolution at the output level. Both paths are connected by means of skip connections that allow to recover the fine-grained information. Our implementation[1] of the model is in Keras[2], with Theano as backend.

The CIMT estimation contains three steps (see Figure 2). First, the biggest connected component which best fits (shape and localization) the CIM region is

[1] Code available after paper acceptance
[2] `https://keras.io/`

selected. Second, simple morphological operations are applied in order to smooth the borders of the region. Third, a minimum thickness value of 0.4mm is imposed following the reference values in [3]. Finally, when analyzing CCA images, the region is limited to 1 cm distal from the bulb component, as it is justified from a clinical standpoint. The CIMT mean value is computed as the mean absolute distance between the bottom border (media-adventitia) and the upper border (lumen-intima) of the region. Afterwards, each image is classified as containing plaque or non-plaque following the Mannheim Consensus (see Section 1).

3 Experiments

3.1 Data Set

We consider a sample of 2,379 subjects from *Anonymous* Heart Register. The images were collected from 2007 to 2010, and the subjects represent general population aged between 35 and 84. Two trained sonographers performed the CA US scans with an Acuson XP128 US system equipped with L75-10 MHz transducer and a computer program extended frequency (Siemens-Acuson). US longitudinal images were obtained in B-mode with resolution 23.5 pixels/mm. All the images were obtained from left and right CA resulting in a total of 4,751 CCA images and 3,733 Bulb images. The CIMT reference values (Ground-Truth, GT) were given by *Anonymous* Medical Center. All the images were analyzed by an expert using the semi-automatic software e-track.

To asses the segmentation results, a subset of 159 CCA images (51 with plaque and 108 without plaque) and 78 Bulb images (12 with plaque and 66 without plaque) from the dataset were manually delineated by an expert (Expert1), using 6 labels for CCA and 4 for Bulb (see Figure 1). The train set contained 141 CCA images and 65 Bulb images, whilst the rest of them were used for testing. Additionally, the test images were manually segmented by a second expert (Expert2) to measure the IOV.

3.2 Validation Setup

To validate the segmentation method, we compared six approaches: four DenseNets models based on Tiramisu [5], the U-Net method [11] and the shallow method, Random Forest (RF). Regarding the Tiramisu model, the networks trained were: Tiramisu56 (a total of 56 layers, 4 per dense block) and Tiramisu103 (a total of 103 layers, from 4 to 12 per block). In order to show if the Semantic Segmentation helps in the CIM region segmentation, we compared the results provided by the same Tiramisu models but using only two labels (CIM region and background). We call this binary approach Binary Segmentation (BS), and the one with more than two labels, Semantic Segmentation (SS). U-Net was also trained for comparison to demonstrate the adequacy of using DenseNets, since its main difference with Tiramisu is that U-Net uses standard convolutions instead of dense blocks. We call RF2 the two-step approach in

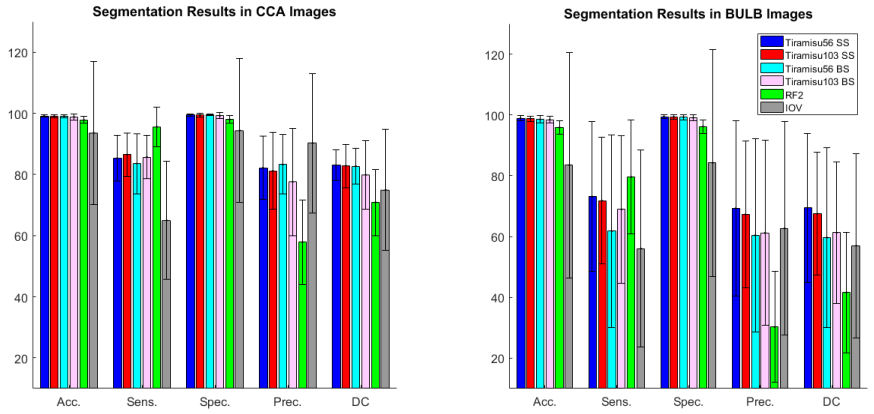

Fig. 3. Pixel-wise metrics results (mean and standard deviation) of different segmentation methods and IOV.

which a ROI is first extracted (pre-processing) and then RF (multi-class) is used for pixel-wise classification. Moreover, a post-processing specifically designed for this method [16] has been applied. When the post-processing is applied, we call the method RF2-PP.

All the NN models were initialized with HeUniform and trained with RMSprop, as in [5]. The training process was carried out in two steps: 1) pre-training with crop images (crop resolution: 224×224 px) for data augmentation, learning rate $1e-3$, and batch size 3; 2) fine-tuning with full size images (image resolution: 470×445 px), learning rate $1e-4$ and batch size 1.

We evaluated the performance with the following metrics. For segmentation results, the pixel-wise metrics Accuracy (Acc.), Specificity (Spec.), Sensitivity (Sens.), Precision (Prec.), and Dice Coefficient (DC). For CIMT estimation, correlation coefficient (cc), and we also applied a Bland-Altman analysis. For plaque detection, Acc., Spec., and Sens. as statistical measures.

3.3 Results

Segmentation Results. In Figure 3, we compare the segmentation approaches in CCA and Bulb test images. It can be seen that the different Tiramisu architectures drastically improve the RF2 results. Moreover, making the Tiramisu model deeper by increasing the number of parameters (from 56 to 103) does not improve the results. Note that although the BS is equivalent to SS in CCA images, the semantic information is crucial for the proper CIMT estimation in these images (see Section 2). Note that the improvement is more evident in Bulb images using SS. Regarding U-Net, its results are slightly worse than Tiramisu103 BS and are not included in the graphic. Finally, IOV result for DC is 74.95 in CCA images which is better than in Bulb images, where the DC is 56.97. These results

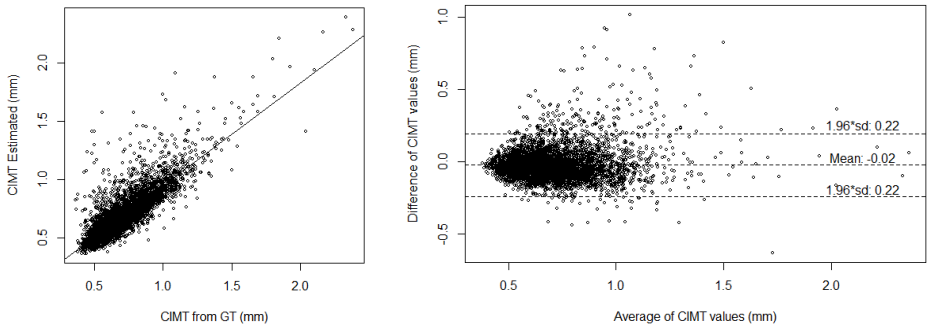

Fig. 4. Left: Correlation between CIMT values. Right: Bland-Altman analysis. Both plots show the relation between GT and the estimated values with our proposal.

and the high values in standard deviation show the difficulty in analyzing CA images, specially in Bulb. It is worth noting that all the measures have been computed using the GT of the Expert1, but the values are equivalent for the GT of the Expert2.

CIMT Measurement Results. Figure 4 (left) shows the correlation plot between the CIMT values (GT and predicted) in CCA images for the best method, "Tiramisu56 SS+CIMT estimation", which reaches a cc of 0.81. The result is very similar to Tiramisu103, cc of 0.80, in contrast to RF2-PP, which reaches a cc of 0.72. On the other hand, the cc obtained for Bulb images is quite low, 0.58. This fact is probably because of the small set of train images in Bulb and, at the same time, their low quality. However, our proposal, "Tiramisu56 SS+CIMT estimation", it improves RF2-PP, which only reaches a cc of 0.41. In Figure 4 (right), the Bland-Altman plot shows the difference, also in CCA images, between the CIMT of the corresponding two values ("Tiramisu56 SS+CIMT est." and GT) against the average of both values. This plot shows a high degree of agreement between the two measures, especially in the cases where the CIMT is small (<0.5mm) which corresponds to healthy population [3]. Furthermore, this plot shows that the predicted CIMT is, on average, slightly underestimated (mean -0.02).

Plaque Detection Results. Table 1 shows the plaque detection results in CCA and Bulb images. Our proposal, achieves promising results in CCA. The small number of plaques in the data set gives lower sensitivity values than specificity values. Results in Bulb images are still unsatisfactory (very low sensitivity), mainly due to the poor quality of these images, as we commented before.

Figure 5 shows qualitative examples of the CIM segmentation results, as well as plaque detection results in CCA and in Bulb images. The two first images, at the left (CCA images), are correctly segmented; the first without plaque

Images	Method	# Plaques/ Total images	Accuracy	Sensitivity	Specificity
CCA	RF2	50/4,722	50.05%	100.00%	49.00%
	RF2-PP	50/4,722	94.08%	86.00%	94.16%
	Our proposal	50/4,751	96.45%	80.00%	96.63%
Bulb	RF2	240/3,539	35.09%	98.33%	30.49%
	RF2-PP	240/3,539	78.50%	69.58%	79.15%
	Our proposal	262/3,733	93.25%	31.30%	97.95 %

Table 1. Results of Plaque Detection. RF2-PP stands for two-step RF2 with pots-processing.

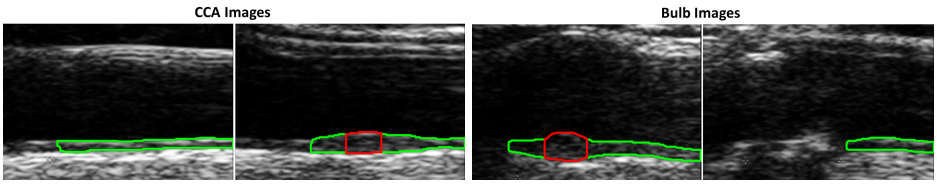

Fig. 5. Qualitative results of the CIM segmentation for 4 different images. Green lines are the CIM region boundaries and red lines the detected plaque boundaries. Images are cropped for visualization purpose.

and the second with plaque. The two last images, at the right (Bulb images), contain plaques. In the third image, the plaque is correctly segmented. But, in the fourth image, the plaque is not detected, probably due to the fact that the plaque composition produces shadows, which prevent a correct segmentation.

4 Conclusions and Future Work

In this paper, we have presented, for the first time in the literature, a single-step approach, based on DenseNets, for semantic Carotid Artery segmentation. The proposed method accurately localizes the CIM region in CCA. Given the segmentation, we have validated the CIMT estimation and the detection of atherosclerotic plaque with a large dataset of more than 8,000 images. We have compared the results obtained by the proposed method with those of other deep models and a shallow approach, demonstrating more accurate results of the segmentation, CIMT measurement and plaque detection. This superior performance is attributed to the effective use of semantic segmentation together with a CIMT estimation approach. In our future work, we will further improve the method results in the Bulb by increasing the number of bulb images in the train set.

Acknowledgments. Anonymous.

References

1. Bastida-Jumilla, M., Menchón-Lara, R., Morales-Sánchez, J., Verdú-Monedero, R., Larrey-Ruiz, J., Sancho-Gómez, J.: Frequency-domain active contours solution to

evaluate intimamedia thickness of the common carotid artery. Biomedical Signal Processing and Control 16(Complete), 68–79 (2015)
2. Carvalho, D., Akkus, Z., van den Oord, S., Schinkel, A., van der Steen, A., et al.: Lumen segmentation and motion estimation in b-mode and contrast-enhanced ultrasound images of the carotid artery in patients with atherosclerotic plaque. Medical Imaging, IEEE Transactions on 34(4), 983–993 (2015)
3. Grau, M., Subirana, I., Agis, D., Ramos, R., et al.: Grosor íntima-media carotídeo en población española: valores de referencia y asociación con los factores de riesgo cardiovascular. Revista Española de Cardiología 65(12), 1086–1093 (2012)
4. Huang, G., Liu, Z., Weinberger, K.Q., van der Maaten, L.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 4700–4708 (2017)
5. Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., Bengio, Y.: The One Hundred Layers Tiramisu: Fully convolutional DenseNets for Semantic Segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 1175–1183 (2017)
6. Loizou, C.P.: A review of ultrasound common carotid artery image and video segmentation techniques. Medical & Biological Engineering & Computing 52(12), 1073–1093 (Dec 2014)
7. Menchón-Lara, R.M., Sancho-Gmez, J.L.: Fully automatic segmentation of ultrasound common carotid artery images based on machine learning. Neurocomputing 151(Part 1), 161–167 (2015)
8. Molinari, F., Zeng, G., Suri, J.S.: A state of the art review on intima–media thickness (IMT) measurement and wall segmentation techniques for carotid ultrasound. Computer Methods and Programs in Biomedicine 100(3), 201–221 (2010)
9. Petroudi, S., Loizou, C., Pattichis, C.: Atherosclerotic carotid wall segmentation in ultrasound images using markov random fields. In: 10th IEEE International Conference on Information Technology and Applications in Biomedicine (2010)
10. Qian, C., Yang, X.: An integrated method for atherosclerotic carotid plaque segmentation in ultrasound image. Computer Methods and Programs in Biomedicine 153, 19 – 32 (2018)
11. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241 (2015)
12. Saba, L., Montisci, R., Famiglietti, L., Tallapally, N., Acharya, U.R., et al.: Automated Analysis of Intima-Media Thickness: Analysis and Performance of CARES 3.0. Journal of Ultrasound in Medicine 32(7), 1127–1135 (2013)
13. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39(4), 640–651 (2017)
14. Shin, J.Y., Tajbakhsh, N., Hurst, R.T., Kendall, C.B., Liang, J.: Automating carotid intima-media thickness video interpretation with convolutional neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2526–2535 (2016)
15. Touboul, P.J., Hennerici, M., Meairs, S., Adams, H., Amarenco, P., et al.: Mannheim carotid intima-media thickness and plaque consensus (2004-2006-2011). Cardiovascular Diseases 34(4), 290–296 (2012)
16. Zhang, C., Vila, M.M., Radeva, P., Elosua, R., Grau, M., Betriu, A., Fernandez-Giraldez, E., Igual, L.: Carotid artery segmentation in ultrasound images. In: MICCAI-Workshops on Computing and Visualization for Intravascular Imaging and Computer Assisted Stenting (CVII-STENT) (2015)