Computer Methods and Programs in Biomedicine 223 (2022) 106954

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

An end-to-end framework for intima media measurement and atherosclerotic plaque detection in the carotid artery

Lucas Gago[a,*], Maria del Mar Vila[a,b,c], Maria Grau[d,b,c], Beatriz Remeseiro[e,1], Laura Igual[a,1]

[a] Dept. de Matemátiques i Informática, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, Barcelona, 08007, Spain
[b] Dept. Epidemiologia i Salut Pública, IMIM, Institut Hospital del Mar d'Investigacions Médiques, Dr. Aiguader 88, Barcelona, 08003, Spain
[c] CIBER Enfermedades Cardiovasculares, Instituto de Salud Carlos III, Monforte de Lemos 3–5, Pabellón 11, Madrid, 28029, Spain
[d] Dept. de Medicina, Universitat de Barcelona, Carrer Casanova 143, Barcelona, 08036, Spain
[e] Dept. of Computer Science, Universidad de Oviedo, Campus de Gijón s/n, Gijón, 33203, Spain

ARTICLE INFO

Article history:
Received 28 February 2022
Revised 22 May 2022
Accepted 13 June 2022

Keywords:
Deep learning
Semantic segmentation
CIMT estimation
Atherosclerotic plaque

ABSTRACT

Background and objectives: The detection and delineation of atherosclerotic plaque are usually manually performed by medical experts on the carotid artery. Evidence suggests that this manual process is subject to errors and has a large variability between experts, equipment, and datasets. This paper proposes a robust end-to-end framework for automatic atherosclerotic plaque detection.

Methods: The proposed framework is composed of: (1) a semantic segmentation model based on U-Net, with EfficientNet as the backbone, that obtains a segmentation mask with the carotid intima-media region; and (2) a convolutional neural network designed using Bayesian optimization that simultaneously performs a regression to get the average and maximum carotid intima media thickness, and a classification to determine the presence of plaque.

Results: Our approach improves the state-of-the-art in both co and bulb territories in the REGICOR database, with more than 8000 images, while providing predictions in real-time. The correlation coefficient was 0.89 in the common carotid artery and 0.74 for bulb region, and the F1 score for atherosclerotic plaque detecting was 0.60 and 0.59, respectively. The experimentation carried out includes a comparison with other fully automatic methods for carotid intima media thickness estimation found in the literature. Additionally, we present an extensive experimental study to evaluate the robustness of our proposal, as well as its suitability and efficiency compared to different versions of the framework.

Conclusions: The proposed end-to-end framework significantly improves the automatic characterization of atherosclerotic plaque. The generation of the segmented mask can be helpful for practitioners since it allows them to evaluate and interpret the model's results by visual inspection. Furthermore, the proposed framework overcomes the limitations of previous research based on ad-hoc post-processing, which could lead to overestimations in the case of oblique forms of the carotid artery.

1. Introduction

The term atherosclerosis refers to a progressive disease characterized by the accumulation of lipids and fibrous substances in the large arteries [1]. Data from several studies suggest that this process can begin in early childhood [2] and worsens with age, while it can eventually lead to reduced blood flow through the affected vessel [3]. On top of that, atherosclerosis affecting the carotid artery (CA) is considered to be the main clinical manifestation of cardiovascular disease.

Cardiovascular disease (CVD) is the main cause of death in developed countries, and one of the leading causes of disease burden [4]. For these reasons, it is clinically essential to be able to accurately detect and mark plaque formation, thus allowing the progress of atherosclerosis to be controlled and monitored. Carotid intima-media thickness (CIMT), which estimates the width of the two deepest layers of arterial walls, is the most common sign of

* Corresponding author.
E-mail address: lgagogag69@alumnes.ub.edu (L. Gago).
[1] These authors jointly supervised this work.

https://doi.org/10.1016/j.cmpb.2022.106954

atherosclerosis development. CIMT and atherosclerotic plaque formation have also been shown to be a risk factor for stroke [5], coronary artery disease [6], and myocardial infarction [7]. Mean CIMT and maximum CIMT are both included in this study for their clinical interest. Firstly, according to previous studies, mean CIMT is associated with CVD risk factors [8]. Secondly, maximum CIMT is used to detect atherosclerotic plaque, which is the common basis of CVDs [9].

The procedure of detecting early atherosclerotic vascular diseases —CIMT estimation— through ultrasound imaging is a safe, non-invasive, and cost-effective method [9]. Carotid arterial wall assessment may include the common (CCA), internal, or bulb territories of the carotid artery. Measurements and monitoring of CIMT, atherosclerotic carotid plaque, and CA diameter are crucial and are regularly evaluated with high-resolution ultrasound images and videos. For the most part, CA delineations are manually performed by medical experts, but evidence suggests that they are subject to errors and have significant variability between different experts, equipment, and datasets [10]. Consequently, the availability of automatic methods for a robust and rigorous CIMT measurement and plaque delineation is highly desirable [11].

There is a growing interest in developing and implementing computer vision systems that can be integrated into real clinical practice. The most successful type of model used for computer vision tasks to date is convolutional neural networks (CNNs), which are made up of multiple layers of convolutional filters that progressively transform the input to extract some relevant features that are ultimately used to solve a learning task.

In this context, fully convolutional networks (FCNs) have proven effective in semantically segmenting different regions of the CA in ultrasound images [12]. More specifically, del Mar et al. [12] proposed a fully automatic method based on semantic segmentation for CIMT estimation and plaque detection. The main drawback of this method is that it applies an ad-hoc post-processing procedure after the semantic segmentation step, which needs domain knowledge and thus limits its generalization ability.

Our study aims to contribute to this growing area of research by exploring fully automatic methods based on semantic segmentation and CNNs. The objective is to define a robust end-to-end framework, without any prior knowledge and handcrafted algorithms, to assist medical practitioners in accurately determining CIMT and detecting plaque through ultrasound imaging.

1.1. Related work

Among the primary techniques for CIM region delineation we can find edge detection [13], active contours [14], and snakes [15]. More recently, machine learning [16] and deep learning approaches [12,17–19] have been proposed.

The interested reader is referred to [20] for an updated review study in which the methods are classified into three. The first-generation technologies were low-level segmentation approaches that employed traditional image processing techniques based on thresholding to get the lumen-intima and media-adventitia boundaries and then measured the mean distance using a caliper-based solution. The second generation used contour-based procedures that utilized parametric or geometric curves. Some of them were semi-automatic, requiring user interaction for initialization and/or correction of the results. In contrast, fully automatic methods do not require any user interaction, therefore being more scalable and reproducible. The third-generation models use artificial intelligence technologies such as machine learning and deep learning.

Table 1 presents a comparison of different fully automatic methods for CIMT estimation found in recent literature and includes some useful information such as the artery territories con-

Table 1
Comparison of fully automatic methods for CIMT estimation.

Article	Year	Method	Artery territory	No. of images	Mean CIMT error (mm)
Molinari et al. [13]	2012	Edge detection	CCA	365	0.078 ± 0.112
Menchón-Lara and Sancho-Gómez [17]	2014	Autoencoders for feature extraction	CCA	55	0.0499 ± 0.0489
Ikeda et al. [21]	2017	Bulb edge point detection	Bulb	649	0.0106 ± 0.0031
Qian and Yang [16]	2018	Patch-based machine learning	CCA	29	0.34 ± 0.10
Biswas et al. [18]	2018	Fully convolutional networks	CCA	396	0.126 ± 0.134
Biswas et al. [19]	2020	Patch classification and segmentation	CCA	250	0.093 ± 0.0637
del Mar et al. [12]	2020	FCN Semantic Segmentation with ad-hoc post-processing	CCA	4751	0.022 ± 0.1254
			Bulb	3733	0.06 ± 0.2749
Lian et al. [22]	2021	Poly-line estimation with DQN	CCA	4351	0.06 ± 0.04
Our proposal	2022	FCN Semantic Segmentation and CNN for CIMT estimation	CCA	4727	0.0058 ± 0.0902
			Bulb	3721	0.0096 ± 0.1791

sidered, the number of images used for evaluation purposes, and the mean CIMT error.

Molinari et al. [13] proposed an automatic procedure that uses edge detection and statistical classification for CA recognition and segmentation for IMT measurement. The experimentation, carried out on 365 CCA images, showed that their method underestimated the IMT values, despite outperforming previous approaches. Also focused on CCA territory, Menchón-Lara and Sancho-Gómez [17] presented a deep learning method based on artificial neural networks and autoencoders to identify CIMT boundaries. The proposed method was tested on a set of 55 longitudinal ultrasound images of CCA.

More recently, Ikeda et al. [21] proposed an automatic method based on carotid geometry and pixel classification to locate the bulb edges and used them for IMT measurement. The method was evaluated on 649 images with different types of plaque and image resolutions, achieving a coefficient of correlation of 0.998.

On the other hand, Qian and Yang [16] proposed a model that integrates random forest and an auto-context model for pixel-wise classification, reporting a dice similarity coefficient of 0.81 in only 29 ultrasound images.

Biswas et al. [18] presented a two-stage deep learning system composed of a convolution layer-based encoder for feature extraction and a fully convolutional network-based decoder for image segmentation. Further work by Biswas et al. [19] includes a segmentation procedure on image patches previously classified.

For their part, del Mar et al. [12] proposed a single-step approach for automatic CA image interpretation, able to be trained in both CCA and bulb territories. This method was composed of a semantic segmentation model and an ad-hoc post-processing module for CIMT estimation. It was tested on REGICOR database [8], which is composed of more than 8000 images. Their results reach a correlation coefficient of 0.81 in CIMT estimation, and a CIMT mean error of 0.02 and 0.06mm in CCA and bulb images, respectively. Performing a segmentation of the ultrasound image before CIMT estimation showed promising results. However, the ad-hoc post-processing applied to the segmentation mask, which includes predefined morphological operations and cropping off part of the image, may limit the robustness of the proposal.

Focusing on automatic estimation but with less amount of data, Lian et al. [22] used a deep Q-network to adjust 15 points to the near-wall, far-wall, and intima-lumen interface, to later adjust poly-lines. In addition, in the reward calculation step, they incorporated anatomical priors related to straightness and parallelism, increasing the performance but limiting the model's ability to detect clinically significant outliers.

Finally, it is worth noting that some of these studies [12,13] compute CIMT by dividing the CIM segmentation mask into vertical regions, which could lead to over-estimations in case of oblique forms of the CA [23].

1.2. Contributions

In this paper, we present an end-to-end framework for CIMT estimation and atherosclerotic plaque detection. The proposed framework is, in fact, a fully automatic system comprised of two modules: (1) a FCN for semantic segmentation and (2) a CNN for classification and regression. The segmentation model is composed of a light architecture with a pre-trained feature extractor as its backbone and aggressive data augmentation. The CNN takes as input both the original image and the mask provided by the segmentation module and generates a prediction of the average CIMT, the maximum CIMT, and the presence of atherosclerotic plaque.

The main contributions of this research are the following:

1. An end-to-end framework composed of two modules: (1) a semantic segmentation model, based on U-Net with EfficientNet as the backbone, achieving state-of-the-art results in CCA and bulb territories; and (2) a regression and classification model, based on a CNN designed using Bayesian optimization, capable of making real-time predictions of the maximum and average CIMT and the presence of atherosclerotic plaque in CCA and bulb territories.
2. An evaluation of the robustness of the proposed framework as well as its suitability and efficiency compared to different versions of it. In particular, a comparative study showing that training a model with a joint feature extractor for CIMT and plaque values can lead to better results than using three independent models.
3. A comparison with other fully automatic methods for CIMT estimation found in the literature, including a detailed comparison with a previous state-of-the-art approach using REGICOR database, with more than 8000 images corresponding to CCA and bulb territories.

The rest of the paper is structured as follows. Section 2 details the proposed end-to-end framework for CIMT estimation and plaque detection. Section 3 introduces the datasets used for evaluation purposes and the design of the experiments performed, followed by the results achieved. Finally, Section 4 closes the paper with the conclusions and future challenges.

2. Methodology

We propose a fully automatic framework that predicts both maximum and average CIMT, and the presence of atherosclerotic plaque without any domain knowledge or metadata from the image. Fig. 1 depicts the architecture of the proposed framework, which is comprised of a semantic segmentation network followed by a CNN that performs classification and regression tasks. A single CNN is used to simultaneously predict three target values, assuming this could lead to a better feature extractor and better performance on the plaque detection problem (see Section 3.4.3 for the validation of this assumption). The input size is 445x470 pixels, which corresponds to the original resolution of the REGICOR database. The two modules of the framework are subsequently described in depth.

2.1. Semantic segmentation

Carotid artery semantic segmentation consists of classifying each pixel of the input ultrasound image as one of lumen, far wall, near wall, CIM, bulb, and CIM-bulb region. Fig. 2 shows an example for each territory, with a legend that details the segmentation labels. The semantic segmentation changes the representation of the image into something more meaningful and easy to analyze.

A U-Net [24] was used to perform the carotid artery semantic segmentation. U-Net is an asymmetrical segmentation network that presents skip connections between down-sampling and up-sampling paths to improve the quality of the segmentation mask by providing local information to the encoded global information in the up-sampling process. The network is composed of three parts: the down-sampling, bottleneck, and up-sampling. Down-sampling is consists of four blocks containing 3×3 convolutional layers with batch normalization [25] followed by 2×2 max-pooling layers. At the end of each block, a skip connection is sent to the symmetric up-sampling module. The bottleneck is built from two convolutional layers with batch normalization and dropout [26] to reduce the overfitting. The up-sampling path also consists of four blocks, made

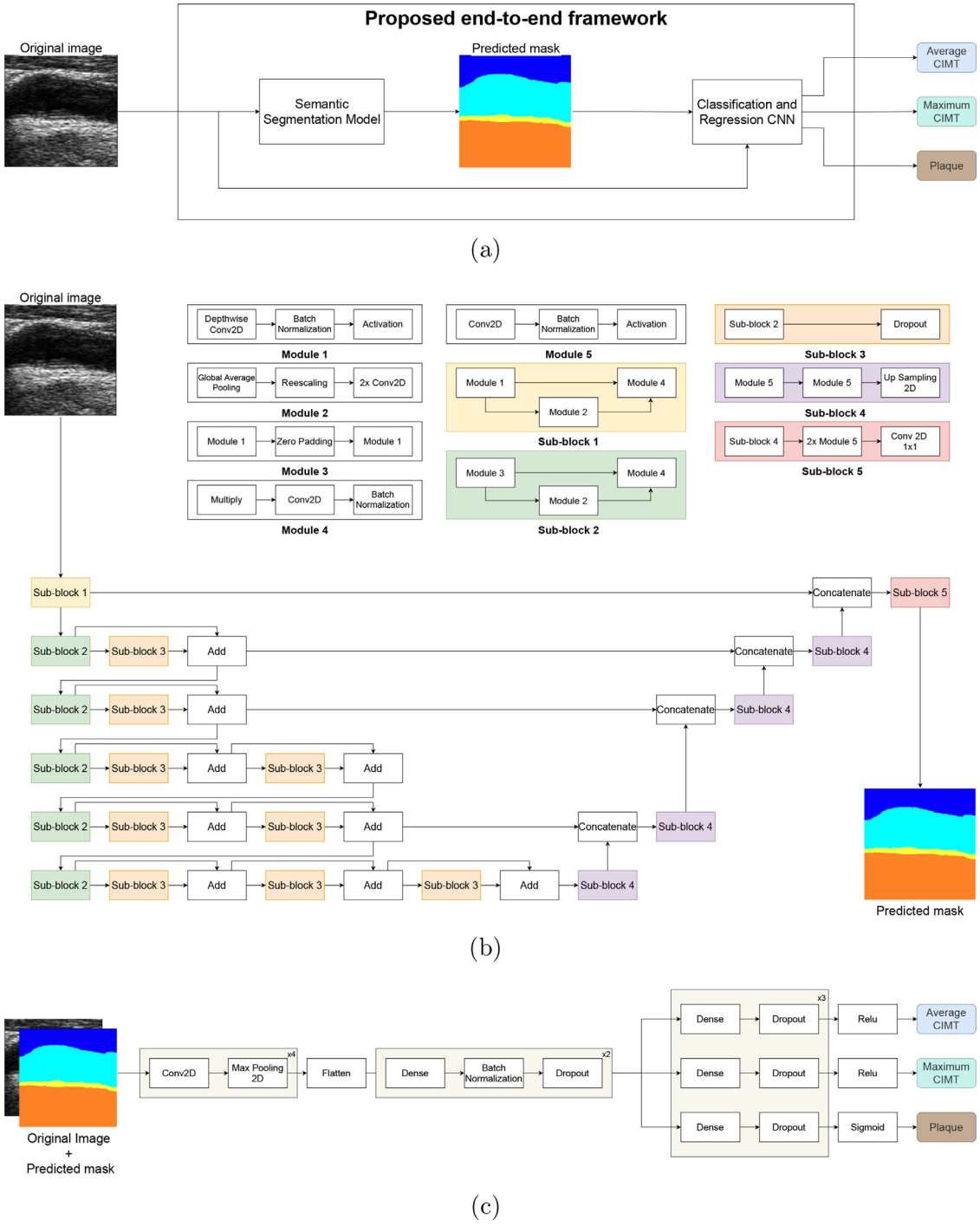

Fig. 1. (a) The proposed end-to-end framework composed of two modules: (b) the semantic segmentation model, and (c) the classification and regression CNN model architecture defined by Bayesian optimization.

up of transposed convolutions with stride 2, a concatenation with the corresponding feature map from down-sampling (skip connection), and 3×3 convolutional layers with batch normalization.

In our proposed architecture, we use EfficientNet B0 [27] as a lightweight feature extractor, pre-trained on ImageNet [28]. In particular, we replaced the down-sampling component of the U-Net with a pre-trained EfficientNet B0, while the bottleneck and up-sampling maintain the original U-Net architecture. Skip connections are sent from the first, second, third, and fifth blocks of EfficientNet B0, while the output is connected to the bottleneck part of U-Net.

The EfficientNet family of networks was generated using neural architecture search and has been proven capable of achieving high accuracy despite being much smaller and faster than previous models. EfficientNet B0 is the smallest architecture of the ones proposed, achieving 93.5% top-5 accuracy on the ImageNet validation set with only 5.3 million parameters. The EfficientNet networks are composed of different combinations of the sub-blocks from one to five, defined in Fig. 1(a).

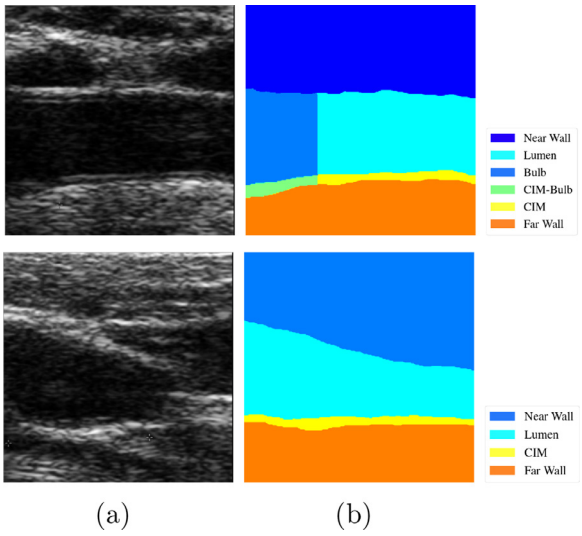

Fig. 2. (a) Two ultrasound images, and (b) their corresponding segmentation masks from the ground truth. The top row corresponds to the CCA territory (six labels), whilst the bottom one corresponds to the bulb (four labels).

Light architectures such as EfficientNet B0 helps to avoid overfitting and reduce training times. Moreover, data augmentation techniques are useful to prevent the overfitting caused by training models with a low number of samples. For this reason, we applied several data augmentation techniques, including scale transformations, elastic transformations, perspective transformations, affine transformations, JPEG compression, speckle noise, motion blur, variation of hue and saturation, histogram equalization, cropping, and padding. The motivation behind the design of the data augmentation pipeline was to improve the generalization capacity of the system and make it more robust to ultrasound imaging artifacts like speckle [29].

Regarding the loss function of the semantic segmentation network, we defined a custom loss (Loss) composed of the Dice coefficient (DC) and the focal loss (FL) [30], defined as follows:

$$\text{Loss} = \text{DC} + \text{FL} \tag{1}$$

$$\text{DC} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{2}$$

$$\text{FL}(p_t) = -\alpha_t (1 - p_t)^{\gamma} \log(p_t) \tag{3}$$

$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise} \end{cases} \tag{4}$$

where TP, FP, and FN stand for true positives, false positives, and false negatives, respectively; $p \in [0, 1]$ is the models estimated probability for the class with label y, p_t is defined for notational convenience, α is a weighting factor, and γ is the focusing parameter Notice that the focal loss component has been proven useful in segmentation problems with under-represented classes [30], as is our case with the CIM region class significantly less represented than the other classes.

2.2. CIMT estimation and plaque detection

We propose a regression and classification CNN model to predict average and maximum CIMT values as well as the presence of plaque in ultrasound images. Previous research relies on ad-hoc post-processing of the segmentation results [12], limiting the ability of the system to generalize. Keeping that in mind, our motivation is to eliminate any prior knowledge and handcrafted algorithms from the pipeline, making it fully automatic and able to perform well without depending on the origin, scale, or quality of the input data. The combination between the semantic segmentation model and CNN for CIMT estimation and plaque detection (see Fig. 1) allows us to have an end-to-end fully automatic framework that can be trained without any domain-specific knowledge given tagged examples.

The CNN input is the concatenation of the original image and the segmentation mask provided by the segmentation model previously applied, with the idea of maximizing the information available to the model when predicting the presence of atherosclerotic plaque. Regarding the CNN architecture, it was tuned using a Bayesian optimization [31] given some design constraints. The system selects a variable number of convolutional layers followed by a max-pooling operation. Then, the last layer is flattened and the optimization selects a series of configurable dense layers followed by batch normalization and dropout, with a rate of 25%. At this point, the network is divided into three branches to predict one target value per branch (average CIMT, maximum CIMT, plaque detection), made up of a variable number of blocks of dense layers followed by dropout (see Fig. 1c).

The activation function used for average and maximum CIMT estimation is the rectified linear unit (ReLU) [32], and the sigmoid for plaque detection. As for the loss functions, we used the mean squared error (MSE) for average and maximum CIMT and binary cross-entropy (BCE) for plaque detection, defined as:

$$\text{MSE} = \frac{\sum_{i=1}^{n} \left(\hat{y}_i - y_i\right)^2}{n} \tag{5}$$

$$\text{BCE} = -\frac{1}{n} \sum_{i=1}^{n} y_i \cdot \log \hat{y}_i + (1 - y_i) \cdot \log\left(1 - \hat{y}_i\right) \tag{6}$$

where y_i is the truth value, and $\hat{y}_i$ is the predicted value.

3. Experimental study

This section introduces datasets used for evaluation purposes and implementation details of the proposed framework. Moreover, experiments carried out and the results achieved are presented and discussed, including a comparison with the state-of-the-art in REGICOR dataset.

3.1. Datasets

The proposed framework was evaluated on the REGICOR database, one of the largest image collections available for the problem at hand [12]. REGICOR consists of a sample of 2379 subjects from Girona's Heart Registry [8]. Images were collected from 2007 to 2010, and the subjects represent general population aged 35 to 84. Two trained sonographers performed the CA ultrasound (US) scans with an Acuson XP128 US system equipped with an L75-10 MHz transducer and a computer program extended frequency (Siemens-Acuson). US longitudinal images were obtained in B-mode with a resolution of 23.5 pixels/mm. Original images were saved in DICOM format and then converted to PNG. The set of images collected for each patient was obtained from left and right CA in two different territories (CCA and bulb), resulting in a total of 8448 images (4727 CCA images, and 3721 bulb images). CIMT reference values, given by the Amsterdam Medical Center, were used as the ground-truth (GT) for CIMT estimation. Regarding the GT for plaque detection, it was obtained using the provided CIMT reference values and applying the Mannheim consensus [9].

Table 2
Summary of the different datasets of US images used in this research.

Dataset	Territory	Information available	No. images	Images with plaque
REGICOR CCA-Seg	CCA	Manually segmented masks	159	50 (31.47%)
REGICOR Bulb-Seg	Bulb	Manually segmented masks	172	68 (39.53%)
REGICOR CCA	CCA	Avg. and max. CIMT and plaque	4727	50 (1.06%)
REGICOR Bulb	Bulb	Avg. and max. CIMT and plaque	3721	262 (7.04%)

Furthermore, images containing plaque were finally supervised by an expert.

Besides the GT for CIMT estimation and plaque detection, a segmentation GT was defined for a subset of REGICOR images [12]. In order to obtain it, an expert manually delineated and labeled different regions of original images, using six labels for CCA and four for the bulb (see Fig. 2). Only a representative subset of REGICOR images was labeled, including 159 CCA images (50 with plaque and 109 without plaque), and 172 bulb images (68 with plaque and 104 without). These labeled subsets will be referred to as REGICOR CCA-Seg and REGICOR Bulb-Seg from now on, respectively. Table 2 presents a summary of the datasets and their main characteristics.

3.2. Implementation details

Our proposed framework is implemented on Tensorflow [33], and the code will be publicly available after paper acceptance[2]. It is made up of two modules, the semantic segmentation model and the regression and classification CNN, the implementation of which is given below.

3.2.1. Semantic segmentation model

For the semantic segmentation module, we used a U-Net network with EfficientNet pre-trained on ImageNet as the backbone (see Section 2.1). The model is composed of 10.1 million parameters and was trained to segment all available classes, even though we are mainly interested in CIMT region. The reason is that models achieve better performance in CIMT class after being trained with all the classes, as demonstrated in [12]. For the custom loss, the weighting factor α was set to 0.25 and the focusing parameter γ to 2.0. Aggressive data augmentation was performed, ensuring that the image scale is not altered or no transformation is performed that could lead to information loss.

We used Adam [34] as the optimizer, a gradient descent method that is based on adaptive estimation of first-order and second-order moments. The parameters β_1 and β_2, representing the exponential decay rate for the first and second-moment estimates, were set to 0.9 and 0.999, respectively.

Note that early stopping was used with a patience of 20, and the learning rate was reduced on plateau with a patience of 10 and a factor of 0.1. The initial learning rate was 1e−4. The models were trained with full resolution images and a batch size of 4.

3.2.2. Regression and classification CNN

The regression and classification CNN model defined in Section 2.2 takes the mask predicted by the segmentation network and the original image as input and uses them to predict three outputs per image: average CIMT, maximum CIMT, and plaque. Notice that the three outputs share the feature extraction part of the CNN, which was designed using a Bayesian optimization tuned with a Gaussian process trained on REGICOR CCA dataset. The motivation here is that adding more information could lead to an improvement in performance, offsetting the additional calculation.

In this case, we also used Adam as the optimizer, with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The loss for the CNN was weighted, focusing more on the average and maximum CIMT outputs. This was motivated by the convergence problems for plaque prediction in REGICOR CCA dataset due to the class imbalance. Weights were set to 0.4 for average and maximum CIMT outputs, and 0.2 for plaque. In the plaque classification problem, classes were not weighted because, in our preliminary experiments, the results did not improve.

Early stopping was used with a patience of 50, and the learning rate was reduced on plateau with a patience of 15 and a factor of 0.1. Initial learning rate was 1e−3. The models were trained with full resolution images and a batch size of 16. It should be noted that mixed-precision was used to accelerate the training process and reduce the memory footprint.

3.3. Performance measures

For the evaluation of the semantic segmentation module, we computed a standard metric in this type of problem: the intersection over union (IoU). This metric measures the number of pixels in common between the target and prediction segmentation masks divided by the total number of pixels present across both masks:

$$\text{IoU} = \frac{\text{target} \cap \text{prediction}}{\text{target} \cup \text{prediction}} \tag{7}$$

Precision and sensitivity were also used as performance metrics:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

where TP, FP, and FN stand for true positives, false positives, and false negatives, respectively. Notice that, for evaluation purposes, the CIM region is considered positive and the background (i.e., the combination of all other classes) negative, since the CIM region is our main focus.

To evaluate the performance of the method in predicting CIMT, we used the Pearson correlation coefficient (CC), the mean absolute error (MAE), and the mean squared error (MSE). Pearson correlation coefficient is defined as:

$$\text{CC}_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} \tag{10}$$

where X and Y are a pair of random variables, cov is the covariance, and σ_X and σ_Y are standard deviations of X and Y, respectively.

The MAE is defined as follows:

$$\text{MAE} = \frac{\sum_{i=1}^{n} \left| \hat{y}_i - y_i \right|}{n} \tag{11}$$

where y_i is the truth value and $\hat{y}_i$ is the prediction. See Eq. (5) for the definition of the MSE.

For the plaque classification task, in addition to the sensitivity defined in Eq. (8), we also computed the accuracy, specificity, and F1 Score:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{12}$$

[2] https://github.com/gagolucasm/DL_CIMT_and_plaque_estimation

Table 3
Summary of the experiments designed for evaluation purposes.

Experiments	Datasets	Train split	Performance metrics
1) Evaluation the of semantic segmentation module	REGICOR CCA-Seg REGICOR Bulb-Seg	90% train 10% test	DICE, IoU, Precision, Sensitivity
	REGICOR CCA REGICOR Bulb	100% test	Accuracy, Sensitivity, Specificity, F1 Score
2) Evaluation of the impact of input data in the regression and classification CNN	REGICOR CCA REGICOR Bulb	60% train 20% val 20% test	MAE, MSE, CC, Accuracy, Sensitivity, Specificity, F1 Score
3) Evaluation of one versus three CNNs for plaque detection	REGICOR CCA REGICOR Bulb	60% train 20% val 20% test	Accuracy, Sensitivity, Specificity, F1 Score

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{13}$$

$$\text{F1 Score} = \frac{TP}{TP + \frac{1}{2}(FP + FN))} \tag{14}$$

Note that, for these measures, the presence of plaque is considered positive and its absence, negative.

3.4. Experimental results

Three experiments were defined to evaluate the robustness of the proposed end-to-end network.

- Experiment 1: Evaluation of the semantic segmentation module. The objective is to measure the performance of the proposed module to segment different regions of CA and to analyze its suitability for plaque detection in comparison with previous research.
- Experiment 2: Evaluation of the impact of input data in the regression and classification CNN. The objective is to analyze whether the additional information and computation of the semantic segmentation model are necessary, and the combined use of the predicted segmentation mask and the original image as input data.
- Experiment 3: Evaluation of one versus three CNNs for plaque detection. The objective is to gain insight into its effect on the overall performance of training a single model or three independent models to predict the three target values (average CIMT, maximum CIMT, and plaque).

Table 3 summarizes the three experiments carried out, including datasets and performance measures used. Note that all the models were trained with an NVIDIA GeForce RTX 2080ti 11GB GPU, and the code to reproduce all the experiments will be publicly available with the implementation of the framework (see Section 3.2).

3.4.1. Experiment 1: Evaluation of the semantic segmentation module

Experiment 1 was divided into two parts: (1) a quantitative and qualitative analysis of the segmentation results on the two databases considered (REGICOR CCA-Seg and Bulb-Seg), and (2) a comparison of results for CIMT estimation and plaque detection on REGICOR CCA and Bulb full datasets using the post-processing proposed in [12].

Table 4 shows the results obtained with the two methods, the segmentation network used in [12] and our segmentation module, on both datasets (REGICOR CCA-Seg and Bulb-Seg). As can be observed, our model outperforms the one proposed by del Mar et al. [12], regardless of the metric considered. More specifically, our model performs better on the CCA-Seg dataset, but there is a more significant improvement over previous results on the Bulb-Seg dataset. In this sense, it is worth mentioning that image quality in the bulb region is lower, with poorer contrast and more affected by noise.

Figure 3 depicts some representative examples of predictions obtained with the two methods, for a qualitative comparison. As can be seen, our proposal shows better region connectivity and does not generate erroneous isolated regions. Prediction time is 0.026 seconds, an order of magnitude faster than [12] (0.79s) due to the smaller network size. Note that time per frame was measured using a GeForce Titan X (Pascal) 12GB GPU from NVIDIA in [12], while we used a GeForce RTX 2080ti 11GB GPU also from NVIDIA.

Semantic segmentation models provide a segmentation mask categorizing each pixel into a class. Based on this information, del Mar et al. [12] proposed an ad-hoc post-processing procedure for CIMT estimation. In this experiment, we applied the same procedure to the results obtained with our semantic segmentation model to measure their impact on plaque detection. The post-processing procedure is based on morphological operators and prior domain knowledge and is detailed in [12].

Table 5 shows the results for plaque detection compared with the most competitive ones reported so far for REGICOR datasets [12]. Its important to mention that the class plaque is underrepresented in both datasets, with 1.06% of total images in REGICOR CCA and 7.04% in REGICOR Bulb. As can be seen, the results obtained with the proposed method are consistently better than those obtained in [12] regardless of the metric applied, thus confirming the superior performance of our proposal not only in terms of segmentation but also in plaque detection.

3.4.2. Experiment 2: Evaluation of the impact of input data in the regression and classification CNN

For the second experiment, we conducted an ablation study in which we eliminate the segmentation module from our input to determine the contribution of the component to the overall system. The regression and classification CNN was fed with different input data: (1) only the original image, (2) only the predicted segmentation mask, or (3) the concatenation of both (our proposal). Achieving similar results with only the original image as input data would mean that we could further reduce the complexity of the system, with the disadvantage of decreasing the interpretability of the model.

Table 6 shows the results obtained in this experiment for maximum and average CIMT values. Here, we can appreciate a clear improvement over previous work [12] in every metric and for both datasets, all with a p-value of < 0.001. Using both the original image and the segmentation mask, our proposal, translates into a consistent improvement of results with respect to CNN versions from segmentation mask only and from original image only. Table 6 contains information about the mean average error (MAE), mean squared error (MSE), and correlation coefficient (CC) for the maximum and average CIMT prediction. This model performs significantly better in CCA than in bulb, where the qual-

Table 4
Results for semantic segmentation on REGICOR CCA-Seg and Bulb-Seg test sets. Only the CIMT region and background are considered. The best results per dataset are marked in bold face.

Dataset	Method	IoU	Dice	Precision	Sensitivity
REGICOR	del Mar et al. [12] seg.	0.7121	0.8299	0.8217	0.8530
CCA-Seg	Our seg. proposal	**0.8712**	**0.9959**	**0.9976**	**0.9942**
REGICOR	del Mar et al. [12] seg.	0.5711	0.6945	0.6930	0.7313
Bulb-Seg	Our seg. proposal	**0.9273**	**0.9694**	**0.9741**	**0.9721**

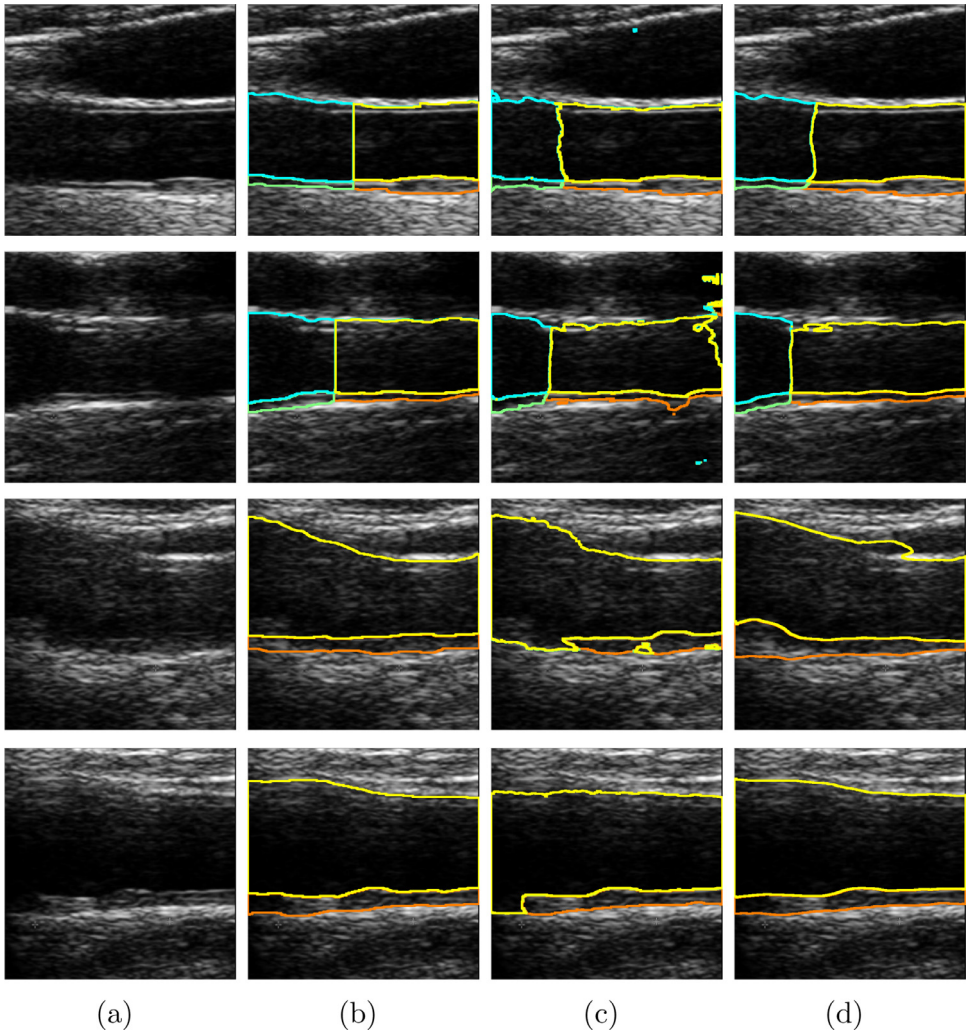

Fig. 3. Qualitative results: (a) Four original images, (b) the ground truth, (c) the segmentation results obtained in [12], and (d) the segmentation results achieved with our proposed method. The first two rows correspond to the REGICOR CCA-Seg, whilst the last two are from REGICOR Bulb-Seg.

Table 5
Results for plaque detection on REGICOR CCA and Bulb full datasets using the post-processing proposed in [12]. The best results per dataset are marked in bold face.

Dataset	Method	Plaques/total images	Accuracy	Sensitivity	Specificity	F1 Score
REGICOR	del Mar et al. [12] seg.	50/4727	0.9645	0.8000	0.9663	0.3226
CCA	Our seg. proposal	50/4727	**0.9725**	**0.8800**	**0.9735**	**0.4037**
REGICOR	del Mar et al. [12] seg.	262/3721	0.7809	0.7832	0.7500	0.3262
Bulb	Our seg. proposal	262/3721	**0.8126**	**0.9014**	**0.8054**	**0.4197**

ity of the segmentation is lower. In CCA, it obtains best results in every measured metric, showing a clear use of the additional information provided in the input. In REGICOR Bulb, the CNN from segmentation mask performs best at the maximum estimate of CIMT, while our proposal outperforms in predicting average CIMT.

Results obtained when applying the post-processing presented in [12] to our segmentation masks (Experiment 1) can be considered a more demanding baseline and are still lower on each metric, indicating that the performance increase cannot be explained solely by new segmentations. There seems to be a positive impact from using a CNN for regression and classification. It is worth

Table 6
Results for maximum and average CIMT on REGICOR CCA and Bulb tests sets. The best results per dataset are marked in bold face.

				Maximum CIMT			Average CIMT		
Dataset	Exp.	Input data	Method	MAE (mm)	MSE (mm^2)	CC	MAE (mm)	MSE (mm^2)	CC
REGICOR	-	Seg. mask [12]	del Mar et al. [12] post-processing	0.1539	0.0633	0.6213	0.0906	0.0170	0.8006
CCA	1	Our seg. mask	Vila et al. [12] post-processing	0.1330	0.0477	0.6960	0.0824	0.0159	0.8312
(test	2	Our seg. mask	One single CNN	0.0921	0.0164	0.8367	0.0665	0.0089	0.8820
set)	2	Original image	One single CNN	0.1105	0.0266	0.6896	0.0818	0.0162	0.7519
	2	Orig. image + Our seg. mask	One single CNN *	**0.0896**	**0.0148**	**0.8431**	**0.0659**	**0.0082**	**0.8870**
REGICOR	-	Seg. mask [12]	del Mar et al. [12] post-processing	0.4460	0.4252	0.2673	0.1936	0.0806	0.3899
Bulb	1	Our seg. mask	Vila et al. [12] post-processing	0.3700	0.2833	0.5071	0.1901	0.0650	0.6111
(test	2	Our seg. mask	One single CNN	**0.1646**	**0.0589**	**0.7225**	0.1328	0.0335	0.7279
set)	2	Original image	One single CNN	0.2267	0.1038	0.4048	0.1781	0.0568	0.4056
	2	Orig. image + Our seg. mask	One single CNN *	0.1669	0.0595	0.7228	**0.1311**	**0.0321**	**0.7362**

* Our proposal

Table 7
Results for atherosclerotic plaque detection on REGICOR CCA and Bulb test sets. The best results per dataset are marked in bold face.

Dataset	Exp.	Input data	Method	Accuracy	Sensitivity	Specificity	F1 Score
REGICOR	-	Seg. mask [12]	del Mar et al. [12] post-processing	0.9651	**0.8571**	0.9659	0.2667
CCA	1	Our seg. mask	Vila et al. [12] post-processing	0.9725	**0.8571**	0.9733	0.3158
(test	2	Our seg. mask	One single CNN	0.9926	0.5714	0.9957	0.5333
set)	2	Original image	One single CNN	0.9884	0.0000	0.9957	0.0000
	2	Orig. image + Our seg. mask	One single CNN *	0.9915	**0.8571**	0.9925	**0.6000**
	3	Our seg. mask	Three CNNs (only plaque output)	**0.9947**	0.5714	0.9979	**0.6000**
	3	Original image	Three CNNs (only plaque output)	0.9926	0.0000	**1.0000**	0.0000
	3	Orig. image + Our seg. mask	Three CNNs (only plaque output)	0.9905	0.5714	0.9936	0.4706
REGICOR	-	Seg. mask [12]	del Mar et al. [12] post-processing	0.7903	0.7551	0.7928	0.3217
Bulb	1	Our seg. mask	Vila et al. [12] post-processing	0.8427	**0.9184**	0.8374	0.4348
(test	2	Our seg. mask	One single CNN	0.9476	0.4898	0.9799	0.5517
set)	2	Original image	One single CNN	0.9301	0.2245	0.9799	0.2973
	2	Orig. image + Our seg. mask	One single CNN *	**0.9516**	0.5306	0.9813	**0.5909**
	3	Our seg. mask	Three CNNs (only plaque output)	0.9422	0.3673	0.9827	0.4557
	3	Original image	Three CNNs (only plaque output)	0.9368	0.0408	**1.0000**	0.0784
	3	Orig. image + Our seg. mask	Three CNNs (only plaque output)	0.9341	0.0000	**1.0000**	0.0000

* Our proposal

noting that CNN architecture was optimized for REGICOR CCA, and this could show penalizing results in REGICOR Bulb.

Table 7 shows the results for plaque classification in the test set for all the experiments. Our proposal performs significantly better on REGICOR Bulb, achieving an F1 score of 0.5909, indicating a balance between sensitivity and precision not exhibited by other experiments. On REGICOR CCA, it keeps the sensitivity level of both the Vila et al. [12] proposal and [12] post-processing applied to our segmentation masks, but with a significant increase in precision, reflected in the F1 score of 0.6000, up from 0.2667 and 0.3158 from these previous experiments. Using only the original image as input results in a very low sensitivity in the plaque prediction problem, meaning that the system cannot correctly predict images with high CIMT values, underestimating the plaque prediction. Using the segmentation mask as the only input achieves a lower performance than our proposal, which indicates that the additional information provided by the concatenation of the original image and segmentation mask has a positive impact on the plaque detection problem. In REGICOR CCA, while the F1 score is 0.6000 using only the segmentation mask and using the concatenation of the segmentation mask and the original image, the sensitivity of the latter is higher, which is 0.8571 compared to 0.5714.

Figure 4 includes a representation of Bland-Altman and scatter plots for the results of our proposal on the test sets of both REGICOR CCA and REGICOR Bulb. The model performs well in REGICOR CCA, with a mean error of -0.02mm. The variability seems to be consistent without any appreciable trend between CIMT values of 0.7 and 0.9mm. Predictions of average CIMT in images with values between 0.45 and 0.7mm are invariably near 0.55mm, indicating the model has difficulties in analyzing cases in this range. As for REGICOR Bulb, there is a trend towards underestimating CIMT as it gets higher, showing difficulties in the prediction of outliers. The mean error is -0.04mm, indicating a slight underestimation.

The average processing time of this CNN block is 0.014 seconds, adding to 0.040 seconds if the segmentation step is considered; that is, the processing time is almost 20 times faster than in the previous work [12].

3.4.3. Experiment 3: Evaluation of one versus three CNNs for plaque detection

We conducted a final experiment to assess whether using three CNN models, one for each target value, offers an advantage over a single CNN for the three target values (our proposal). CNN architectures used for this experiment are the same as we used in our proposal, except having one output line instead of three (see Fig. 5). As for the training settings, they are the same as in Experiment 2.

Table 7 reports the results obtained from plaque classification using three independent CNNs, denoted as “only plaque output”. Results suggest that our proposal provides better results than using three individual CNNs. On REGICOR CCA, the models trained only on plaque and only to predict the maximum CIMT cannot classify any image in the plaque category, whereas if they are trained together, they can. The same situation occurs in REGICOR Bulb dataset: the results of our model with multiple outputs and one feature extractor are better for plaque detection than using three independent CNNs. The impossibility of detecting the presence of atherosclerotic plaque with this architecture, as well as the need to train and predict three independent models, make this solution a poor option for the problem at hand, thus demonstrating the suitability of our proposed end-to-end framework.

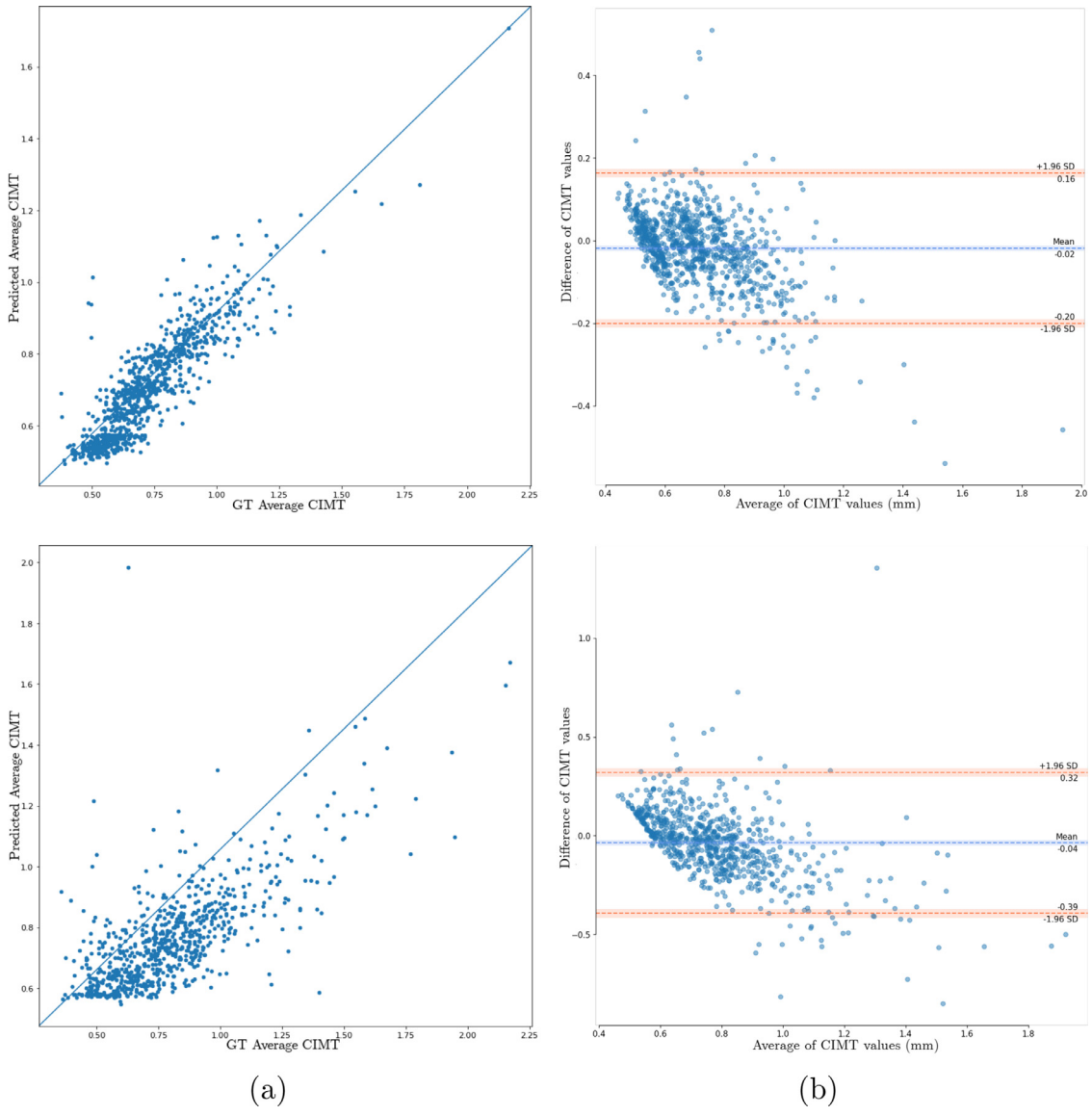

Fig. 4. Results of average CIMT predictions obtained with our proposal (Experiment 3). (a) Correlation between average CIMT values, and (b) BlandAltman analysis for the predicted average CIMT values. Note that the top row corresponds to the REGICOR CCA dataset, whilst the bottom row is for REGICOR Bulb.

4. Conclusions and future work

The intima-media thickness and the presence of atherosclerotic plaque in the carotid artery are the most common signs of cardiovascular disease development. In this context, we present an end-to-end framework to predict average CIMT, maximum CIMT, and presence of plaque on ultrasound images of two different carotid artery territories (CCA and bulb). Our approach is composed of a semantic segmentation module to anatomically segment the input image, followed by a CNN for regression and classification of three target values (average CIMT, maximum CIMT, and plaque) fed with the original image and the predicted segmentation mask. This approach can be useful for practitioners since it allows them to evaluate and interpret the results of the model by visually inspecting the predicted segmentation masks. Moreover, the method is able to estimate CIMT in a fast and useful manner for large image datasets and enables us to eliminate the inter-observer variability usually associated with manual CIMT estimation. The proposed framework achieves state-of-the-art results in REGICOR database and reduces prediction time from 0.79 to 0.04 seconds per image, with a processing speed of 25 frames per second. We compared the semantic segmentation model with previous work, qualitatively and quantitatively, demonstrating more accurate results. Moreover, our experiments also confirm the improvement in terms of CIMT prediction and atherosclerotic plaque detection on 8290 images. Furthermore, the proposed framework overcomes the limitations of previous research [12], based on ad-hoc post-processing that computed CIMT by dividing the mask into vertical regions, which could lead to over-estimations in case of oblique forms of the CA [23]. Instead, we proposed a fully automatic method concatenating two NN models with no need for domain knowledge, or tuning, in the dataset considered.

We also performed an ablation study, eliminating the segmentation module of the proposed framework and finding the need to use the information provided by it to achieve accurate results. Additionally, a study was performed in order to gain insight into the effect on the overall performance of training a single CNN model or three independent networks to predict three target values. The

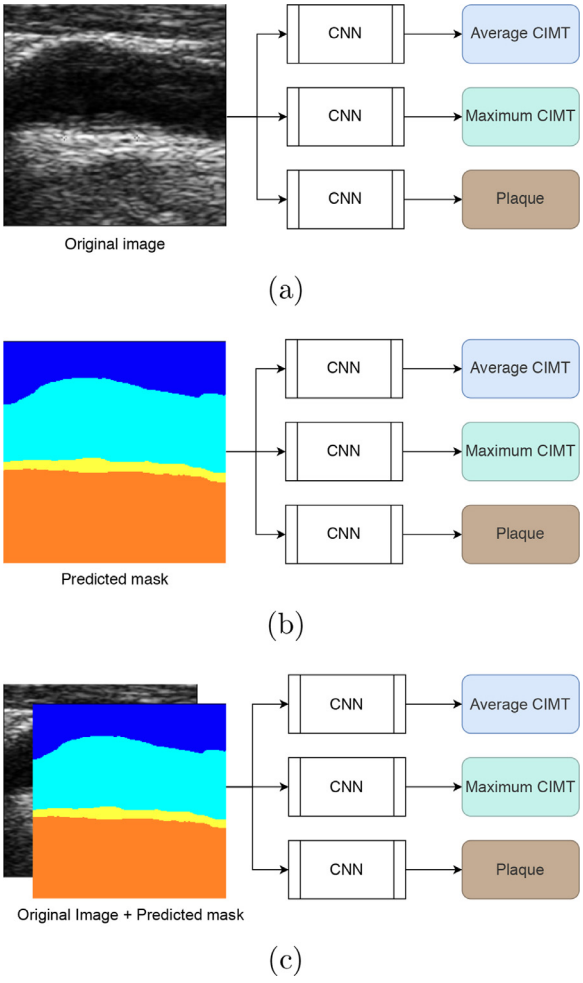

Fig. 5. Models tested to analyze if training a single CNN with multiple outputs leads to better detection of plaque than using three individual CNNs (Experiment 3). Different input data for the CNNs: (a) the original image, (b) the segmentation mask obtained with our proposed model, and (c) a concatenation of the original image and the predicted segmentation mask. Note that the CNN blocks are the same as in our proposal, but with one single output.

results conclusively show that the original idea of training a single CNN with multiple outputs leads to better results for atherosclerotic plaque detection.

Regarding the limitations of this study, highly unbalanced datasets with a small number of plaque images can be a problem in achieving results comparable to our proposal for the single CNN that uses only the original image as input. Additionally, in any medical application, explainability may be required. While the segmentation module of our proposal provides relevant information to the user, the classification and regression module is a black-box.

Our future research includes a study on other datasets to further confirm our conclusions, mainly in terms of generalization power. Therefore, we plan to revise this research work when more data from different institutions and different acquisition systems become available. This research paves the way for a fully automated evaluation of CIMT and plaque. We aim to make a more data from different institutions and different acquisition systems become available.

Furthermore, the method presented is a baseline framework to integrate information from the image and clinical data, which are very relevant to assess other pathologies or events related to atherosclerosis, such as cardiovascular risk. Moreover, since the output of the framework is for regression or classification targets, it can be easily adapted for medical purposes as cardiovascular risk prediction or risk stratification. In our future work, we intend to explore this field as another potential line of research.

Declaration of Competing Interest

No potential competing interest was reported by the authors.

Acknowledgements

This work was supported in part by the MICINN Grant RTI2018-095232-B-C21 and 2017 SGR 1742.

References

[1] A.J. Lusis, Atherosclerosis, Nature 407 (2000) 233–241.
[2] Y.M. Hong, Atherosclerotic cardiovascular disease beginning in childhood, Korean Circ J 40 (1) (2010) 1–9.
[3] T. Head, S. Daunert, P.J. Goldschmidt-Clermont, The aging risk and atherosclerosis: a fresh look at arterial homeostasis, Front Genet 8 (2017) 216.
[4] T.A. Gaziano, A. Bitton, S. Anand, S. Abrahams-Gessel, A. Murphy, Growing epidemic of coronary heart disease in low- and middle-Income countries, Curr Probl Cardiol 35 (2) (2010) 72–115.
[5] D.H. O'Leary, J.F. Polak, R.A. Kronmal, T.A. Manolio, G.L. Burke, S.K. Wolfson, Carotid-artery intima and media thickness as a risk factor for myocardial infarction and stroke in older adults, N top N. Engl. J. Med. 340 (1) (1999) 14–22.
[6] A. Doltra, P. Stawowy, T. Dietrich, C. Schneeweis, E. Fleck, S. Kelle, Magnetic resonance imaging of cardiovascular fibrosis and inflammation: from clinical practice to animal studies and back, Biomed Res Int 2013 (10) (2013) 1–2.
[7] M.L. Bots, A.W. Hoes, P.J. Koudstaal, A. Hofman, D.E. Grobbee, Common carotid intima-media thickness and risk of stroke and myocardial infarction: the rotterdam study, Circulation 96 (5) (1997) 1432–1437.
[8] M. Grau, I. Subirana, D. Agis, R. Ramos, X. Basagaña, R. Martí, E. De Groot, R.J. Arnold, J. Marrugat, N. Künzli, R. Elosua, Carotid intima-media thickness in the Spanish population: reference ranges and association with cardiovascular risk factors, Revista Española de Cardiologia 65 (12) (2012) 1086–1093.
[9] P.-J. Touboul, M.G. Hennerici, S. Meairs, H. Adams, P. Amarenco, et al., Mannheim carotid intima-media thickness and plaque consensus (2004-2006-2011), Cardiovasc Dis 34 (4) (2012) 290–296.
[10] C.P. Loizou, A review of ultrasound common carotid artery image and video segmentation techniques, Med. Biol. Eng. Comput. 52 (12) (2014) 1073–1093.
[11] C.P. Loizou, M. Pantziaris, C.S. Pattichis, E. Kyriakou, M-mode state based identification in ultrasound videos of the atherosclerotic carotid plaque, in: 4th International Symposium on Communications, Control, and Signal Processing, 2010, pp. 1–6.
[12] V.M. del Mar, B. Remeseiro, M. Grau, R. Elosua, A. Betriu, E. Fernandez-Giraldez, L. Igual, Semantic segmentation with DenseNets for carotid artery ultrasound plaque segmentation and CIMT estimation, Artif Intell Med 103 (2020) 101784.
[13] F. Molinari, C.S. Pattichis, G. Zeng, L. Saba, U.R. Acharya, R. Sanfilippo, A. Nicolaides, J.S. Suri, Completely automated multiresolution edge snappera new technique for an accurate carotid ultrasound IMT measurement: clinical validation and benchmarking on a multi-institutional database, IEEE Trans. Image Process. 21 (3) (2011) 1211–1222.
[14] M.C. Bastida-Jumilla, R.M. Menchón-Lara, J. Morales-Sánchez, R. Verdú-Monedero, J. Larrey-Ruiz, J.L. Sancho-Gómez, Frequency-domain active contours solution to evaluate intima-media thickness of the common carotid artery, Biomed Signal Process Control 16 (2015) 68–79.
[15] C.P. Loizou, C.S. Pattichis, M. Pantziaris, T. Tyllis, A. Nicolaides, Snakes based segmentation of the common carotid artery intima media, Medical & Biological Engineering & Computing 45 (1) (2007) 35–49.
[16] C. Qian, X. Yang, An integrated method for atherosclerotic carotid plaque segmentation in ultrasound image, Comput Methods Programs Biomed 153 (2018) 19–32.
[17] R.-M. Menchón-Lara, J.-L. Sancho-Gómez, Fully automatic segmentation of ultrasound common carotid artery images based on machine learning, Neurocomputing 151 (2015) 161–167.
[18] M. Biswas, V. Kuppili, T. Araki, D.R. Edla, E.C. Godia, L. Saba, H.S. Suri, T. Omerzu, J.R. Laird, N.N. Khanna, et al., Deep learning strategy for accurate carotid intima-media thickness measurement: an ultrasound study on japanese diabetic cohort, Comput. Biol. Med. 98 (2018) 100–117.
[19] M. Biswas, L. Saba, S. Chakrabartty, N.N. Khanna, H. Song, H.S. Suri, P.P. Sfikakis, S. Mavrogeni, K. Viskovic, J.R. Laird, et al., Two-stage artificial intelligence model for jointly measurement of atherosclerotic wall thickness and plaque burden in carotid ultrasound: a screening tool for cardiovascular/stroke risk assessment, Comput. Biol. Med. 123 (2020) 103847.
[20] M. Biswas, L. Saba, T. Omerzu, A.M. Johri, N.N. Khanna, K. Viskovic, S. Mavrogeni, J.R. Laird, G. Pareek, M. Miner, et al., A review on joint carotid intima–media thickness and plaque area measurement in ultrasound for cardiovascular/stroke risk monitoring: artificial intelligence framework, J Digit Imaging 34 (3) (2021) 581–604.

[21] N. Ikeda, N. Dey, A. Sharma, A. Gupta, S. Bose, S. Acharjee, S. Shafique, E. Cuadrado-Godia, T. Araki, L. Saba, et al., Automated segmental-IMT measurement in thin/thick plaque with bulb presence in carotid ultrasound from multiple scanners: stroke risk assessment, Comput Methods Programs Biomed 141 (2017) 73–81.
[22] S. Lian, Z. Luo, C. Feng, S. Li, S. Li, APRIL: Anatomical prior-guided reinforcement learning for accurate carotid lumen diameter and intima-media thickness measurement, Med Image Anal 71 (2021) 102040.
[23] E. Bianchini, C. Giannarelli, R. Maria Bruno, S. Armenia, L. Landini, F. Faita, V. Gemignani, S. Taddei, L. Ghiadoni, Functional and structural alterations of large arteries: methodological issues, Curr. Pharm. Des. 19 (13) (2013) 2390–2400.
[24] 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, 2015, pp. 234–241.
[25] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in: 32nd International Conference on Machine Learning, volume 37, 2015, pp. 448–456.
[26] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, Journal of Machine Learning Research 15 (2014) 1929–1958.
[27] M. Tan, Q. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, in: International Conference on Machine Learning, 2019, pp. 6105–6114.
[28] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, ImageNet: a large-scale hierarchical image database, in: IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255.
[29] C.P. Loizou, C.S. Pattichis, C.I. Christodoulou, R.S.H. Istepanian, M. Pantziaris, A. Nicolaides, Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery, IEEE Trans Ultrason Ferroelectr Freq Control 52 (10) (2005) 1653–1669.
[30] T.Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, in: IEEE International Conference on Computer Vision, 2017, pp. 2999–3007.
[31] J. Snoek, H. Larochelle, R.P. Adams, Practical bayesian optimization of machine learning algorithms, in: Advances in Neural Information Processing Systems, 2012, pp. 2951–2959.
[32] V. Nair, G.E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: 27th International Conference on Machine Learning, 2010, pp. 807–814.
[33] A. Martín, A. Ashish, B. Paul, B. Eugene, C. Zhifeng, C. Craig, G.S.C.A. Davis, D. Jeffrey, D. Matthieu, et al., Tensorflow: large-scale machine learning on heterogeneous distributed systems, arXiv preprint arXiv:1603.04467 (2016) 1–19.
[34] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, in: 3rd International Conference on Learning Representations, 2015, pp. 1–15.