Early-stage atherosclerosis detection using deep learning over carotid ultrasound images

Rosa-María Menchón-Lara[a,*], José-Luis Sancho-Gómez[a], Andrés Bueno-Crespo[b]

[a]*Departamento de Tecnologías de la Información y las Comunicaciones, Universidad Politécnica de Cartagena, Murcia, Spain.*
[b]*Departamento de Informática de Sistemas, Universidad Católica San Antonio, Murcia, Spain.*

Abstract

This paper proposes a computer-aided diagnosis tool for the early detection of atherosclerosis. This pathology is responsible for major cardiovascular diseases, which are the main cause of death worldwide. Among preventive measures, the Intima-Media Thickness (IMT) of the common carotid artery stands out as early indicator of atherosclerosis and cardiovascular risk. In particular, IMT is evaluated by means of ultrasound scans. Usually, during the radiological examination, the specialist detects the optimal measurement area, identifies the layers of the arterial wall and manually marks pairs of points on the image to estimate the thickness of the artery. Therefore, this manual procedure entails subjectivity and variability in the IMT evaluation. Instead, this article suggests a fully automatic segmentation technique for ultrasound images of the common carotid artery. The proposed methodology is based on Machine Learning and artificial neural networks for the recognition of IMT intensity patterns in the images. For this purpose, a Deep Learning strategy has been developed to obtain abstract and efficient data representations by means of Auto-Encoders with multiple hidden layers. In particular, the considered deep architecture has been designed under the concept of Extreme Learning Machine (ELM). The correct identification of the arterial layers is achieved in a totally user-independent and repeatable manner, which not only improves the IMT measurement in daily clinical practice but also facilitates the clinical research. A database consisting of 67 ultrasound images has been used in the validation of the suggested system, in which the resulting automatic contours for each image have been compared with the average of four manual seg-

*Corresponding author: rmml@alu.upct.es (R. M. Menchón-Lara)

mentations performed by two different observers (ground-truth). Specifically, the IMT measured by the proposed algorithm is 0.625±0.167 mm (mean ± standard deviation), whereas the corresponding ground-truth value is 0.619±0.176 mm. Thus, our method shows a difference between automatic and manual measures of only 5.79±34.42 µm. Furthermore, different quantitative evaluations reported in this paper indicate that this procedure outperforms other methods presented in the literature.

Keywords: Deep learning, Auto-encoders, Extreme learning machine,
Intima-media thickness, Image segmentation

1. Introduction

Cardiovascular diseases (CVD) remain the major cause of death in the world [1]. A large proportion of CVD are caused by an underlying pathological process known as atherosclerosis. Thus, its early diagnosis is critical for preventive purposes. Atherosclerosis involves a progressive thickening of the arterial walls by fat accumulation, which hinders blood flow and reduces the elasticity of the affected vessels.

The Intima-Media Thickness (IMT) of the Common Carotid Artery (CCA) is considered as an early and reliable indicator of atherosclerosis [2] and it is extracted from ultrasound scans [3], i.e. by means of a non-invasive technique. As can be seen in Fig. 1 (left), blood vessels present three different layers, from innermost to outermost: intima, media and adventitia. The IMT is defined as the distance from the lumen-intima interface (LII) to the media-adventitia interface (MAI) . The use of different protocols and the variability between observers are recurrent problems in the IMT measurement procedure. To ensure the repeatability and reproducibility of the process, according to the Mannheim consensus [2], the IMT should be measured preferably on the far wall of the CCA within a region free of atherosclerotic lesions (plaques), where a double-line pattern corresponding to the intima-media-adventitia layers can be clearly observed (see Fig. 1, right).

Figure 1: Diagram of the arterial layers in a transverse section (left) and longitudinal view of the CCA in an ultrasound image (right)

Usually, the IMT is manually measured by the specialist, who marks pairs of points corresponding to the LII and MAI on the ultrasound. It is possible to reduce the subjectivity and variability of manual approaches and detecting the IMT throughout the artery length by means of image segmentation algorithms.

In the last two decades, several solutions have been developed to perform the carotid wall segmentation in ultrasound images [4, 5] for the IMT measurement. Most of the proposed methods require user interaction [6–10]. However, some fully automatic approaches have already been published [11–17].

It is possible to make a classification of techniques according to the used methodology. In this sense, we can find algorithms based on edge detection and gradient-based techniques [6, 8, 9, 18], and other proposals based on dynamic programming [19–24], active contours [7, 13, 14, 25–27], neural networks [11, 12] or in a combination of techniques [10, 16]. There are also techniques based in statistical modelling [17, 28] or in Hough transform [15, 29].

This work addresses a fully automated segmentation technique completely based on Machine Learning to recognize IMT intensity patterns in the carotid ultrasound images. In particular, the developed system intends to emulate the procedure followed by the specialist in the manual protocol. That is, firstly, the detection of the optimal measurement area and, then, the identification of the arterial wall layers. With this purpose, a Deep Learning strategy has been designed to obtain abstract and efficient feature representations by means of Auto-Encoders based on Extreme Learning Machine (ELM). The proposed method jointly extracts the LII and MAI from ultrasound CCA images in a totally user-independent and repeatable manner. Therefore, it improves the reproducibility and objectivity of the IMT evaluation to assist in the early diagnosis of atherosclerosis.

The remainder of this paper is structured as follows: Sect. 2.1 describes the dataset of ultrasound CCA images and the manual segmentations, while Sect. 2.2 introduces the machine learning concepts used in this work. In Sect. 2.3, the proposed segmentation method is explained in detail. The obtained results are shown in Sect. 3. Finally, the main extracted conclusions close the paper.

2. Material and Methods

2.1. Image Database and Manual Segmentations

The set of images used in this work consists of 67 ultrasounds of the CCA taken with a Philips iU22 Ultrasound System using three different ultrasound transducers or probes, with frequency ranges of 9-3 MHz, 12-5 MHz and 17-5 MHz. All of them were provided by the Radiology Department of Hospital Universitario Virgen de la Arrixaca (Murcia, Spain). The parameters of the scanner were adjusted in each case by the radiologist. The spatial resolution of the images ranges from 0.029 to 0.081 mm/pixel, with mean and standard deviation equal to 0.051 and 0.015 mm/pixel, respectively. Some blurred and noisy images, affected by intraluminal artifacts, and some others with partially visible boundaries are included in the studied set.

To assess the performance of the proposed segmentation method, it is necessary to compare the automatic results with some indication of reference values. In our case, the *ground-truth* corresponds to the average of four manual segmentations for each ultrasound image. In particular, two different observers delineated each image twice, with a mean period of one month between tracings. Each manual segmentation of a given ultrasound image includes tracings for the LII and MAI on the far carotid wall. The delineations were performed by marking at least 10 points over the images for each contour, which were subsequently interpolated. Once the four manual contours have been interpolated, the ground-truth for each IMT interface (LII and MAI, separately) is assessed by averaging these in a column-wise manner, i.e., along the longitudinal axis of the image. Figure 2 illustrates the process for the manual segmentations and the definition of the ground-truth for each contour. Hereinafter, we will refer to the different segmentations as:

- MA1: first manual segmentation from observer A.
- MA2: second manual segmentation from observer A.
- MB1: first manual segmentation from observer B.
- MB2: second manual segmentation from observer B.
- GT: ground-truth, average of MA1, MA2, MB1 and MB2.
- AUT: proposed automatic segmentation.

Figure 2: Manual segmentations: application for manual delineation of the IMT interfaces (top); definition of the ground-truth (GT, green line) from four manual segmentations made by two different observers (bottom)

2.2. *Machine Learning Techniques*

In the last decade, Extreme Learning Machine (ELM) has emerged as a powerful tool in the learning process of Single-Layer Feed-Forward Networks (SLFN) by providing good generalization capability at fast learning speed [30]. Given N arbitrary distinct samples ($\mathbf{x}_n, \mathbf{t}_n$), where $\mathbf{x}_n \in \mathbb{R}^d$ is an input vector and $\mathbf{t}_n \in \mathbb{R}^m$ its corresponding target vector, the output of a SLFN with M hidden neurons and activation function $f(\cdot)$ is given by

$$\mathbf{y}_n = \sum_{j=1}^{M} \beta_j f(\mathbf{w}_j \mathbf{x}_n + b_j), n = 1, ..., N; \tag{1}$$

where $\mathbf{w}_j = [w_{j1}, w_{j2}, ..., w_{jd}]$ is the input weight vector connecting the input units and the j-th hidden neuron, $\beta_j = [\beta_{j1}, \beta_{j2}, ..., \beta_{jm}]$ is the output weight vector connecting the j-th hidden neuron and the output units, and b_j is the bias of the j-th hidden neuron. If it is assumed that SLFN can approximate these N samples with zero error, then, there exist β_j, $\mathbf{w}_j$ and b_j such that

$$\sum_{i=j}^{M} \beta_j f(\mathbf{w}_j \mathbf{x}_i + b_j) = \mathbf{t}_i, i = 1, ..., N. \tag{2}$$

ELM is based on the randomly initialization of the input weights and biases of SLFN. Thus, the network can be considered as a linear system and the N equations in the expression (2) can be written compactly in the following form:

$$\mathbf{HB} = \mathbf{T}; \tag{3}$$

where $\mathbf{T} \in \mathbb{R}^{N \times m}$ is the targets matrix, $\mathbf{B} \in \mathbb{R}^{M \times m}$ is the output weights matrix and $\mathbf{H} \in \mathbb{R}^{N \times M}$ is the hidden layer output matrix, which is defined as

$$\mathbf{H} = \begin{bmatrix} f(\mathbf{w}_1 \mathbf{x}_1 + b_1) & \dots & f(\mathbf{w}_M \mathbf{x}_1 + b_M) \\ \vdots & \dots & \vdots \\ f(\mathbf{w}_1 \mathbf{x}_N + b_1) & \dots & f(\mathbf{w}_M \mathbf{x}_N + b_M) \end{bmatrix} \tag{4}$$

Thereby, the training is reduced to solve the linear system in Eq. (3), whose smallest norm least-squares solution is given by:

$$\hat{\mathbf{B}} = \mathbf{H}^{\dagger} \mathbf{T}; \tag{5}$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse matrix of $\mathbf{H}$. Moreover, in order to improve the robustness and generalization performance, a regularization term (C) can be added to the solution [31]:

$$\hat{\mathbf{B}} = \left(\frac{\mathbf{I}}{C} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{T} \tag{6}$$

Although ELM provides an efficient training for SLFN, the performance of machine learning methods and applications highly depends on the selected features for the representation of the problem. Thus, to make progress towards the Artificial Intelligence (AI), the new perspectives in Machine Learning are necessary based on learning data representations that make more accurate classifiers/predictors [32]. In this sense, Deep Learning has emerged as set of algorithms that attempt to model more abstract and useful representation of the data by means of architectures with multiple non-linear transformations [33].

Among the various deep learning architectures, this work focuses on deep networks based on Auto-Encoders (AE). In particular, the ELM Auto-Encoders (ELM-AE) introduced in [34] have been used to solve our segmentation task. Auto-encoders are SLFN performing unsupervised learning in the sense that an AE is trained to reconstruct its own inputs, i.e. $\mathbf{t}_n = \mathbf{x}_n$ (see Fig. 3). Therefore, in the hidden layer of the AE takes place a feature mapping: if $M < d$ (number of hidden neurons $<$ input data dimension), a compressed data coding is obtained as hidden layer output ($\mathbf{H}$); while if $M > d$, the result is a sparse representation of data.

Figure 3: Structure of a generic Auto-Encoder

2.3. Segmentation Procedure

Figure 4 shows an overview of the proposed segmentation methodology. As can be seen in Fig. 1 (right), the raw images contain not only the CCA ultrasound, but also a frame with patient data and additional information is incorporated. Therefore, in order to remove this unwanted frame, the images are automatically cropped at the start. This is done by using Mathematical Morphology to determine the adequate borders of the ultrasound region, because the DICOM fields that provide these parameters are often empty. In particular, an image binarization takes place firstly, then, a procedure is applied to the obtained binary image in order to remove spurious objects and to fill regions or holes. It is based on two basic morphological operators: opening and reconstruction. In this way, a binary mask that matches with the ultrasound region is obtained and the unwanted frame with metadata can be cropped. The process is simple and it does not show any error on the tested images, i.e., the 67 ultrasound images are correctly cropped.

Once the CCA ultrasound is isolated, it is pre-processed to automatically detect the region of interest (ROI), which is the far wall of the blood vessel. Then, those pixels belonging to the ROI are classified for the LII and MAI recognition. The final contours are extracted from the obtained classification results and the IMT can be evaluated on them.

Figure 4: Flow chart of the proposed method for the CCA segmentation

2.3.1. *ROI Detection*

This section describes the first stage of the proposed methodology, in which the carotid far wall (ROI, where the IMT will be evaluated) is located by means of a a system for Pattern Recognition. The scheme of the adopted strategy for this purpose is shown in Fig. 5.

Figure 5: Overview of the strategy for the ROI (far wall of the artery) detection in CCA ultrasounds. An ELM-AE provides a compressed representation of input image blocks at its hidden layer output to improve the classification performance

Specifically, a given CCA ultrasound is split into blocks (squared sections) to proceed with the processing. An ELM-AE has been designed to obtain useful and efficient representations of image blocks for their posterior classification as *'ROI-block'*, if a typical pattern of the far wall is recognized, or *'non-ROI-block'*, otherwise. The size of the image blocks to process is 39×39 pixels, which ensures that the intima-media complex can be contained in a block even if the arterial wall is thick and the radiologist selects high resolutions. Thus, the ELM-AE has an input data dimension of 1521 features.

For the configuration of the ELM-AE, an exhaustive search of the number of hidden neurons and the regularization parameter (M and C, respectively) by means of a validation procedure has been performed. As it is verified later in Sect. 3.1.1, the optimal coding is obtained with 850 hidden sigmoidal neurons. Once the architecture of the AE is fixed, the connections between the new features (hidden layer outputs, $\mathbf{h}$) and the system output ($\mathbf{y}$) are analytically calculated according to Eq. (6).

The dataset used in the design of this system consists of 13776 observations (50% from each class): two thirds for training and the remaining for testing. In particular, the samples were carefully taken from five heterogeneous images (with different orientations of the CCA, spatial resolutions, IMT values, etc.) to assemble a representative and consistent dataset. Table 1 specifies the distribution of the selected samples.

Table 1: Specification of samples used in the design of the system for far wall detection

2.3.2. Arterial Layers Recognition

The segmentation of the LII and MAI in the ultrasound images is carried out by means of a classification of pixels belonging to the ROI. In particular, the intensity values from a certain neighbourhood centred on the pixel to classify provide the necessary contextual information to the classifier for the recognition of the arterial layers. Specifically, the neighbourhoods consist of 51×5 pixels (i.e., 255 input features) and four different classes have been considered as possible system outputs: *'LII-pixel'*, *'MAI-pixel'*, *'IMC-pixel'* (intima-media complex) and *'non-IMC'* (out of the intima-media complex). As in the previous stage, representational learning techniques have been applied to improve the precision of the classifier. With the aim of obtaining meaningful representations from the inputs corresponding to LII and MAI, two different multilayer ELM-AE (with multiple hidden layers) have been implemented. These multilayer auto-encoders (MLAE) are based on the concept set forth in [34], where ELM is used to perform layer-by-layer unsupervised learning. The diagram of the proposed deep architecture can be seen in Fig. 6.

Figure 6: Deep-architecture designed for the LII and MAI segmentation. Two different multilayer ELM-AE produce sparse coding of the input patterns at the output of their second hidden layer. Then, the union of the learned representations is classified for the recognition of the arterial layers

To perform the design and training process of this architecture, a labelled dataset composed of 38908 patterns was assembled by taking samples from 8 manually segmented images. The distribution per class and image of the samples is shown in Table 2: 50% of them belong to the IMT boundaries ('LII' and 'MAI' classes), and the remaining 50% are distributed between 'IMC' class (8904 samples) and 'non-IMC' class (10554 samples). Besides, the dataset is carefully

Table 2: Specification of dataset used in the design of the system for arterial layers recognition

divided into three subsets: one-third of samples for testing, 80% of the remaining two thirds for training and 20% for validation.

The configuration of the LII-MLAE has been done by means of a layer-wise unsupervised training with the 5877 LII samples (training and validation sets). As commented before, the learning process is carried out in a layer-by-layer manner, i.e., each hidden layer is trained as a simple single layer ELM-AE (see Fig. 6) by taking the hidden layer outputs of the previous AE as inputs and desired outputs (unsupervised learning, $\mathbf{t}_n = \mathbf{x}_n$). Therefore, the final LII-MLAE architecture consists of the succession of the hidden layers of the designed single layer auto-encoders along with their corresponding connection weights. Moreover, it is necessary to establish the optimal number of layers for the LII-MLAE. Thus, after designing each layer, it is added to the LII-MLAE architecture only if it allows an improvement in the recognition of LII patterns. In order to know if this happens, the performance of a binary auxiliary classifier ('*LII-pixels*' or '*non-LII-pixels*') is examined using the whole dataset (more details in Sect. 3.1.2). Taking into account this consideration, the optimal architecture for the LII-MLAE consists of two stacked stages, which perform a '255-1100-1900' feature mapping.

On the other hand, using exclusively MAI samples (5672 for training and 1417 for validation), the MAI-MLAE has been configured in a similar manner to represent better the MAI patterns. In this case, the optimal coding is also obtained with two hidden layers ('255-1000-1900' mapping).

In accordance with the suggested system (return to Fig. 6), a given input $\mathbf{x}$ is transformed into two different feature vectors with 1900 dimensions each one. These new representations of the system input ($\mathbf{h}_{\mathbf{LII}}$ and $\mathbf{h}_{\mathbf{MAI}}$) are then joined for proceeding to their classification. The connections between the union of outputs from the second hidden layer of both AE ($\mathbf{h}$) and the system output ($\mathbf{y}$) are computed in accordance with the expression (6).

2.3.3. Extraction of Final Contours

Once a CCA ultrasound is processed by means of the proposed system, the IMT boundaries are properly identified (see Fig. 7, right-central, where the LII and MAI pixels detected are depicted in red and blue, respectively). However, due to the poor definition of the ultrasound images, thick boundaries are obtained. This happens because the system finds the searched intensity patterns in all these pixels. In fact, the classification results cover the variability of the manual segmentations, as can be seen in Fig. 7 (right-central), where the points marked by the two specialist are superimposed.

Figure 7: Example of a good quality processed image: far wall detected (left); ROI with manual segmentations (right-top); recognition of IMT boundaries and manually marked points (right-central); final LII and MAI contours (right-bottom)

Therefore, it is necessary to define the final contours from the system output. For this purpose, the gradient image is evaluated by using morphological operators (Fig. 8, top). Then, the search for the peaks of the intensity gradient is performed (Fig. 8, central). Specifically, the points of maximum gradient which fall within pixels classified as '*LII-pixel*' or '*MAI-pixel*' are considered (Fig. 8, bottom). From these points, two curves corresponding to LII and MAI interfaces are defined by means of a smoothing process based on a moving average filter. The final contours of IMT are determined in this way (see Fig. 7, right-bottom).

Figure 8: Extraction of final contours: gradient image corresponding to Fig. 7 (top); peaks of the intensity gradient (central); points of maximum gradient which fall within pixels classified as '*LII-pixel*' or '*MAI-pixel*' by the system (bottom)

3. Results and Discussion

3.1. Architecture Configuration and Classification Performance

This section includes the results of the performed study for the configuration of the system, as well as the evaluation of its classification performance. For this analysis, several metrics have been used: the *accuracy* (ACC), *specificity* (SPEC), *sensitivity* (SEN), and the *Mathews correlation coefficient* (MCC) of a given classification, which are defined as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$SPEC = \frac{TN}{TN + FP} \tag{8}$$

$$SEN = \frac{TP}{TP + FN} \tag{9}$$

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)\cdot(TP+FN)\cdot(TN+FP)\cdot(TN+FN)}}, \tag{10}$$

where TP is the number of true positives; TN is the number of true negatives; FP and FN are the number of false positives and false negatives, respectively.

3.1.1. ROI Detection

As commented in Sect. 2.3.1, a validation procedure becomes necessary for the configuration of the ELM-AE. The design parameters to select are the number of hidden neurons (M) and the regularization term (C). In our case, 28 different values for M (10, 20, ..., 100, 150, 200, ..., 1000) and 38 different values for C (2^{-18}, 2^{-17}, ..., 2^{19}) have been considered. The ELM-AE was retrained 50 times for every pair of values (50×28×38) and its mean performance has been analysed. Moreover, 20% of training samples were randomly selected as validation set in each trial.

The learning process is carried out in an unsupervised manner, i.e. with target values identical to inputs. However, the selection of its design parameters has been performed by analysing the classification accuracy of the system over the validation samples. For this purpose, a provisional output is computed according to Eq. (5) in each case. Thus, it is possible to analyse the surface of the classification performance (see Fig. 9, left). In view of this result, the optimal compressed coding is obtained with 850 hidden neurons and $C_1 = 2^{-6}$.

Once the single hidden layer of the ELM-AE is adjusted, the output connections of the system must be determined according to Eq. (6). Therefore, a new tuning of another regularization parameter ($C_2 = 2^{-18}$, 2^{-17}, ..., 2^{24}, 2^{25}) is performed. The analysis of the mean validation accuracy (see Fig. 9, right) shows that the optimal value is $C_2 = 2^{19}$, because from that point, the saturation is reached.

Figure 9: Classification accuracy over validation set. Mean performance of the ELM-AE (50 trials) according to M and C_1, i.e. number of hidden neurons and regularization term (left). Analysis for the tuning of the second regularization parameter (C_2) related to the computation of the output connections (right)

Finally, note that the proposed system has proved its good performance. The confusion matrix of the system over the test samples, which are completely unrelated to the training/validation process, is shown in Table 3. From this information, it is possible to deduce that the accuracy of the classification between ROI and non-ROI image blocks is 98.45 $\pm$ 0.06 % (mean and standard deviation from 50 trials). Moreover, the sensitivity is 99.38 $\pm$ 0.06 % and the specificity is 97.56 $\pm$ 0.11 %, which describe the ability of the system to identify positive results (ROI observations) and negative results (non-ROI observations), respectively.

Table 3: Confusion matrix of the system for ROI detection

3.1.2. Detection of IMT Boundaries

Two different multilayer ELM-AE are part of the system developed for the recognition of the arterial layers in CCA ultrasounds (see Fig. 6). The configuration of these machines has been performed in a layer-wise unsupervised manner. For each layer, the number of hidden neurons and the regularization parameter were varied as follows: $M = \{10, 20,..., 500, 550,..., 1000, 1100,..., 2000\}$; and $C = \{2^{-18}, 2^{-17}, ..., 2^{50}\}$. Fifty trials were conducted for each pair of values. In each case, the Root-Mean-Square Error (RMSE) between the equivalent inputs and outputs of the LII-MLAE and MAI-MLAE was evaluated. The optimal parameters (M_{opt} and C_{opt}) for every layer have been selected according to the minimal validation RMSE obtained.

An additional parameter of the architecture to optimize is the number of hidden layers constituting the LII-MLAE and the MAI-MLAE. As it is mentioned in Sect. 2.3.2, a layer is appended to the LII-MLAE or MAI-MLAE architecture only if this fact implies an improvement in the recognition of LII patterns or MAI patterns, respectively. With the aim of determining this enhancement, the whole dataset is passed through the corresponding AE and the performance of a binary classification between '*LII-pixels*' and '*non-LII-pixels*' (or between '*MAI-pixels*' and '*non-MAI-pixels*') is analysed. Note that this provisional labelling of the dataset involves an imbalanced class distribution. Therefore, the connections between the outputs of the hidden layer under analysis and these provisional binary outputs have been assessed according to the Weighted-ELM [35].

The results obtained in the design process of the LII-MLAE and the MAI-MLAE are summarized in Table 4, where ACC and MCC represent the accuracy and Mathews correlation coefficient of the related binary classification, respectively. The latter is generally regarded as a balanced measure which can be used even if the classes are of very different sizes. Thus, the optimal architecture for the LII-MLAE consists of two stacked stages, which perform a '255-1100-1900' feature mapping; whereas in the case of the MAI-MLAE, the optimal coding is also obtained with two hidden layers ('255-1000-1900' mapping).

Table 4: Specification of the analysis for LII-MLAE and MAI-MLAE configuration

The final deep architecture achieves a high success rate over the testing dataset. Table 5 shows the confusion matrix of the system in terms of mean and standard deviation from 50 trials. The overall success rate is 99.44±0.05% (considering the four classes), with 99.76±0.03% of accuracy in the recognition of LII patterns and 99.69±0.04% for the classification of MAI samples.

Table 5: Confusion matrix of the system for the arterial layers segmentation

3.2. Visual Results

The proposed segmentation method has been tested on a set of 67 ultrasound images of the common carotid artery. Fig. 7 shows an example of processed image. Left image depicts the result of the stage for the far wall detection, whereas right pictures show the ROI in detail, where the manual segmentations (top), the classification results and manually marked points (central) and the final IMT contours (bottom) are superimposed on the ultrasound. The correct detection of the far wall is achieved in all the tested images, even in noisy and blurred ones. Some examples can be seen in Fig. 10.

Figure 10: Examples of correct far wall detection in CCA ultrasounds

To ensure an optimal visualization of the IMT boundaries in the ultrasound, a straight and horizontal appearance of the carotid artery in the image is desirable. However, this projection is not always possible. Sometimes, the CCA may be tilted or curved because of the probe position or the own anatomy of the subject. In the case of algorithms using human interaction, the operator can select the optimal area of the image for the IMT measurement. Nevertheless, fully automatic methods must be able to correctly handle the different morphologies of the artery. The examples of final results included in Fig. 11 reveals that the fully automatic segmentation approach proposed in this paper is robust against the orientation and appearance of the CCA in the ultrasound image (slope and curvature).

Figure 11: Final IMT boundaries obtained for the images in Fig. 10

3.3. Segmentation Accuracy

In order to validate the precision of the obtained segmentation results, four manual tracings performed by two different experts are taken into consideration. On the one hand, manual segmentations are compared between themselves in order to characterize the uncertainty and variability of the manual procedure. Thus, the intra-observer errors (MA1 vs MA2 and MB1 vs MB2) as well as the inter-observer errors (MA1 vs MB1, MA1 vs MB2, MA2 vs MB1 and MA2 vs MB2) have been evaluated. On the other hand, the inter-method error is evaluated by comparing our automatic segmentations with those considered as ground-truth (AUT vs GT). In addition, in order to complete the characterization of the automatic IMT contours, comparisons between each of the four manual segmentations and the automatic ones (AUT vs MA1, AUT vs MA2, AUT vs MB1 and AUT vs MB2) have been studied. The different segmentation errors are calculated separately for LII and MAI contours using the Mean Absolute Difference (MAD), which is the most used quantitative metric to evaluate IMT and the accuracy of a segmentation method [4, 5]. This metric represents the average of the vertical distances between two contours along the longitudinal axis of an image.

The box-plot in Fig. 12 shows the distributions of the segmentation errors for LII and MAI over the 67 tested ultrasound images. Moreover, Table 6 includes the maximum, minimum, mean and standard deviation values of the different errors over the image database. Since the scale resolution varies from one image to another, the results are expressed in µm and pixels, for a better description of the difference between segmentations. This statistical analysis reveals that a greater variability exists for the MAI, which is much more noticeable between manual segmentations. This is due to the fact that, in general, transitions from lumen to intima layer are clearer than transitions from media to adventitia layer.

The difference between manual tracings of LII ranges, on average, from 29.7 µm to 40.6 µm, whereas manual segmentation error for MAI varies between 43.5 and 53.9 µm. Despite the greater error and the higher dispersion of the error for the MAI boundaries, there is a good agreement between manual tracings, since the mean differences are around one pixel.

Nevertheless, when the comparisons are made between automatic contours and GT, the segmentation errors for LII and MAI are considerably reduced. Besides, although the MAI error remains slightly greater on average, its distribution is more comparable to the distribution of LII error. In view of the results, it is possible to appreciate that our automatic segmentation reduces the uncertainty and variability of the manual procedure and, therefore, it will lead to a more reliable and precise measurement of the IMT.

Figure 12: Statistical distribution of Mean Absolute Difference (mm) between different segmentations for LII and MAI. Box-plot: in each box, the whiskers extend to the most extreme not outliers values (marked as black crosses), upper and lower box limits represent the 75th and 25th percentile, respectively; the median is depicted by the inner line in the box

Table 6: Mean absolute difference between different segmentations for LII and MAI

3.4. IMT Measurements

Given an ultrasound image and the corresponding boundaries of the arterial wall, manually or automatically segmented contours, the IMT is estimated by using three different metrics: Mean Absolute Difference (MAD), Poly-Line Distance (PLD) and Center Line Distance (CLD). As commented in Sect. 3.3, MAD is the most used metric to evaluate IMT. It is based on the vertical distance between contours along the longitudinal axis of an image. In particular, it is necessary that both contours have the same number of points (N) to calculate the average of these vertical distances (see Fig. 13, top) as follows:

$$IMT_{MAD} = \frac{1}{N}\sum_{x=1}^{N} |LII(x) - MAI(x)| \tag{11}$$

Nevertheless, MAD may deviate from the actual distance between LII and MAI when these contours present certain slope or curvature. To avoid the overestimation in these cases, PLD was proposed in [36] as a more robust and reliable indicator of the distance between two boundaries. It is based on trigonometry and, in this case, it is not a necessary condition that the contours to compare have the same number of points. Given the IMT contours, LII with N_1 points and MAI with N_2 points (see Fig. 13, centre), the distance between a vertex $v = (x_0, y_0)$ in LII and the segment s in MAI (from $v_1 = (x_1, y_1)$ to $v_2 = (x_2, y_2)$) is defined as:

$$d(v,s) = \begin{cases} \mid d_\perp \mid, \text{ if } 0 \leq \lambda \leq d_{12} \\ \min(d_1, d_2), \text{ otherwise} \end{cases} \tag{12}$$

where d_1 and d_2 are the euclidean distances between the vertex v and the vertices in the segment s (v_1 and v_2, respectively); whereas d_{12} is the euclidean distance between v_1 and v_2. As can be seen in Fig. 13 (centre), $d_\perp$ is the perpendicular distance from s to v, and λ is the distance along the segment s between v_1 and the intersection with the perpendicular. In this way, the distance from $v \in$ LII to MAI is calculated as:

$$d(v, MAI) = \min_{s \in MAI} d(v,s) \tag{13}$$

The distance between LII and MAI is evaluated as the sum of the distances from all the vertices in LII to the closest segment in MAI:

$$d(LII, MAI) = \sum_{v \in LII} d(v, MAI) \tag{14}$$

Similarly, the distance from MAI to LII is assessed ($d(MAI, LII)$). And finally, the IMT can be measured by using PLD in the following form:

$$IMT_{PLD} = \frac{d(LII, MAI) + d(MAI, LII)}{N_1 + N_2} \tag{15}$$

The third of the three considered metrics is CLD [37], which also takes into account the local orientation of the IMT boundaries. CLD is based on the calculation of the center line between LII and MAI (see Fig. 13, bottom). Once this line is found, a segment perpendicular to the center line, which intersects with the two boundaries, is considered at each point, and CLD is defined as the mean length of all these segments:

$$IMT_{CLD} = \frac{1}{N} \sum_{i=1}^{N} l_i \tag{16}$$

where l_i is the length of the i-th segment and N is the number of points of the center line. In this case, just like in the MAD metric, the number of points of LII and MAI must be the same.

Figure 13: Diagrams of the three different metrics used to evaluate the IMT: Mean Absolute Difference, MAD (top); Poly-Line Distance, PLD (centre); Center Line Distance, CLD (bottom)

For each ultrasound image, the IMT has been evaluated for manual and automatic segmentations using the aforementioned metrics (MAD, PLD and CLD) to quantify the distance between the corresponding LII and MAI contours. Similar distributions of the IMT values are obtained for the 4 manual measures and the automatic one, as can be seen in Table 7. It is possible to note the slight overestimation produced by the MAD metric in the IMT measurement.

Table 7: Statistics of the IMT measurements (n = 67 images) in millimetres using different metrics (MAD, PLD and CLD)

As in Sect. 3.3, intra and inter-observer IMT errors have been estimated for the manual measurements to be compared with the error between automatic IMT and GT. Given two different segmentations S_1 and S_2, the degree of agreement between its IMT measures over the 67 images is assessed by calculating three figures of merit: the correlation coefficient (ρ), the absolute error value ($\varepsilon_{IMT_i} = |IMT_i^{S_1} - IMT_i^{S_2}|$, for each image) and the difference between measurements ($\Delta_{IMT_i} = IMT_i^{S_1} - IMT_i^{S_2}, i = 1, ..., 67$). Table 8 shows the results of the IMT measurement error analysis.

The IMT intra-observer reproducibility is of 98.4% for observer A and practically 98% for observer B (see the corresponding correlation coefficients in Table 8). Moreover, the inter-observer reproducibility of the IMT measurements is around 97%. This high grade of agreement between manual measures confirms the goodness of the 4 manual segmentations and, consequently of the GT, which is defined as the average of these ones. The absolute errors and the standard deviations of the differences indicate a greater IMT inter-observer error in comparison with the error between IMT measures from the same observer, which seems logical.

Table 8: Comparison between IMT measurements (MAD, PLD and CLD metrics) from different segmentations. n = 67 images; ρ: Correlation coefficient; ε_{IMT}: Mean $\pm$ standard deviation of the absolute errors; Δ_{IMT}: Mean $\pm$ standard deviation of the differences

The difference between automatic IMT and GT is of 5.8 $\pm$ 34 µm for MAD and PLD metrics (6.7 $\pm$ 34 µm for CLD), whereas the absolute error of the automatic measurements is 27.3 $\pm$ 21 µm for MAD and PLD (27.2 $\pm$ 22 µm for CLD). These values reveal that the measurement error associated with the proposed method is lower than the inter-observer errors and it is in the rank of the intra-observer errors. In addition, the correlation coefficient (98.1%) is comparable to the intra-observer variability.

Figure 14 (right) depicts the linear regression analysis between automatic IMT and GT (MAD, PLD and CLD metrics), where the high degree of agreement can be observed. Furthermore, Fig. 14 (left) shows the corresponding Bland-Altman plots (MAD, PLD and CLD metrics) with the following limits of agreement (mean $\pm$ 2 $\times$ standard deviation): 5.8 $\pm$ 68.8 µm for MAD, 5.8 $\pm$ 68.5 µm for PLD, and 6.7 $\pm$ 68.5 µm for CLD. The vertical axis in the Bland-Altman plot represents the difference between AUT and GT measures of the IMT, whereas the horizontal axis represents the average of the values compared. Therefore, the precision in the automatic IMT measurements is full well justified.

Figure 14: Analysis of automatic IMT measurements: Bland-Altman plot (left) and linear regression analysis (right)

4. Conclusions

This paper proposes a fully automated segmentation method for CCA ultrasounds to accurately measure the IMT, an early indicator of atherosclerosis and cardiovascular risk. This proposal is completely based on Machine Learning, in order to detect the arterial far wall and to extract the IMT contours (LII and MAI) in a reliable and automatic way. In particular, the suggested architecture is based on the Extreme Learning Machine (ELM). Furthermore, Auto-Encoders (AE) and Deep Learning concepts have been used to obtain useful data representations for solving the segmentation task, which is posed as a Pattern Recognition problem. Following the developed strategy, the IMT can be measured in a totally user-independent and repeatable manner.

The method has been tested over a database of 67 images with different spatial resolutions. The validation of the technique is carried out by comparing the automatic contours with the average of four manual segmentations performed by two different observers. The results show a mean segmentation error of 0.028 mm for the LII and 0.035 mm for the MAI and demonstrate that the proposed methodology reduces the uncertainty and variability of the manual procedure. In reference to the IMT measurements, a high grade of agreement between manual and automatic observations is obtained with a difference of only 5.8 ± 34 μm (mean and standard deviation).

With these error values, the algorithm outperforms both automatic and semi-automatic methods presented in the literature. Table 9 summarizes the results reached by other IMT measurement methods. All methods included in Table 9 do not consider ultrasound images with atherosclerotic plaques, and use MAD metric to evaluate the IMT and the different errors. Although direct comparisons between different studies are difficult due to the dependence on the measurement protocol, characterization of the results, number and type of patients, tissue to be segmented and image quality, it can be seen that our automatic segmentation compares favourably to other semi-automatic and automatic algorithms.

Table 9: Some IMT measurement methods

It is important to pay special attention to the works [11] and [12], which correspond to previous contributions from the authors of this paper. In the previous papers [11, 12], as well as in this one, a novel point of view is presented for this specific application and the posed problem is solved by means of Pattern Recognition techniques. This new proposal is oriented to improve the generalization capability and the performance of the method in order to achieve the best system configuration.

In particular, the new approach is completely based on Machine Learning (ML), both the recognition of the carotid far wall (region of interest, ROI) and the identification of the IMT contours (LII and MAI). Whereas in [11, 12], different image processing techniques were applied to detect the ROI, none of them based on ML. The use of a new pattern recognition strategy based on ML for the recognition of the ROI implies that the system is able to adapt to the optimal area for the measurement, by avoiding those uncertain regions in which the characteristic IMT pattern is unclear, blurred or even hidden. Moreover, in this work, it has been studied the utilisation of ELM multilayer AE to obtain sparse data representations with the aim of obtaining a high performance classifier. With the proposed architecture, the obtained results show an overall success rate exceeding 99% in the classification of the nearly 13000 test samples. In [12], single-layer AE were considered only to reduce the data dimension. ELM algorithm was not used in the previous contributions [11, 12], but in the present study it has provided advantages in the learning process (training and design) of the proposed system, because of its good performance at fast speed even with high-dimensional data. Furthermore, the suggested strategy has been designed to recognize jointly LII and MAI and it is able to identify and differentiate both contours by means of the developed multiclass classifier (4 classes); whereas in the previous works [11, 12], only binary classifications were considered. In this way, the post-classification stage of the new proposal does not require hard efforts for debugging the results and for the extraction of the final IMT boundaries and it has been simplified notably.

To summarize, the main contributions and improvements of the proposed method are the following: a greater intelligence and autonomy of the system; the high degree of robustness against the anatomical and instrumental variability of the ultrasound images; and the noteworthy reliability and precision in the evaluation of the IMT. Finally, it is important to emphasize that these positive aspects are crucial for a fully automatic method which has been designed to assist in the early detection of atherosclerosis and the prevention of cardiovascular diseases. According to this, it can be concluded that the proposed methodology is suitable for the evaluation of IMT not only in the daily clinical practise, but also in clinical research studies.

Acknowledgment

Authors would like to thank the Radiology Department of 'Hospital Universitario Virgen de la Arrixaca' (Murcia, Spain) for their kind collaboration and for providing all the ultrasound images used.

[1] WHO, Global atlas on cardiovascular disease prevention and control, online. URL `www.who.int/cardiovascular_diseases/en/`

[2] P.-J. Touboul, et al., Mannheim carotid intima-media thickness and plaque consensus (2004-2006-2011), Cerebrovasc. Dis. 34 (2012) 290–296.

[3] K. S. Nikita, Atherosclerosis: The evolving role of vascular image analysis, Comput. Med. Imaging Graph. 37 (1) (2013) 1–3. doi:http://dx.doi.org/10.1016/j.compmedimag.2012.12.001.

[4] F. Molinari, G. Zeng, J. S. Suri, Review: A state of the art review on intima-media thickness (imt) measurement and wall segmentation techniques for carotid ultrasound, Comput. Methods Prog. Biomed. 100 (3) (2010) 201–221. doi:10.1016/j.cmpb.2010.04.007.

[5] C. Loizou, A review of ultrasound common carotid artery image and video segmentation techniques, Medical & Biological Engineering & Computing 52 (12) (2014) 1073–1093. doi:10.1007/s11517-014-1203-5.

[6] C. Liguori, A. Paolillo, A. Pietrosanto, An automatic measurement system for the evaluation of carotid intima-media thicknesss, Instrumentation and Measurement, IEEE Transactions on 50 (6) (2001) 1684–1691.

[7] D.-C. Cheng, A. Schmidt-Trucksass, K.-S. Cheng, H. Burkhardt, Using snakes to detect the intimal and adventitial layers of the common carotid artery wall in sonographic images, Computers Methods and Programs in Biomedicine 67 (1) (2002) 27–37.

[8] J. H. Stein, C. E. Korcarz, M. E. Mays, P. S. Douglas, M. Palta, H. Zhang, T. LeCaire, D. Paine, D. Gustafson, L. Fan, A semiautomated ultrasound border detection program that facilitates clinical measurement of ultrasound carotid intima-media thickness, Journal of the American Society of Echocardiography 18 (3) (2005) 244–251.

[9] F. Faita, V. Gemignani, E. Bianchini, C. Giannarelli, L. Ghiadoni, M. Demi, Real-time measurement system for evaluation of the carotid intima-media thickness with a robust edge operator, Journal of ultrasound in medicine 27 (9) (2008) 1353–1361.

[10] R. Rocha, A. Campilho, J. Silva, E. Azevedo, R. Santos, Segmentation of the carotid intima-media region in b-mode ultrasound images, Image and Vision Computing 28 (4) (2010) 614–625.

[11] R.-M. Menchón-Lara, M.-C. Bastida-Jumilla, J. Morales-Sánchez, J.-L. Sancho-Gómez, Automatic detection of the intima-media thickness in ultrasound images of the common carotid artery using neural networks, Med. Biol. Eng. Comput. 52 (2) (2014) 169–181. doi:10.1007/s11517-013-1128-4.

[12] R.-M. Menchón-Lara, J.-L. Sancho-Gómez, Fully automatic segmentation of ultrasound common carotid artery images based on machine learning, Neurocomputing 151, Part 1 (0) (2015) 161 – 167. doi:http://dx.doi.org/10.1016/j.neucom.2014.09.066.

[13] 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 intimamedia thickness of the common carotid artery, Biomedical Signal Processing and Control 16 (0) (2015) 68 – 79. doi:http://dx.doi.org/10.1016/j.bspc.2014.08.012.

[14] S. Delsanto, F. Molinari, P. Giusetto, W. Liboni, S. Badalamenti, J. Suri, Characterization of a completely user-independent algorithm for carotid artery segmentation in 2-d ultrasound images, IEEE Transactions on Instrumentation and Measurement 56 (4) (2007) 1265–1274.

[15] S. Golemati, J. Stoitsis, E. G. Sifakis, T. Balkizas, K. S. Nikita, Using of the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery, Ultrasound in Medicine & Biology 33 (12) (2007) 1918–1932.

[16] F. Molinari, G. Zeng, J. Suri, Intima-media thickness: setting a standard for a completely automated method of ultrasound measurement, Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on 57 (5) (2010) 1112–1124.

[17] D. Ilea, C. Duffy, L. Kavanagh, A. Stanton, P. Whelan, Fully automated segmentation and tracking of the intima media thickness in ultrasound video sequences of the common carotid artery, Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on 60 (1) (2013) –. doi:10.1109/TUFFC.2013.2547.

[18] R. H. Selzer, W. J. Mack, P. L. Lee, H. Kwong-Fu, H. N. Hodis, Improved common carotid elasticity and intima-media thickness measurements from computer analysis of sequential ultrasound frames, Atherosclerosis 154 (1) (2001) 185 – 193.

[19] T. Gustavsson, Q. Liang, I. Wendelhag, J. Wikstrand, A dynamic programming procedure for automated ultrasonic measurement of the carotid artery, in: Computers in Cardiology 1994, 1994, pp. 297–300.

[20] I. Wendelhag, Q. Liang, T. Gustavsson, J. Wikstrand, A new automated computerized analysis system simplifies readings and reduces the variability in ultrasound measurement of intima-media thickness, Stroke 28 (1997) 2195–2200.

[21] Q. Liang, I. Wendelhag, J. Wikstrand, T. Gustavsson, A multiscale dynamic programming procedure for boundary detection in ultrasonic artery images, Medical Imaging, IEEE Transactions on 19 (2) (2000) 127–142.

[22] D.-C. Cheng, X. Jiang, Detections of arterial wall in sonographic artery images using dual dynamic programming, Information Technology in Biomedicine, IEEE Transactions on 12 (6) (2008) 792 –799.

[23] N. Santhiyakumari, M. Madheswaran, Non-invasive evaluation of carotid artery wall thickness using improved dynamic programming technique, Signal, Image and Video Processing 2 (2) (2008) 183–193.

[24] Y.-B. Lee, Y.-J. Choi, M.-H. Kim, Boundary detection in carotid ultrasound images using dynamic programming and a directional haar-like filter, Computers in Biology and Medicine 40 (8) (2010) 687 – 697.

[25] C. P. Loizou, C. S. Pattichis, M. Pantziaris, T. Tyllis, A. Nicolaides, Snakes based segmentation of the common carotid artery intima media, Medical and biological engineering and computing 45 (1) (2007) 35–49.

[26] F. Molinari, K. M. Meiburger, L. Saba, U. R. Acharya, M. Ledda, A. Nicolaides, J. S. Suri, Constrained snake vs. conventional snake for carotid ultrasound automated imt measurements on multi-center data sets, Ultrasonics 52 (7) (2012) 949 – 961.

[27] S. Petroudi, C. Loizou, M. Pantziaris, C. Pattichis, Segmentation of the common carotid intima-media complex in ultrasound images using active contours, IEEE Transactions on Biomedical Engineering 59 (11) (2012) 3060–3069.

[28] F. Destrempes, J. Meunier, M.-F. Giroux, G. Soulez, G. Cloutier, Segmentation in ultrasonic b-mode images of healthy carotid arteries using mixtures of nakagami distributions and stochastic optimization, Medical Imaging, IEEE Transactions on 28 (2) (2009) 215 –229.

[29] X. Xu, Y. Zhou, X. Cheng, E. Song, G. Li, Ultrasound intima-media segmentation using hough transform and dual snake model, Computerized Medical Imaging and Graphics 36 (3) (2012) 248–258.

[30] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: Theory and applications, Neurocomputing 70 (1 - 3) (2006) 489 – 501. doi:http://dx.doi.org/10.1016/j.neucom.2005.12.126.

[31] G.-B. Huang, H. Zhou, X. Ding, R. Zhang, Extreme learning machine for regression and multiclass classification, Syst. Man Cybern. Part B-Cybern., IEEE Trans. 42 (2) (2012) 513–529. doi:10.1109/TSMCB.2011.2168604.

[32] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, Pattern Anal. Mach. Intell., IEEE Trans. 35 (8) (2013) 1798–1828. doi:10.1109/TPAMI.2013.50.

[33] L. Deng, D. Yu, Deep learning: Methods and applications, Tech. Rep. MSR-TR-2014-21 (January 2014).
URL `http://research.microsoft.com/apps/pubs/default.aspx?id=2093`

[34] L. L. C. Kasun, H. Zhou, G.-B. Huang, C. M. Vong, Representational learning with extreme learning machine for big data, IEEE Intelligent Systems 28 (6) (2013) 31–34.

[35] W. Zong, G.-B. Huang, Y. Chen, Weighted extreme learning machine for imbalance learning, Neurocomputing 101 (0) (2013) 229 – 242. doi:http://dx.doi.org/10.1016/j.neucom.2012.08.010.

[36] J. S. Suri, R. M. Haralick, F. H. Sheehan, Greedy algorithm for error correction in automatically produced boundaries from low contrast ventriculograms, Pattern Analysis & Applications 3 (1) (2000) 39–60. doi:10.1007/s100440050005.

[37] F. H. Sheehan, D. K. Stewart, H. T. Dodge, S. Mitten, E. L. Bolson, B. G. Brown, Variability in the measurement of regional left ventricular wall motion from contrast angiograms, Circulation 68 (3) (1983) 550–559.