Bimodal Automated Carotid Ultrasound Segmentation using Geometrically Constrained Convolutional Neural Networks

Carl Azzopardi, *Member, IEEE,* Kenneth P. Camilleri, *Senior Member, IEEE,* and Yulia Hicks, *Senior Member, IEEE*

***Abstract*—For asymptomatic patients suffering from carotid stenosis, the assessment of plaque morphology is an important clinical task which allows monitoring of the risk of plaque rupture and future incidents of stroke. Ultrasound Imaging provides a safe and non-invasive modality for this, and the segmentation of media-adventitia boundaries and lumen-intima boundaries of the Carotid artery form an essential part in this monitoring process. In this paper, we propose a novel Deep Convolutional Neural Network as a fully automated segmentation tool, and its application in delineating both the media-adventitia boundary and the lumen-intima boundary. We develop a new geometrically constrained objective function as part of the Network's Stochastic Gradient Descent optimisation, thus tuning it to the problem at hand. Furthermore, we also apply a novel, bimodal fusion of envelope and phase congruency data as an input to the network, as the latter provides an intensity-invariant data source to the network. We finally report the segmentation performance of the network on transverse sections of the carotid. Tests are carried out on an augmented dataset of 81,000 images, and the results are compared to other studies by reporting the DICE coefficient of similarity, modified Hausdorff Distance, sensitivity and specificity. Our proposed method is shown to yield results of comparable accuracy over this larger dataset, with the advantage of it being fully automated. We conclude that Deep Convolutional Neural Networks provide a reliable trained manner in which carotid ultrasound images may be automatically segmented, using envelope data and intensity invariant phase congruency maps as a data source.**

***Index Terms*—Ultrasound, Segmentation, Deep Convolutional Networks, Carotid Artery, Phase Congruency.**

I. Introduction

ACCORDING to statistics presented by the American Hearth Association, cerebrovascular disease is amongst the leading causes of death in the United States, with an estimated 7 million Americans above 20 years of age having had a stroke between 2013 and 2016 [1]. In 2016 alone, stroke accounted for approximately 1 in every 19 deaths, and it thus ranks as the fourth leading cause of disease mortality in the United States after heart disease, cancer and chronic lower respiratory disease [1].

The underlying cause of these diseases is atherosclerosis - a vascular pathology which is characterised by the thickening and hardening of blood vessel walls [2]. When fatty substances such as cholesterol, triglycerides, or cellular waste products such as calcium and fibrin, start to accumulate on the inner linings of an artery, they cause a progressive narrowing of the lumen and consequently restrict the free flow of blood [3]. This is shown in Figure 1. The carotid is one such artery which is susceptible to such atherosclerotic deposits - or plaque. Located on either side of the neck, it has the vital function of supplying blood to the brain and to the muscles in the face [4]. When atherosclerotic plaque ruptures in the carotid artery, there is a significant risk that the blood clot which forms will eventually travel upstream to occlude a narrower vessel in the brain - ultimately leading to a stroke [4].

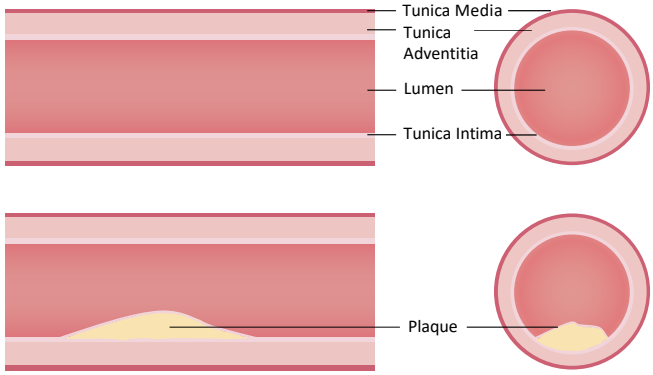


Fig. 1: [Top] An example of a healthy carotid artery and [Bottom] an example of an artery with an accumulation of atherosclerotic plaque.

Localisation and grading of the severity of a stenosis, forms a large part of the diagnostic process that clinical practitioners use to assess the risk of rupture. Techniques such as Digital Subtraction Angiography (DSA) and Magnetic Resonance Angiography (MRA) are presently considered to be the gold standard in assessing carotid disease severity [5]–[7]. However, they are also considered to be somewhat invasive, and pose risks to the patients due to the risk of emboli or nephrotoxicity arising from the contrast agent used [7]. Ultrasound imaging has therefore widely gained popularity due to its low cost

C. Azzopardi and Y. Hicks are with the Sensors, Signals and Imaging Research Group, School of Engineering, Cardiff University, Wales. E-mail: (AzzopardiC@cardiff.ac.uk)

K. P. Camilleri is with Centre for Biomedical Cybernetics, University of Malta, Msida, Malta.

Manuscript received May 18, 2019

and and non-invasive nature, permitting a quick assessment of vessel geometry, degree of stenosis, as well as of plaque morphology [4], [5], [7]. Classically, the intima-media thickness (IMT) has been the clinical measure used in Ultrasound imaging to monitor plaque progression and burden [8], [9]. However, given that plaque progresses faster along the length of the vessel as opposed to in thickness, it may be better indicated to monitor plaque progression in 3D, as this would provide a better indication of sites which are at risk of rupture [9].

In order for accurate measurement of vessel geometry to take place - be it for later three dimensional reconstruction of the artery or for further assessment of plaque burden using metrics such as Total Plaque Volume (TPV) or Vessel Wall Volume (VWV), two specific wall interfaces need to be identified: the media-adventitia boundary (MAB) and the Lumen - Intima boundary (LIB) [10]. Both these interfaces need to be delineated in a robust and reproducible manner, and manual methods have been shown to be tedious, labour intensive [10], and prone to variability [11]. Thus, considerable attention has been devoted to developing automated or semi-automated carotid segmentation algorithms which facilitate this process.

II. Related Work

Previous literature shows certain patterns in the development of the carotid segmentation techniques, allowing us to categorise them by the nature of their approach. The first distinction is that between addressing the segmentation problem on longitudinal sections of the carotid, or segmentation on transverse sections of the carotid. The majority of studies available address the former [10], since this type of segmentation then easily lends itself to evaluating the intima-media thickness. A comprehensive review by Molinari *et al.* [8] addresses some of the major works on the subject matter.

A. Segmentation in Longitudinal Images

A popular approach for longitudinal carotid segmentation has been through using edge tracking or gradient-based techniques, as shown in a number of studies [12]–[14]. In the longitudinal section, the common carotid artery may be considered as a dark region surrounded by two bright line patterns - the near wall and the far wall of the artery [8]. By considering the intensity profile, or the intensity gradient across a section cutting across the artery, the adventitial walls may be clearly identified, and the IMT estimate may be obtained as the distance between these two points [12]–[14].

Another widely used segmentation approach is based on Active Contours, or Snakes. This entails having a set of vertices connected by line segments [8] which dynamically move to settle around the desired contour, under the action of defined forces. Snake models however have issues that effect their performance. They require correct fine tuning of parameters for them to be correctly attracted to edges; they depend on the initialisation of the snake model, and they are also prone to leaking through edges which are not clearly defined [8]. Other segmentation approaches for longitudinal sections described in [8] include the use of Dynamic Programming, Nakagami modelling, the use of the Hough Transform, and the use of motion estimation and Bayesian frameworks.

B. Segmentation in Transverse Images

A number of studies have also addressed the problem of segmenting carotid arteries in the transverse section, using either native 2D images or else from transverse slices extracted from 3DUS images. In 2009, Seabra *et al.* [15] proposed a semi-automatic technique for plaque segmentation in transverse images based on a manually initialised 2D active contour algorithm. Another study by Yang *et al.* [16] proposed to use active shape models to segment both the MA and LI interfaces. Ukwatta *et al.* [10] proposed a novel semi-automated technique based on a level-set method to segment the MA and LI interfaces. The operator was asked to provide anchor points as high-level domain knowledge, and this together with the incorporation of local and global image statistics with boundary separation-based constraints allowed accurate segmentation of the MA and LI interfaces. Other segmentation approaches for transverse sections included the use of deformable models [11], [17], modified Cohen Snakes [18] and a Star algorithm improved by Kalman filtering [19].

Alternative methods have been proposed to segment the carotid structures or plaque morphology natively in three dimensions. Gill *et al.* [7] proposed a semi-automatic method based on a dynamic balloon model in 2000. In 2010, Solovey *et al.* [20] also proposed an LI interface segmentation algorithm on native 3D images based on a level-set method. In 2015, Hossain *et al.* [9] presented a semi-automatic method for segmenting both MA and LI interfaces using a distance regularized level set algorithm, with a novel stopping criterion and a modified energy function.

One notes further that, particularly in the case of transverse segmentation, studies have aimed to segment either the LI interface alone [7], [11], [17], [20]–[22], or else both the LI and the MA interfaces [9], [10], [16], [23]. The latter approach has increasingly gained interest, due to new volumetric parameters such as vessel wall volume (VWV) and total plaque volume (TPV) which have been proposed to characterise plaque burden [24]. Although these parameters have been used in a number of research studies, they have not yet gained widespread clinical acceptance due to certain challenges, amongst which are the tediousness and inter/intra-observer variability when manually delineating the LI and MA interfaces [9]. Studies have therefore sought to develop automatic or semi-automatic algorithms to segment these interfaces.

C. Our Contribution

Deep networks have recently garnered much interest, as they have driven forward the state-of-the-art in computer vision tasks such as image classification, object detection and

segmentation [25]. Such advancements have also been picked up by the medical imaging research community. The survey by Litjens *et al.* in 2017 [26] provides a comprehensive review of studies employing Deep networks for a variety of tasks and application areas within medical imaging. Litjens *et al.* review over 300 contributions since 2012, with approximately 20 of these being contributions within the ultrasound imaging field, and in turn 6 of these being related to cardiac / vascular applications. None of these however treat the subject of carotid segmentation. The study by Menchon-Lara *et al.* in [27] addresses longitudinal carotid ultrasound segmentation using a single layer perceptron network preceded by an autoencoder, for the purpose of intima-media thickness estimation. The segmentation of the MAB interface in transverse and longitudinal carotid images using deep convolutional neural networks is thus a novel application, proposed by ourselves in our preliminary work reported in [28]. In [28], we evaluate various network configurations to find the optimal network size and depth, as well as the optimal filter dimension. We further propose a novel fusion of envelope and phase congruency data as an input to the network, as the latter provides an intensity-invariant data source to the network.

In this work, we propose a novel, geometrically constrained Deep Convolutional Neural Network (DCNN), as well as its novel application to concurrently segment both the MAB interface and LIB interface on a larger data set of symptomatic and asymptomatic transverse carotid ultrasound images. The segmentation of both MAB and LIB is necessary if the total plaque burden contained in between these two interfaces is to be quantified. We propose a new geometrically constrained objective cost function which is constructed and tuned towards the segmentation of carotid structures. An enhanced data augmentation strategy is also employed to improve and enlarge further the size of the dataset. Finally, we present an extensive set of experiments to test the performance of our technique, and we compare it to the performance of other methods by reporting the DICE coefficient of similarity, modified Hausdorff Distance, the Sensitivity and Specificity.

III. Methods

A. Study subjects and Image Acquisition

A total of fifty transverse ultrasound images were obtained from across 15 subjects, having carotid arteries which display varying degrees of stenosis of between 0% (negligible) and 60% (European Carotid Surgery Trial ECST Criteria). Subjects had ages spanning between 60 to 80 years. Subjects provided signed consent to participate in the study, which was in turn approved by the University of Malta Research Ethics Committee. The 2D ultrasound images were acquired using an Ultrasonix Sonix RP ultrasound machine (Analogic Corporation, Peabody, MA, USA), equipped with a 14 MHz L14-5 Linear Probe. Scanner settings were set as follows: Frequency: 6.6MHz, Depth: 3.0cm, Sector width: 100%, Gain: 51%, Dynamic Range: 92dB, Persist Setting: 0, Map Setting: 9, Chroma Setting: 0; Power Setting: 0; and were kept constant across all subject acquisitions. Subjects were asked to lie supine on a couch, and the probe was placed against the neck while an image sequence was acquired for transverse carotid sections. The probe was kept in the same spatial location while the image sequence was acquired at a rate of 24 Hz.

B. Data Pre-processing

Image Pre-Processing. All data acquired from the Sonix RP Ultrasound Scanner was saved in raw RF format and transferred to a workstation for data processing with MATLAB (Mathworks, USA). For the creation of B-mode images, a Hilbert transform was used to demodulate the amplitude information from the RF sinusoids. The amplitude data was then passed through a logarithmic function to adjust for dynamic range, and then decimated by a factor of four. The resulting B-mode image was scan converted to obtain correct image geometry. A median speckle reduction filter was then implemented as per methodology described in [29], and applied over all the B-mode images to reduce the effect of speckle noise.

Manual Segmentation. All the acquired transverse ultrasound images were manually and independently traced with the assistance of 2 radiographers, and used as labelled training data for the DCNN. Each radiographer manually traced the image sets twice, with a period of 2 weeks in between sessions. The radiographers traced both the MAB and LIB in the transverse sections of the carotid arteries.

C. Data Augmentation.

An abundance of training data is an important prerequisite for correctly training a neural network. If the data is scarce, simple distortions such as rotations, translations or skewing, may be applied to the original training data as an easy way of expanding the size of the data set [30]. In the case of the ultrasound data sets, transverse image sets, as well as the corresponding labelled datasets, were first scaled twice by a factor of 1.2 and 1.5. The scaled images were then cropped back to their original dimension. Each image, having dimensions of 256 x 256 pixels was then patch-wise sampled 9 times along an equally spread 3x3 grid centred about the middle, using a 120 x 120 pixel window with a constant overlap between each region. Finally, the images were rotated through 90, 180 and 270 degrees. All the additionally generated image sets and corresponding labels were concatenated into separate augmented dataset-pairs. From 15 different patients, originally having a total of 50 images each, the dataset was augmented to a total of 81,000 images.

D. Extracting Phase Congruency Information

The presence of speckle noise, low contrast, and local changes of intensity make ultrasound image segmentation a fairly difficult problem [31]. Methods which seek to delineate a contour of interest based on a globally set intensity threshold are often problematic - particularly because relevant contours might not be visible due to tissue-dependent attenuation, transducer orientation or structure depth [32]. There is also another inherent variable which is the end user. Preferences

on intensity gain settings vary across end-users and across machines, thus making it difficult to find optimum parameter values which apply across the board. It is therefore desirable to approach the ultrasound segmentation problem using features which are intensity-invariant and robust to attenuation. The use of phase information provides one such possibility, since phase is theoretically amplitude-invariant, and preserves the structural information of a signal [31], [32].

In 1987, Morrone *et al.* [33] proposed a model of feature perception called the local energy model. The said model postulates that features may be perceived within an image, at the points where the Fourier components are maximally in phase. At the point of an edge transition, all the constituent Fourier components of a square wave are exactly in phase - and hence the *phase congruency* is said to reach a maximum value. More specifically, we note that the phase is 0° at positive edges, and 180° at negative edges. At all other points in the square wave, the phase congruency will be low [34]. In their work, Morrone *et al.* [33] define phase congruency as follows:

$$PC(x) = \max_{\bar{\phi}(x)\in[0,2\pi]} \frac{\sum_n A_n \cos(\phi_n(x) - \bar{\phi})}{\sum_n A_n} \tag{1}$$

where A_n is the amplitude of the *n*th Fourier component of a one-dimensional signal $I(x) = \sum_n A_n cos(\phi_n(x))$, and ϕ_n is the local phase. The value of $\bar{\phi}(x)$, over which the equation is maximised, is the amplitude weighted mean local phase angle of all Fourier components at the point being considered. Venkatesh *et al.* [35] however show that points of maximum phase congruency may also be found by searching for peaks in the local energy function, and that the energy function is equal to the phase congruency scaled by the sum of the Fourier amplitudes:

$$E(x) = PC(x) \sum_n A_n \tag{2}$$

One clearly notes therefore that the local energy model is directly proportional to the phase congruency function, and thus, peaks in local energy correspond to peaks in phase congruency.

Kovesi [34] proposed a more convenient method of computing phase congruency via the local energy model, by convolving the signal with a filter bank of quadrature logarithmic Gabor filters. These allow for an arbitrarily large bandwidth filter to be constructed, while still maintaining a zero DC component in the even-symmetric filter. The log Gabor function has the following transfer function on the linear frequency scale [34]:

$$G(\omega) = \exp - \frac{(\log(\omega/\omega_0))^2}{2(\log(\kappa/\omega_0))^2} \tag{3}$$

where ω_0 is the filter's centre frequency, and the term κ/ω_0 ensures a constant shape-bandwidth ratio by keeping it constant over varying ω_0. If the even-symmetric (cosine) and odd-symmetric (sine) wavelets are denoted by M_n^e and M_n^o respectively at a scale n, then, given signal I, the responses of each quadrature pair of filters may be seen as follows:

$$[e_n(x), o_n(x)] = [I(x) * M_n^e, I(x) * M_n^o] \tag{4}$$

The amplitude and phase at a given scale n is therefore given by the following equations:

$$\begin{aligned} A_n(x) &= \sqrt{e_n(x)^2 + o_n(x)^2} \\ \phi_n(x) &= \arctan 2(e_n(x), o_n(x)) \end{aligned} \tag{5}$$

where A_n and ϕ_n are response vectors defined at each point x in the signal. The response vectors form the localised representation of the signal, and may be used in the same way that Fourier components are used to calculate phase congruency [34]. Thus, given that:

$$PC(x) = \frac{E(x)}{\sum_n A_n(x)} \tag{6}$$

and that:

$$E(x) = \sqrt{e(x)^2 + o(x)^2} \tag{7}$$

it follows that:

$$PC(x) = \frac{\sqrt{(\sum_n e_n(x))^2 + (\sum_n o_n(x))^2}}{\sum_n \sqrt{e_n(x)^2 + o_n(x)^2}} \tag{8}$$

To mitigate the problem of noise resistance, Kovesi further suggests to modify Eq. (7) as follows:

$$PC(x) = \frac{E(x) - T}{\sum_n A_n(x) + \varepsilon} \tag{9}$$

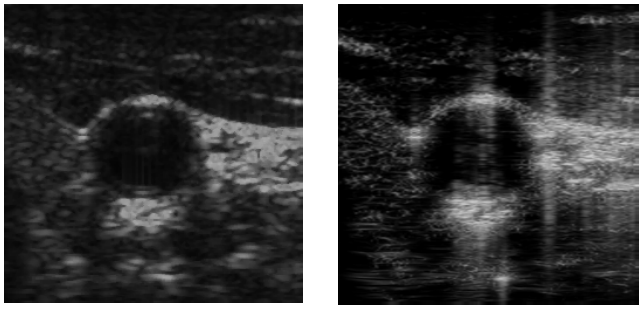
where ε is a small positive constant which is included to avoid dividing by zero, and T is a noise threshold.

The phase congruency calculation for a one dimensional signal has required the formation of a 90 degree phase shift of the original signal using odd-symmetric filters [34]. For the analysis of two dimensional data, we note that one cannot construct rotationally symmetric odd-symmetric filters, and therefore we must apply our one dimensional analysis over a number of separate orientations, and then combine the result to obtain a single measure of edge significance [32]. Kovesi *et al.* [34] suggested to construct a series of orientable 2D filters by spreading a Log-Gabor function into 2D. Thus, considering the one dimensional Log-Gabor filters defined earlier with geometrically increasing centre frequencies and bandwidths, we now mask these with an angular Gaussian tuned to a particular orientation ϕ_0 as follows [32]:

$$G(\omega_r, \phi) = \exp^{-\left(\frac{(log(\omega_r/\omega_{r0}))^2}{2(log(\kappa/\omega_0))^2} + \frac{(\phi-\phi_0)^2}{2\sigma_\phi^2}\right)} \tag{10}$$

where σ_ϕ defines the standard deviation of the Gaussian spreading function in the angular direction.

From these filters, Kovesi *et al.* [36] proposes a phase-based measure called *feature asymmetry* (FA), which combines



(a) B-Mode image (b) Phase Congruency Map

Fig. 2: A Transverse B-Mode image of a Carotid Artery, and the corresponding Phase Congruency Map

the orientation information and identifies step-like edges. The 2D FA measure proposed by Kovesi was first applied to echocardiography images by Mulet Parada *et al.* in [32], whereby it was extended to a 2D + time measure for boundary detection. Their method showed an improvement in the number of spurious feature responses due to speckle. We opt to use instead the maximum moments of phase congruency, proposed also by Kovesi in [37] as an indication of feature significance. The maximum moment of phase congruency is obtained by computing the Phase Congruency Covariance Matrix for each point in the image as follows:

$$G = \begin{bmatrix} \sum PC_x^2 & \sum PC_x PC_y \\ \sum PC_x PC_y & \sum PC_y^2 \end{bmatrix} \tag{11}$$

where PC_x and PC_y are the x and y components of the Phase Congruency $PC(x)$, for each orientation. The maximum moment may be obtained by taking the maximum singular value of the covariance matrix G. This provides us with a Phase Congruency Map, an example of which may be noted in Figure 2.

E. Deep Convolutional Neural Networks

A DCNN is a multilayer perceptron network which can exploit the stationary nature of natural images by learning features on locally connected pixels. The convolutional layers learn small features from small image patches sampled from the whole image [38]. The sub-sampling layers are used to reduce the computational complexity by summarising the statistics of a feature over a region in the image [38]. Our image segmentation task may be posed as a pixel-by-pixel classification problem, whereby a decision is made for each pixel - classifying it into 'foreground' or 'background'. The output of the network will therefore be a segmentation mask, ideally matching the manual segmentation (ground truth) provided by the expert. This may be defined as an optimisation problem, whereby we attempt to minimize the error between our output mask and the ground truth, by finding the optimal set of parameters, θ and b for the objective function below. First, assume we have a training set $\{(x^{(1)}, y^{(1)}), ..., (x^{(m)}, y^{(m)})\}$ of m training examples:

$$J_O(\theta, b) = \sum_{k=1}^{m} (h_{\theta,b}(x^{(k)}) - y^{(k)})^2 \tag{12}$$

The objective function $J_O(\theta, b)$ is the sum of differences between the predicted output segmentation mask $h_{\theta,b}(x^k)$ and the ground truth labels y^k, over all different training examples $k = 1, 2, 3...m$. The mask $h(x)$ is the result of a feed forward operation carried out through the network. Referring to Figure 3, we note that the full DCNN is constructed using two parts: an encoder section and a decoder section. The encoder is built using stacks of convolutional and subsampling layers, whereas the decoder is built using stacks of convolutional transpose layers and up-sampling layers. Typical CNNs normally resemble just the encoder structure, with their final layer being fully interconnected to a one-dimensional layer of nodes, before feeding on to the output. Such CNNs however have the inherent drawback of loss of image resolution, arising partially out of the convolution operation in the convolution layers, and partially out of the subsampling process designed to summarise the feature space. Since we intend to train our network in an end-to-end fashion, a decoder is appended to the end of the encoding CNN, to expand the extracted features back to full resolution, while concurrently calculating a probabilistic mask. The feedforward equation for a particular layer l in the encoder is provided by the following deterministic function $g^{(l)}$:

$$\begin{aligned} z_j^{(l)} &= g^{(l)}(z_i^{(l-1)}, w_{ij}^{(l)}, b_j^{(l)}) \\ &= \psi\left(\rho\left(\sum_{i=0}^{I} \tilde{w}_{i,j}^{(l)} * \ z_i^{(l-1)} + b_j^{(l)}\right)\right) \end{aligned} \tag{13}$$

where $z_j^{(l)}, j \in [1, F]$ is the j^{th} output feature map for layer l, calculated by convolving the trainable convolution filter $w_{ij}^{(l)}$ with the input to that layer $z_i^{(l-1)}$. The index i denotes the number of input maps available from the preceding layer, F denotes the number of filters, $b_j^{(l)}$ denotes the trainable bias term for layer l, $\tilde{w}$ denotes the flipped version of w [39], and $*$ denotes the convolution operator. The function $\rho(x)$ denotes the rectified linear activation function (ReLu), defined as $\rho(x) = max(0, x)$, whereas the function $\psi(x)$ is used to define the sub-sampling function. Subsampling functions normally implement either a *max pooling* function, whereby the maximum value from the preceding layer of local connections is passed onwards, or a mean pooling function, whereby the average is passed onwards to the next layer instead of the maximum. Within the decoder structure, the feedforward equation is provided by the function $h^{(l)}$ [40]:

$$\begin{aligned} y_j^{(l)} &= h^{(l)}(y_i^{(l-1)}, w_{ij}^{(l)}, b_j^{(l)}) \\ &= \rho\left(\sum_{i=0}^{I} w_{i,j}^{(l)} \Psi(y_i^{(l-1)}) + b_j^{(l)}\right) \end{aligned} \tag{14}$$

where $y_i^{(l-1)}$ in the first instance would be z from the preceding encoder layer. Thereafter it would be simply the output of the previous decoding layer. The function $\Psi(x)$

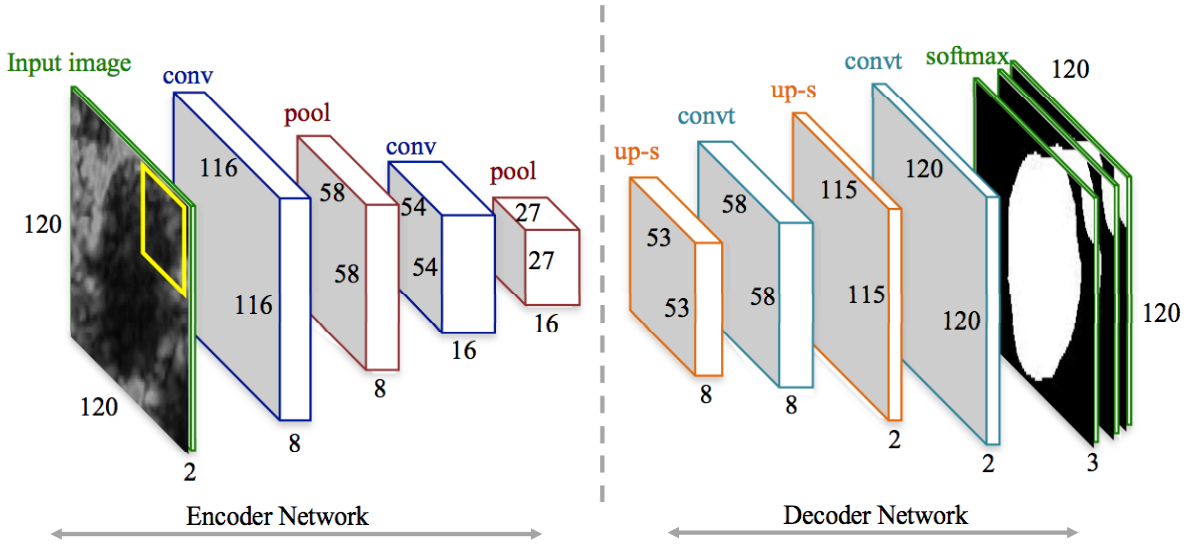


Fig. 3: A graphical representation of the encoder-decoder structure of the Deep Neural Network.

denotes an up-sampling operation, and the operator refers to the transposed full convolution. Each layer is once again followed by a ReLu function $\rho(x)$. At the end of the decoder network, the number of output maps are reduced to three, and fed into a softmax classifier, which provides logistic regression for a three-class problem [38]. The softmax function $\sigma(z)$ has the effect of maximising the maximum value of the outputs, making these close to 1, and the rest close to 0. A graphical representation of the DNN structure utilised is shown in Figure 3.

F. Modification of the Objective Cost Function

The objective function $J_O(\theta, b)$ is defined in equation (13) as the sum of differences between the predicted output segmentation mask $h_{\theta,b}(x^k)$ and the ground truth labels y^k, over all different training examples $k = 1, 2, 3...m$. The output segmentation mask $h_{\theta,b}(x^k)$ is itself defined as the output from the softmax function $\sigma(z)$ at the end of the DCNN. Thus:

$$(h_{\theta,b}(x^{(k)}) = \sigma(y_j^{(l)}) \tag{15}$$

where $y_j^{(l)}$ is defined from equation (16) as being the output of the network prior to softmax normalisation. In order to tune the objective function and make it more sensitive to the nature of the structures that are being segmented, we propose to modify the objective function with three additional cost terms which are defined as follows:

Curvature. The radius of curvature at any point on a curve is equal to the radius of the circular arc which best approximates the curve at the said point as shown in Figure 4. It is also defined as the inverse of curvature, and is expressed mathematically as follows:

$$R = \frac{1}{\kappa} \tag{16}$$

where κ is the curvature at a point. If we let the contour of the artery, from which these penalty terms are going to be derived, be expressed in the form of a curve in two dimensions as $u = f(v)$, then the curvature term κ itself may also be expressed as [41]:

$$\kappa = \frac{\left|\frac{d^2u}{dv^2}\right|}{\left[1 + \left(\frac{du}{dv}\right)^2\right]^{3/2}} \tag{17}$$

The penalty term we propose from the above measure of curvature κ is is defined as follows:

$$C(u) = \alpha\left(\max\left(|\kappa| - |\bar{\kappa}| - 0.5, 0\right)\right) \tag{18}$$

In equation (20), α is a scaling coefficient. The absolute value of κ is used because we may ignore the sign of the value, which is indicative of the direction of curvature, and which is not important for computation of the penalty term. The subtraction of the mean of κ is implemented so that the penalty term is made negligible if the contour is largely smooth and without notches, a situation which would generally yield similar values of κ all throughout the curve. Finally, we subtract a value of 0.5 because we may ignore curvature values below this value as simple noise. Noise values below 0.5, which yield negative values after this subtraction are in fact truncated to 0 by the max function to ensure a 0 penalty in this case.

Solidity. Solidity is a measure of morphological roughness and is sensitive to concavities in a shape or structure. It compares the pixel area of the object to the area of a bounding reference shape, which in this case would be the convex hull. Mathematically therefore, it may be expressed as:

$$SLD = \frac{A(u)}{A_c(u)} \tag{19}$$

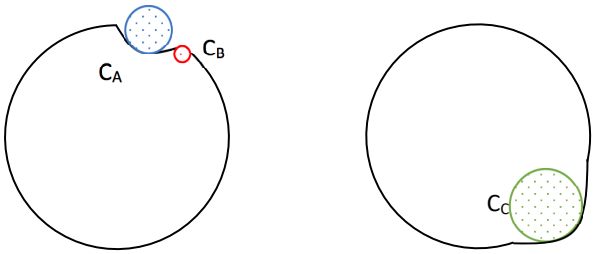


Fig. 4: [Left] A graphical representation of Curvature C_A and C_B, where the penalty at $C_B > C_A$ because the radius of the osculating circle at C_B is smaller than that of C_A. [Right] Curvature C_C is of opposite sign due to opposite direction of deformation.

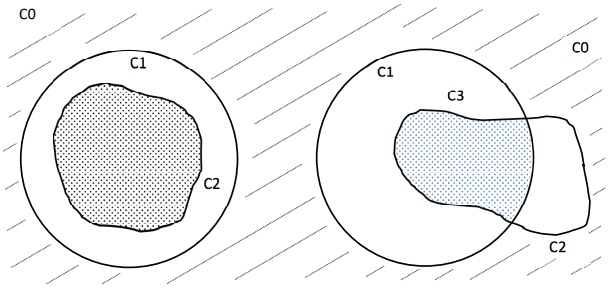


Fig. 5: A graphical representation of incorrect intersection between the MA outer boundary and LI inner boundary.

where $A(u)$ is the area of the object and $A_c(u)$ is the convex area of the shape in question. A solidity of 1 would indicate a perfectly solid shape with an area which is equal to convex area. An irregular shape with concavities would present with a solidity which is < 1 due to the area being smaller than the convex area. With regards to defining a penalty term, the media adventitia boundary of the carotid is expected to be smoothly circular, without any irregular concavities within the perimeter. Thus we would seek to penalise the objective function in instances where the solidity of the segmented mask would again deviate significantly from 1. The cost term we propose is therefore:

$$S(u) = \beta \left(\frac{A_c(u)}{A(u)} - 1 \right) \tag{20}$$

where β is a scaling coefficient and the inverse of SLD is used to have an increasing term in proportion to increased concavities in shape.

Intersection. The third penalty term we introduce to the objective cost function is a term which penalises the instances whereby the lumen intima boundary would not be contained within the media adventitia boundary, as is shown in Figure 5. This is in fact a circumstance which may not anatomically occur, and therefore its occurrence is penalised heavily to reduce the possibility of the neural network from converging to such a result. Referring to Figure 5, if we consider the MAB to be defined as contour C_1, the LIB to be defined as contour C_2, and their joint intersection as contour C_3, it easily follows that any occurrence whereby $Area_{C_2} \neq Area_{C_3}$ should be penalised. The proposed penalty term is therefore defined as:

$$I(u) = \gamma \left(\frac{Area_{C_2} - Area_{C_1 \cap C_2}}{Area_{C_0}} \right) \tag{21}$$

where γ is a scaling coefficient used to scale the normalised summation of pixels outside C_1 to a suitable magnitude of penalty term.

The overall new objective cost function, $J_M(\theta, b)$, may therefore now be represented in the following manner:

$$J_M(\theta, b) = \sum_{k=1}^{m} (h_{\theta,b}(x^{(k)}) - y^{(k)})^2 + C(v) + S(v) + I(v) \tag{22}$$

or alternatively:

$$\begin{aligned} J_M(\theta, b) = & \sum_{k=1}^{m} (h_{\theta,b}(x^{(k)}) - y^{(k)})^2 \\ & + \alpha \left(\max \left(|\kappa| - |\bar{\kappa}| - 0.5, 0 \right) \right) \\ & + \beta \left(\frac{A_c(v)}{A(v)} - 1 \right) \\ & + \gamma \left(\frac{Area_{C_2} - Area_{C_1 \cap C_2}}{Area_{C_0}} \right) \end{aligned} \tag{23}$$

The final intent of a modified objective cost function as described in equation 25, is to firstly require the network to train further, if the cost function results in greater, penalised values, but also to feed back through the backpropagation algorithm a set node errors observed at the output, such that the network may direct correction efforts towards the right filter map nodes further in. Modifying the nature of the backpropagation algorithm itself is beyond the scope of this work, and thus a heuristic approach to feeding back the node errors at the output was implemented, and which allows different penalty coefficients to be applied locally to individual output maps. Furthermore, the modified objective cost function, $J_M(\theta, b)$, was implemented in the DCNN with the scaling coefficients α, β, and γ being expressed as a fraction of the unmodified, original cost function $J_O(\theta, b)$. The rationale behind this was to have the scaling coefficients vary proportionately according to the original cost function, thus applying heavy penalty terms in the beginning when the errors were large, and eventually having diminishing penalty terms when the network was beginning to converge. The scaling coefficients, α, β, and γ were set, after several experiments to: $\alpha = 0.26 J_O(\theta, b)$, $\beta = 0.33 J_O(\theta, b)$ and $\gamma = J_O(\theta, b)$. This allowed for the evaluation of the overall modified cost function $J_M(\theta, b)$ as follows, which served to guide convergence.

$$
\begin{aligned}
J_M(\theta,b) = &\sum_{k=1}^{m}(h_{\theta,b}(x^{(k)}) - y^{(k)})^2 \\
&+ 0.26 J_O(\theta,b)\left(\max\left(|\kappa| - |\bar{\kappa}| - 0.5, 0\right)\right) \\
&+ 0.33 J_O(\theta,b)\left(\frac{A_c(v)}{A(v)} - 1\right) \\
&+ J_O(\theta,b)\left(\frac{Area_{C_2} - Area_{C_1 \cap C_2}}{Area_{C_0}}\right)
\end{aligned} \tag{24}
$$

Then, to feed back information regarding the relative nodes producing errors at the output, an indication of which output nodes are misclassifying results is required first, followed by a manner in which to amplify and accentuate their error to the network. To implement this, an **apportionment** of the penalty scale coefficients α, β, and γ, deriving from how much we wish to penalise errors in the MAB and LIB is first required. Following experiments, we split up the original scaling coefficients, α, β, and γ into two sets of constituent scale coefficients for both MAB and LIB contours. Mathematically therefore:

$$
\begin{aligned}
\alpha &= \alpha_{MAB} + \alpha_{LIB} \\
\beta &= \beta_{MAB} + \beta_{LIB} \\
\gamma &= \gamma_{MAB} + \gamma_{LIB}
\end{aligned} \tag{25}
$$

The coefficients are thus split up and set at: $\alpha_{MAB} = 0.16 J_O(\theta,b)$, $\beta_{MAB} = 0.16 J_O(\theta,b)$ and $\gamma_{MAB} = 0$ for the outer MAB, and $\alpha_{LIB} = 0.10 J_O(\theta,b)$, $\beta_{LIB} = 0.16 J_O(\theta,b)$ and $\gamma_{LIB} = J_O(\theta,b)$ for the inner LIB. The smaller curvature penalty for the LIB was implemented to achieve a somewhat lesser level of smoothening interference on the LIB than the MAB - since the LIB must be allowed to retain less regular shapes than the MAB.

The separate sets of α, β, and γ allow the computation of two temporary, intermediary cost functions which we denote: $J_{MAB}(\theta,b)$ and $J_{LIB}(\theta,b)$ as follows:

$$
\begin{aligned}
J_{MAB}(\theta,b) = &\sum_{k=1}^{m}(h_{\theta,b}(x^{(k)}) - y^{(k)})^2 \\
&+ 0.16 J_O(\theta,b)\left(\max\left(|\kappa| - |\bar{\kappa}| - 0.5, 0\right)\right) \\
&+ 0.16 J_O(\theta,b)\left(\frac{A_c(v)}{A(v)} - 1\right)
\end{aligned} \tag{26}
$$

$$
\begin{aligned}
J_{LIB}(\theta,b) = &\sum_{k=1}^{m}(h_{\theta,b}(x^{(k)}) - y^{(k)})^2 \\
&+ 0.10 J_O(\theta,b)\left(\max\left(|\kappa| - |\bar{\kappa}| - 0.5, 0\right)\right) \\
&+ 0.16 J_O(\theta,b)\left(\frac{A_c(v)}{A(v)} - 1\right) \\
&+ J_O(\theta,b)\left(\frac{Area_{C_2} - Area_{C_1 \cap C_2}}{Area_{C_0}}\right)
\end{aligned} \tag{27}
$$

These are both generally larger than the original unmodified cost function $J_O(\theta,b)$. If we divide both temporary cost functions $J_{MAB}(\theta,b)$ and $J_{LIB}(\theta,b)$ by the unmodified cost function $J_O(\theta,b)$, the resultant ratio is a multiplier > 1 which gauges the proportionate increase in penalty due to classification errors in the output maps of the MAB and LIB respectively. Reverting back to the error maps at the output layers, the respective multipliers are applied to all node deltas in their respective error maps. These node deltas would correspond, spatially, to pixels found on the contour and pixels contained within the said contour for the LIB, and to pixels contained within an annular ring for the MAB. Then, error nodes which already had a significant error value, are correspondingly augmented by the multiplier. Conversely, error nodes which had small error values are augmented by the same factor, but their already small error values ensures that the change remains negligible. The augmented node errors are then fed back through the network via the backpropagation algorithm.

IV. Experimental Setup and Results

A. Experimental Setup

Given the size of the dataset available, the data was first augmented using the strategy described in section III C. A number of schemes were then employed in order to prevent over-fitting and in order to ensure validity of the results. Firstly, during the training of the neural network, a basic early-stopping technique described by Prechelt in [42] was implemented in order to account for over-fitting. In accordance with the technique described by Prechelt, the dataset was split into two sub-datasets: a training dataset and a validation dataset. This split was implemented in a ratio of 66% training data and 33% validation data. Training took place only on the training set, and the error was evaluated on the validation set after each epoch. Once the training is stopped, the weights that the network had in the previous training run are used. In our experiments, we train using stochastic gradient descent, with batches of 5 ultrasound images, a learning rate η of 5×10^{-6}, a momentum of 0.90 and a weight decay of 5×10^{-6}. The weights and bias terms were randomly initialised. The early stopping technique was applied after 20 to 30 epochs.

In addition to the early stopping technique, a 15-fold leave-one-out validation scheme was also utilised to have some assurances on the validity of the results. The data available from the acquisitions was split at patient-scan level during each iteration into two categories: training + validation datasets (described earlier for the early stopping technique) and testing datasets. The training + validation datasets were created by concatenating the various ultrasound images obtained from different patient scans, and then randomising their sequence. The testing dataset in each iteration was then always made up of a hold-out set of ultrasound images from a particular patient scan, which were not present as image in the training data set.

Two experiments were carried out to compare the performance of the network without the modified geometrically

TABLE I: Results for testing DICE coefficient and MHD across [15]-fold Leave-one-out validation

Method	Without modified objective function				With modified objective function			
	MAB DICE	**MAB MHD** in mm	**LIB DICE**	**LIB MHD** in mm	**MAB DICE**	**MAB MHD** in mm	**LIB DICE**	**LIB MHD** in mm
Average	0.92 $\pm$ 0.04	0.27 $\pm$ 0.16	0.89 $\pm$ 0.08	0.36 $\pm$ 0.34	0.94 $\pm$ 0.04	0.24 $\pm$ 0.16	0.91 $\pm$ 0.07	0.33 $\pm$ 0.33

constrained objective function, and that of the network with the said modification. Cross validation iterations were run in both scenarios. The DCNN was built using the MatConvNet toolbox and trained on an Intel Core i7 with a Geforce GT 650M video card. The segmentation results obtained from the DCNN were compared against a manually labelled ground truth dataset which were manually and independently traced with the assistance of 2 radiographers, and used as labelled training data. Each radiographer was asked to manually trace the image sets twice, with a period of 2 weeks in between sessions and an average across both labelling sessions was retained as the final ground truth.

A number of evaluation metrics were used in order to quantify the performance of the convolutional neural network, and in order to allow comparison against other methods in the literature:

Similarity. The similarity between the segmented result and the ground truth is computed using the Dice Coefficient of Similarity. This effectively gauges the degree of overlap between two boundaries.

Modified Hausdorff Distance. The Hausdorff Distance is a measure of distance between two point sets. It provides the largest mismatched points between two boundaries. The Modified Hausdorff distance [43] on the other hand finds the mean distance between two boundaries and is computed as follows:

$$MHD = max(d(A,B), d(B,A)), \tag{28}$$

where

$$d(A,B) = \frac{1}{N_a} \sum_{a \in A} d(a,B) \tag{29}$$

Since the HD is a measure of the largest mismatch between two boundaries, a single point can cause a high HD value. The MHD mitigates this effect.

Sensitivity. The Sensitivity term is defined as the ability of a method or test to correctly classify a sample as being diseased [44], or in other words, the probability of a test being positive when the disease is indeed present. A method which displays high sensitivity is considered reliable when it produces a result that is negative, since it rarely misdiagnoses instances where the disease is present.

Specificity. The Specificity term, is defined as the ability of a method or test to correctly classify a sample as being free from disease [44], or in other words, the probability of a test being negative when the disease is indeed absent. A method which displays high specificity may be considered to reliably exclude the presence of disease when this is in fact absent.

B. Results

As a novel application over our previous work in [28], the network was trained to identify the contours of both the media adventitia boundary as well as the lumen intima boundary contemporarily. Thus the segmentation performance was quantified individually for these two boundaries and reported in our results. The rationale behind this is because the problem of segmenting these two boundaries poses different levels of challenge for the network. The media adventitia boundary normally presents itself as an approximately circular structure with well defined contours. This is shown to be fairly consistent across a number of patient B-Mode scans, as evidenced in Figure 7. The lumen intima boundary on the other hand is prone to having an irregular shape, partly because of the plaque which may be sited between the walls of the intima, and partly because the walls of the intima might not have clearly defined contours. This naturally poses a greater challenge for the radiographer to manually label, and correspondingly, a greater challenge for the network.

In Table 1 we present the averaged performance metrics quantified for both training methods. The DICE coefficient and the Modified Hausdorff Distance defined previously are noted for both MAB and LIB. We observe that in agreement with the noted difficulty of segmentation task, the segmentation performance of the network in delineating the MAB is consistently higher than the performance for delineating the LIB. The average DICE coefficient for the MAB is approximately 3% higher for both instances of training methods. Correspondingly, the average modified hausdorff distance is noted to be approximately 0.09mm lower in the MAB than the LIB, for both training methods. The overall performance of the network is noted to be satisfactory, with DICE coefficients exceeding 90% in both instances of MAB and LIB.

We further observe the performance of the network itself using both training methods. The average DICE coefficients for MAB and LIB respectively with the modified cost function are 0.94 $\pm$ 0.04 and 0.91 $\pm$ 0.07. Whereas the average MHD coefficients for the same modified cost function are 0.24 mm $\pm$ 0.16 mm and 0.33 mm $\pm$ 0.33 mm. These figures are both improved in relation to the corresponding average DICE and MHD coefficients for the MAB and LIB using a standard objective function. This indicates that our proposed modification to the objective function, which imposes geometric constraints on the stochastic gradient descent function, correctly discourages the latter from assuming irregular segmentation results, particularly for the media adventitia boundary which we know

TABLE II: Results for testing Sensitivity (Sens) and Specificity (Spec) across [15]-fold Leave-one-out validation

Method	Without modified objective function				With modified objective function			
	MAB Sens	**MAB Spec**	**LIB Sens**	**LIB Spec**	**MAB Sens**	**MAB Spec**	**LIB Sens**	**LIB Spec**
Average	0.924 $\pm$ 0.056	0.961 $\pm$ 0.025	0.937 $\pm$ 0.029	0.968 $\pm$ 0.026	0.929 $\pm$ 0.051	0.964 $\pm$ 0.025	0.936 $\pm$ 0.035	0.972 $\pm$ 0.020

should have an approximately circular shape. The modification which penalises intersection of borders on the other hand prevents instances whereby the lack of clearly defined LIB border causes the network to result in incorrect interpretations.

The results noted in Table II provide further comparative analyses with additional performance metrics of Sensitivity and Specificity. Here we notice no particular consistency in improvement between the MAB and LIB boundaries, as we have instances where the sensitivity and specificity of the LIB are better than those recorded for the MAB, and vice versa. However we do also note a general improvement in average performance across sensitivity and specificity for the modified objective cost function, when this is compared to the unmodified objective cost function. The only exception is noted in the Sensitivity of the LIB, whereby the sensitivity with the unmodified objective cost function is marginally better. It may suffice to note again however that the performances for both MAB and LIB are in excess of 90% for both sensitivity and specificity. The images in Figure 7 show a qualitative result, comparing the outputs of the ground truth with that of the network with both optimisation methods. One may notice that the results obtained exhibit a good degree of similarity to the manual labelling, because indeed both network types generally achieve DICE similarity coefficients in excess of 90%. The third column shows degraded performance. The reason behind this is that this particular dataset exhibited a higher degree of occlusion and poor contrast in the delineation between the MAB And LIB borders. This naturally poses a harder than usual task for the neural network, which is reflected in its poorer performance. The second and fifth columns exhibit instances where the DCNN with the modified optimisation achieves quantitative and qualitative improvements. Here we observe a similar performance in the outer border but the inner border however exhibits improved performance quantitatively, which translates to a closer degree of congruence in Figure 7 between the labelled inner border and the corresponding result produced by the network with the modified objective function. Taking column 2 for instance, we note that the sharp notch produced by the network on the left of the LIB with the standard objective function is reduced and smoothed in the corresponding result produced by the modified network. Similarly, the kink visible at the top edge of the LIB in the 5th column is smoothed in the corresponding result produced by the modified network. If one notes columns 1 and 4 on the other hand, we may here notice improved smoothed performances on the outer MAB border at the top edge for column 1 and rightmost edge for column 4 respectively. These results are important, because although the 'general' DICE performance of both networks is still fairly high, with similar coefficients in straightforward cases, the geometric constraints imposed by the modified network serve to smooth the contours produced by the network. Of course, care must be taken to tune the effect of such geometric constraints to within limits whereby the smoothing influence they exert on the LIB is smaller and more subtle, to allow the latter to assume irregular shapes which are common of atherosclerosis.

The results noted in Table III show a comparative assessment of performance between our technique and the techniques of other studies whose work similarly addressed the problem of MAB and LIB segmentation. Table III shows that our average results yield similar performance to that of Ukwatta *et al.* in their study in [23] and [10], and superior performance to that of Yang *et al.* in [16] and that of Hossain *et al.* in their work of [9]. In their work, Ukwatta *et al.* propose a semi-automated technique based on a level-set method to segment the MA and LI interfaces, whereby the operator was asked to provide anchor points as high-level domain knowledge. This together with the incorporation of local and global image statistics with boundary separation-based constraints allowed for segmentation of the MA and LI interfaces. In comparison to this, the technique proposed in our work is a fully automated technique which requires no user intervention, barring the initial training on a manually segmented dataset which would have presumably been carried out once in the beginning with a large enough and representative dataset.

V. Conclusion

In this work, we have developed a novel system, based on a bimodal and geometrically constrained Deep Convolutional Neural Network, for segmenting both the Media Adventitia Boundary and the Lumen Intima Boundary in transverse carotid Ultrasound images, using a fully automated approach. We have combined the novel fusion of amplitude and phase data as a bimodal source of input data, and also developed a novel geometrically constrained objective function for the training of the Deep Convolutional Neural Network.

We have shown that our geometrically constrained Deep Convolutional Neural Network shows improvement of approximately 2% in terms of DICE coefficient of performance, in comparison to the DCNN with the standard objective cost function that we have developed in our previous work. Furthermore, we demonstrate that albeit the technique being fully automated and having a larger and more generalisable dataset, it retains a good performance of 94% and 91% for MAB and LIB borders respectively, and that this retains comparable performance to techniques from other studies which test on a much smaller dataset.

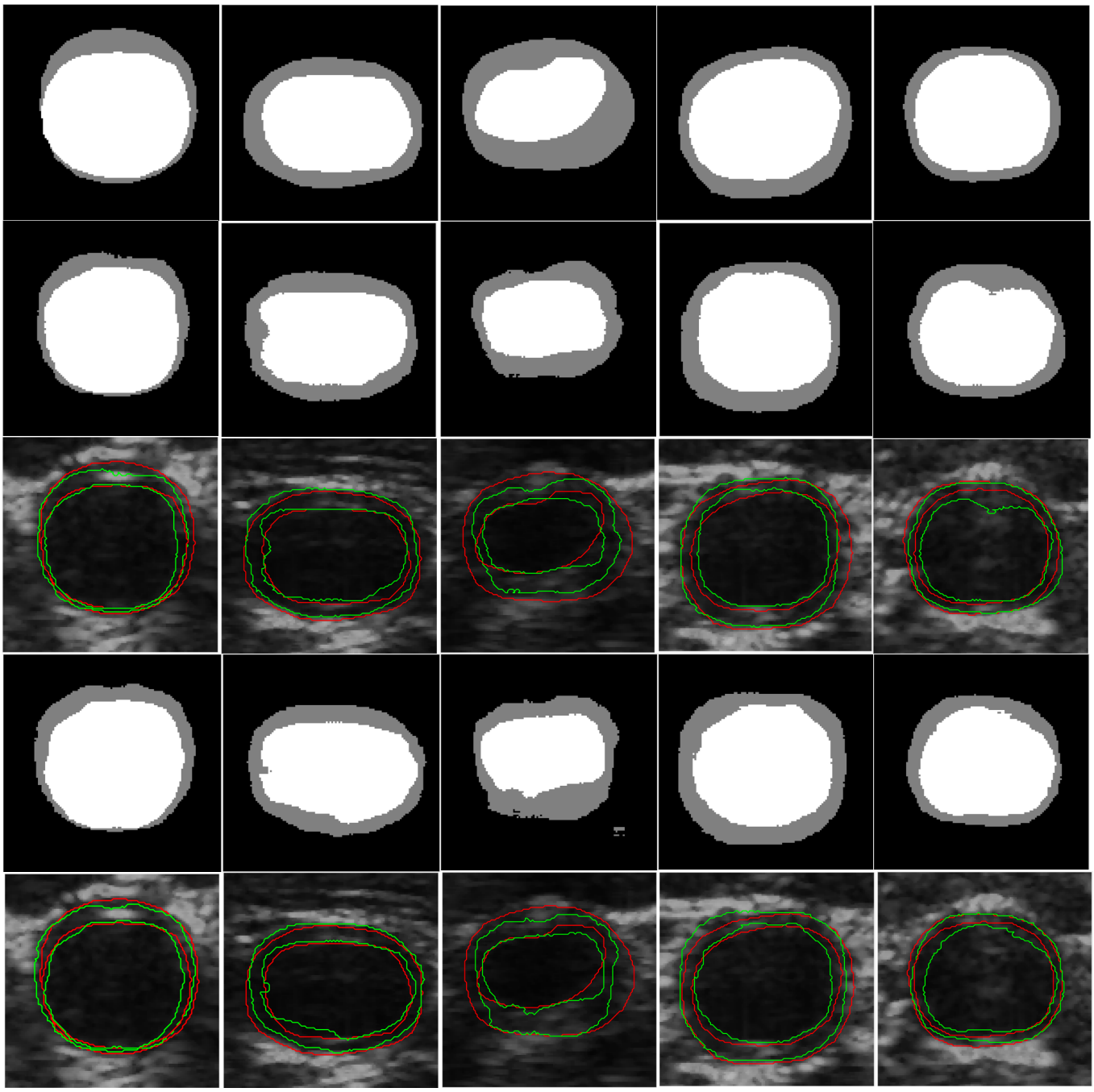


Fig. 6: A sample of segmentation results of outer MAB and inner LIB borders. [Row 1] - Averaged Ground truth; [Row 2] - Classification without optimisation; [Row 3] - Segmentation without optimisation; [Row 4] - Classification with optimisation; [Row 5] - Segmentation with optimisation. N.B red lines denote ground truth, green lines denote network results.

TABLE III: Results across different studies. Note that '-' denotes no results available from said study.

Study	**MAB DICE**	**MAB MHD** in mm	**LIB DICE**	**LIB MHD** in mm
Ukwatta *et al.* [10]	0.954 ± 0.016	- ± -	0.931 ± 0.031	- ± -
Yang *et al.* [16]	0.918 ± 0.035	- ± -	0.936 ± 0.026	- ± -
Ukwatta *et al.* [23]	0.95 ± 0.017	- ± -	0.92 ± 0.042	- ± -
Hossain *et al.* [9]	0.915 ± 0.035	0.25 ± -	0.735 ± 0.169	0.25 ± -
Azzopardi *et al.* [28]	0.988 ± 0.035	0.05 ± -	- ± -	- ± -
This study	0.940 ± 0.040	0.24 ± 0.16	0.910 ± 0.070	0.33 ± 0.33

REFERENCES

[1] E. J. Benjamin *et al.*, "Heart disease and stroke statistics - 2019 update: A report from the american heart association," *Circulation*, p. CIR.0000000000000659, 2019.

[2] A. Al-Mamari, "Atherosclerosis and physical activity." *Oman Med. J.*, vol. 24, no. 3, pp. 173–8, jul 2009.

[3] P. Libby, "Inflammation and Atherosclerosis," *Circulation*, vol. 105, no. 9, pp. 1135–1143, mar 2002.

[4] Z. Akkus, D. D. B. Carvalho, S. C. H. van den Oord, A. F. L. Schinkel, W. J. Niessen, N. de Jong, A. F. W. van der Steen, S. Klein, and J. G. Bosch, "Fully automated carotid plaque segmentation in combined contrast-enhanced and B-mode ultrasound," *Ultrasound Med. Biol.*, vol. 41, no. 2, pp. 517–531, 2015.

[5] A. Long, A. Lepoutre, Corbillon, and A. Branchereau, "Critical Review of Non- or Minimally Invasive Methods for Evaluating Stenosis of the Proximal Internal Carotid Artery," *Eur. J. Vasc. Endovasc. Surg.*, vol. 24, no. 1, pp. 43–52, jul 2002.

[6] B. Randoux, B. Marro, F. Koskas, M. Duyme, M. Sahel, A. Zouaoui, and C. Marsault, "Carotid artery stenosis: prospective comparison of CT, three-dimensional gadolinium-enhanced MR, and conventional angiography." *Radiology*, vol. 220, no. 1, pp. 179–85, jul 2001.

[7] J. D. Gill, H. M. Ladak, D. a. Steinman, and a. Fenster, "Accuracy and variability assessment of a semiautomatic technique for segmentation of the carotid arteries from three-dimensional ultrasound images." *Med. Phys.*, vol. 27, no. 6, pp. 1333–42, 2000.

[8] F. Molinari, G. Zeng, and J. S. Suri, "A state of the art review on intima-media thickness (IMT) measurement and wall segmentation techniques for carotid ultrasound." *Comput. Methods Programs Biomed.*, vol. 100, no. 3, pp. 201–21, dec 2010.

[9] M. M. Hossain, K. AlMuhanna, L. Zhao, B. K. Lal, and S. Sikdar, "Semiautomatic segmentation of atherosclerotic carotid artery wall volume using 3D ultrasound imaging." *Med. Phys.*, vol. 42, no. 4, p. 2029, 2015.

[10] E. Ukwatta, J. Awad, a. D. Ward, D. Buchanan, J. Samarabandu, G. Parraga, and a. Fenster, "Three-dimensional ultrasound of carotid atherosclerosis: semiautomated segmentation using a level set-based method." *Med. Phys.*, vol. 38, no. 5, pp. 2479–2493, 2011.

[11] F. Mao, J. Gill, D. Downey, and A. Fenster, "Segmentation of carotid artery in ultrasound images: Method development and evaluation technique," *Med. Phys.*, vol. 27, no. 8, p. 1961, 2000.

[12] P. J. Touboul, P. Prati, P. Y. Scarabin, V. Adrai, E. Thibout, and P. Ducimetière, "Use of monitoring software to improve the measurement of carotid wall thickness by B-mode imaging." *J. Hypertens. Suppl.*, vol. 10, pp. S37–S41, 1992.

[13] C. Liguori, A. Paolillo, and A. Pietrosanto, "An automatic measurement system for the evaluation of carotid intima-media thickness," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 6, pp. 1684–1691, 2001.

[14] J. H. Stein, C. E. Korcarz, M. E. Mays, P. S. Douglas, M. Palta, H. Zhang, T. LeCaire, D. Paine, D. Gustafson, and L. Fan, "A semiautomated ultrasound border detection program that facilitates clinical measurement of ultrasound carotid intima-media thickness," *J. Am. Soc. Echocardiogr.*, vol. 18, no. 3, pp. 244–251, 2005.

[15] J. C. R. Seabra, L. M. Pedro, J. F. E Fernandes, and J. M. Sanches, "A 3-D ultrasound-based framework to characterize the echo morphology of carotid plaques," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1442–1453, 2009.

[16] X. Yang, J. Jin, W. He, M. Yuchi, and M. Ding, "Segmentation of the common carotid artery with active shape models from 3D ultrasound images," B. van Ginneken and C. L. Novak, Eds., feb 2012, p. 83152H.

[17] A. Zahalka and A. Fenster, "An automated segmentation method for three-dimensional carotid ultrasound images," *Phys. Med. Biol.*, vol. 46, no. 4, pp. 1321–1342, 2001.

[18] D. chuan Cheng, A. Schmidt-Trucksäss, K. sheng Cheng, and H. Burkhardt, "Using snakes to detect the intimal and adventitial layers of the common carotid artery wall in sonographic images," *Comput. Methods Programs Biomed.*, vol. 67, no. 1, pp. 27–37, 2002.

[19] P. Abolmaesumi, M. Sirouspour, and S. Salcudean, "Real-time extraction of carotid artery contours from ultrasound\nimages," *Proc. 13th IEEE Symp. Comput. Med. Syst. CBMS 2000*, pp. 2–7, 2000.

[20] I. Solovey, "Segmentation of 3D Carotid Ultrasound Images Using Weak Geometric Priors," Master of Applied Science, University of Waterloo, 2010.

[21] J. Guerrero, S. E. Salcudean, J. A. McEwen, B. A. Masri, and S. Nicolaou, "Real-time vessel segmentation and tracking for ultrasound imaging applications," *IEEE Trans Med Imaging*, vol. 26, no. 8, pp. 1079–1090, 2007.

[22] M. M. Hossain, K. AlMuhanna, L. Zhao, B. K. Lal, and S. Sikdar, "Semi- automatic segmentation of atherosclerotic carotid artery lumen using 3d ultrasound imaging," *SPIE*, vol. 8869, p. 86694A, 2013.

[23] E. Ukwatta, J. Yuan, D. Buchanan, B. Chiu, J. Awad, W. Qiu, G. Parraga, and a. Fenster, "Three-dimensional segmentation of three-dimensional ultrasound carotid atherosclerosis using sparse field level sets." *Med. Phys.*, vol. 40, no. May, p. 052903, 2013.

[24] M. Egger, J. D. Spence, A. Fenster, and G. Parraga, "Validation of 3D Ultrasound Vessel Wall Volume: An Imaging Phenotype of Carotid Atherosclerosis," *Ultrasound Med. Biol.*, vol. 33, no. 6, pp. 905–914, 2007.

[25] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," *Iclr*, pp. 1–14, 2014.

[26] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60 – 88, 2017.

[27] R. M. Menchon-Lara and J. L. Sancho-Gomez, "Fully automatic segmentation of ultrasound common carotid artery images based on machine learning," *Neurocomputing*, vol. 151, pp. 161–167, 2015.

[28] C. Azzopardi, Y. A. Hicks, and K. P. Camilleri, "Automatic carotid ultrasound segmentation using deep convolutional neural networks and phase congruency maps," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, April 2017, pp. 624–628.

[29] D. Sheet, S. Pal, and A. Chakraborty, "Image quality assessment for performance evaluation of despeckle filters in Optical Coherence Tomography of human skin," *Iecbes, 2010 Ieee*, no. December, pp. 499–504, 2010.

[30] P. Simard, D. Steinkraus, and J. C. Platt, "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis," *Proc. 7th Int. Conf. Doc. Anal. Recognit.*, pp. 958–963, 2003.

[31] A. Belaid, D. Boukerroui, Y. Maingourd, and J.-F. Lerallut, "Phase-Based Level Set Segmentation of Ultrasound Images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 138–147, jan 2011.

[32] M. Mulet-Parada and J. a. Noble, "2D+T acoustic boundary detection in echocardiography." *Med. Image Anal.*, vol. 4, no. 1, pp. 21–30, 2000.

[33] M. Morrone and R. Owens, "Feature detection from local energy," *Pattern Recognit. Lett.*, vol. 6, no. 5, pp. 303–313, dec 1987.

[34] P. Kovesi, "Image Features from Phase Congruency," *Videre*, vol. 1, no. 3, 1999.

[35] S. Venkatesh and R. Owens, "An Energy Feature Detection Scheme," in *Int. Conf. Image Process.*, 1989, pp. 553–557.

[36] P. Kovesi, "Invariant Measures of Image Features from Phase Information," Ph.D. dissertation, University of Western Australia, 1996.

[37] ——, "Phase congruency detects corners and edges," in *Digital Image Computing: Techniques and Applications: Proceedings of the VIIth Biennial Australian Pattern Recognition Society Conference, DICTA 2003*, C. Sun, H. Talbot, S. Ourselin, and T. Adriaansen, Eds., vol. Vol 1., 2003.

[38] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network." *Med. Image Comput. Comput. Assist. Interv.*, vol. 16, no. Pt 2, pp. 246–53, 2013.

[39] T. Brosch, Y. Yoo, L. Y. W. Tang, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep Convolutional Encoder Networks for Multiple Sclerosis Lesion Segmentation," in *Lect. Notes Comput. Sci.*, ser. Lecture Notes in Computer Science, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, vol. 9351, pp. 3–11.

[40] C. Wang, X. Yan, M. Smith, K. Kochhar, M. Rubin, S. M. Warren, J. Wrobel, and H. Lee, "A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks." *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2015, pp. 2415–2418, 2015.

[41] E. Abbena, S. Salamon, and A. Gray, *Modern Differential Geometry of Curves and Surfaces with Mathematica, Third Edition*, ser. Textbooks in Mathematics. Taylor & Francis, 2006.

[42] L. Prechelt, *Early Stopping — But When?* Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 53–67.

[43] M. Dubuisson and A. K. Jain, "A modified hausdorff distance for object matching," in *Proceedings of 12th International Conference on Pattern Recognition*, vol. 1, Oct 1994, pp. 566–568.

[44] R. Parikh, A. Mathai, S. Parikh, G. Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values," *Indian journal of ophthalmology*, vol. 56, pp. 45–50, 03 2008.