Unsupervised domain adaptation method for segmenting cross-sectional CCA images

Computer Methods and Programs in Biomedicine 225 (2022) 107037

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Unsupervised domain adaptation method for segmenting cross-sectional CCA images

Luuk van Knippenberg[a,*], Ruud J.G. van Sloun[a,c], Massimo Mischi[a,c], Joerik de Ruijter[b], Richard Lopata[b,c], R. Arthur Bouwman[c]

[a] Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands
[b] Department of Biomedical Engineering, Eindhoven University of Technology, the Netherlands
[c] Department of Anesthesiology, Catharina Hospital Eindhoven, the Netherlands

ARTICLE INFO

Article history:
Received 16 March 2022
Revised 20 July 2022
Accepted 21 July 2022

Keywords:
Deep learning
Unsupervised domain adaptation
Vessel segmentation
Ultrasound

ABSTRACT

Background and Objectives: Automatic vessel segmentation in ultrasound is challenging due to the quality of the ultrasound images, which is affected by attenuation, high level of speckle noise and acoustic shadowing. Recently, deep convolutional neural networks are increasing in popularity due to their great performance on image segmentation problems, including vessel segmentation. Traditionally, large labeled datasets are required to train a network that achieves high performance, and is able to generalize well to different orientations, transducers and ultrasound scanners. However, these large datasets are rare, given that it is challenging and time-consuming to acquire and manually annotate *in-vivo* data.

Methods: In this work, we present a model-based, unsupervised domain adaptation method that consists of two stages. In the first stage, the network is trained on simulated ultrasound images, which have an accurate ground truth. In the second stage, the network continues training on *in-vivo* data in an unsupervised way, therefore not requiring the data to be labelled. Rather than using an adversarial neural network, prior knowledge on the elliptical shape of the segmentation mask is used to detect unexpected outputs.

Results: The segmentation performance was quantified using manually segmented images as ground truth. Due to the proposed domain adaptation method, the median Dice similarity coefficient increased from 0 to 0.951, outperforming a domain adversarial neural network (median Dice 0.922) and a state-of-the-art Star-Kalman algorithm that was specifically designed for this dataset (median Dice 0.942).

Conclusions: The results show that it is feasible to first train a neural network on simulated data, and then apply model-based domain adaptation to further improve segmentation performance by training on unlabeled *in-vivo* data. This overcomes the limitation of conventional deep learning approaches to require large amounts of manually labeled *in-vivo* data. Since the proposed domain adaptation method only requires prior knowledge on the shape of the segmentation mask, performance can be explored in various domains and applications in future research.

1. Introduction

Ultrasound imaging is a noninvasive, safe and inexpensive imaging modality that is commonly used to evaluate the cardiovascular system [1]. In particular, the common carotid artery (CCA), which provides blood to the neck and head, is one of the major arteries that is easily accessible by ultrasound. Accurately segmenting the artery in ultrasound images is required for purposes such as the assessment of risk for stroke due to atherosclerotic plaques [2–6], diameter estimation [7,8] and image-guided therapy [9].

Another application of vessel segmentation is Doppler angle estimation in cross-sectional ultrasound acquisitions, which is required for angle-corrected velocity estimates. By assuming that the vessel has a circular cross-section, as is common for major arteries such as the CCA [10], the flow can be calculated if the vessel diameter and average blood velocity are known. Doppler ultrasound is used to estimate velocity, where the velocity of red blood cells,

* Corresponding author.
E-mail address: l.a.e.m.v.knippenberg@tue.nl (L. van Knippenberg).

https://doi.org/10.1016/j.cmpb.2022.107037

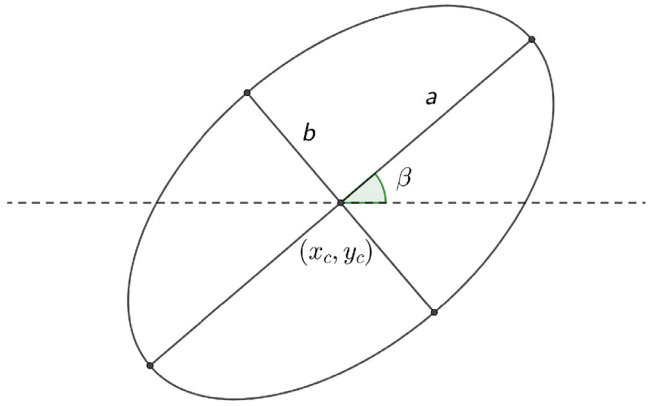

Fig. 1. The intersection between a cylinder and a plane is an ellipse, which can be represented using five parameters: The position of the center point of the ellipse (x_c, y_c), the semi-major axis a, the semi-minor axis b and the rotation of the ellipse β.

v, is proportional to the Doppler frequency:

$$v = \frac{c f_d}{2\cos(\alpha) f_0}, \tag{1}$$

with c being the speed of sound in tissue, f_d the Doppler frequency shift, α the Doppler angle (angle between the insonifying beam and blood flow direction) and f_0 the transmit frequency. Previously it was shown that cross-sectional Doppler results in accurate velocity estimates, without requiring a specific probe orientation [11]. In this case, the intersection of the ultrasound plane and the cylindrical vessel results in an ellipse, of which the parameters are used to estimate the Doppler angle:

$$\cos(\hat{\alpha}) = \cos(\arcsin(b/a))\cos(\beta), \tag{2}$$

where a and b are the semi-major and semi minor axis of the ellipse, respectively, and β is the rotation of the ellipse as depicted in Fig. 1. As a result, a proper vessel segmentation will result in an accurate Doppler angle estimate and therefore a more accurate velocity estimate. Note that the sensitivity of the Doppler angle estimate depends strongly on the probe orientation with respect to the vessel, and the ratio b/a and ellipse orientation β that result from that.

However, vessel segmentation in ultrasound images is a challenging problem due to the quality of the images, which is affected by attenuation, high level of speckle noise, and acoustic shadowing. In addition, the image strongly depends on the ultrasound scanner and the settings chosen by the operator (e.g. gain and focus depth). Moreover, the arterial wall contains three distinct separate layers, namely, intima, media and adventitia [12]. The outer boundary (media-adventitia) results in a strong echo, in contrast to the intima-media boundary that generates a weaker echo. Depending on the imaging orientation, the latter boundary will not always be visible. Manual segmentation requires substantial experience, is labor intensive and is prone to variability and operator dependency. Automatic segmentation methods that use conventional signal processing techniques are mostly based on edge detection, the Hough transform, active contours or a combination of these, as is summarized in multiple review papers [1,13,14].

Given that ideally the vessel interior appears as black on the B-mode image and the vessel wall appears as bright white, the vessel wall can be identified through edge detection. When imaging the vessel cross-sectionally, a seed point is selected within the vessel, from which a number of spokes are drawn (representing a star-shape) [15]. Along each spoke, an edge detector results in a single point on the vessel wall. These detected edges are then connected to identify the vessel wall. This method can be extended to detect edges using a scale-space representation [11] or include a Kalman filter to track the ellipse parameters over time (Star-Kalman algorithm) [5,16]. The segmentation accuracy for such a Star-algorithm is then mainly determined by the implemented edge detector and chosen seed point.

Alternatively, the Hough transform can be used to automatically detect mathematically defined shapes such as straight lines in longitudinal imaging and circles in transverse imaging [17]. The output of the Hough transform can also be used to initialize a snake to further improve the segmentation result [18], where the basic concept is to fit a contour to local image information, such as gradient. Typically, such parametric active contours are implemented by finding the contour that minimizes the sum of an internal energy (quantifying the continuity and smoothness of the contour) and external energy that attracts the contour to edges in the image [2,6,19–23]. As a result, these methods have the drawback that they require careful initialization and often require extensive tuning of parameters for convergence. Moreover, the snake may show leaking (contour bleeding) due to shadows or apparent gaps in the vessel wall. Also, due to the iterative nature of these algorithms, they are difficult to implement in real-time.

On the other hand, deep convolutional neural networks (CNNs) have shown great performance on image segmentation problems, including vessel segmentation [24–31]. In particular the U-net architecture [32] is commonly used in medical image analysis as it is fast and can precisely segment images in all major image modalities using a scarce amount of training data [33]. However, these works all use a supervised approach that requires extensive labeled datasets to successfully train a network that generalizes well. The problem here is that collecting and annotating datasets is an expensive and time-consuming process, and sufficient training data may not always be available. As a result, these networks often do not generalize well to novel datasets, for example captured using a different ultrasound scanner and settings.

Domain adaptation aims to solve this problem in which a model trained on a source distribution is used in the context of a different (but related) target distribution. One possible method to realize this is the use of an adversarial network, which discriminates between the source and target domain during training [34,35]. During training the aim is to maximize the loss of this domain classifier, such that the final classification decisions are made based on features that are both discriminative and invariant to the change of domains. However, this "minimax" approach is often unstable and therefore hard to train optimally.

In this work, we propose an unsupervised domain adaptation method that is based on three main contributions. First, a modified U-net architecture is described that includes model-based processing and can be trained end-to-end. In this case, the model-based processing consists of morphological closing, region splitting, connected component analysis, edge detection and a least-squares ellipse fit. Second, we describe how an a priori known shape can be used to detect unexpected outputs in a model-based approach, rather than relying on a second (adversarial) neural network to discriminate outputs. Third, we demonstrate good performance of a model trained using the proposed unsupervised domain adaptation method on a challenging *in-vivo* dataset. This paper is structured as follows: The network architecture and training process is described inSection 2. The data that was used in this work is presented in Section 3. The results are reported in Section IV and further discussed in Section 5. Finally, conclusions are drawn in Section 6.

2. Materials and methods

In this section, we describe the network architecture and training process.

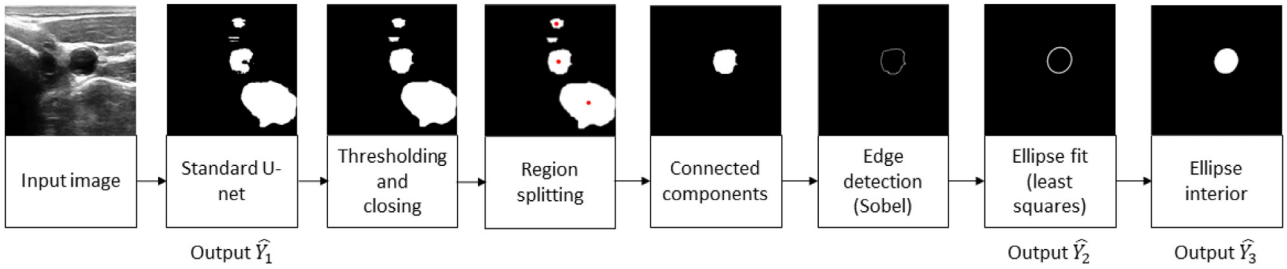

Fig. 2. Demonstration of the model-based processing steps that are appended to the final layer of the U-net architecture. The processing steps consists of thresholding probabilities (0.5), morphologically closing with a circular element, region splitting, connected components analysis, edge detection and ellipse fitting, resulting in an elliptical mask.

2.1. Network architecture

In this work, the basis of the network is a standard U-net architecture [32], consisting of a contracting path to capture context and a symmetric expanding path that enables precise localization. The input to the network is a B-mode image sized 256x256 pixels and the output of the U-net (after thresholding the probabilities) is a binary image, where for each pixel it is indicated whether that pixel belongs to the CCA (1) or not (0). From this semantic segmentation we then estimate the vessel diameter and Doppler angle. The prior knowledge that the prediction should be elliptical is included in the network by appending several basic image processing steps to the final layer of the network (output 1, $\hat{Y}_1$), as is summarized in Fig. 2.

First, the probabilities are thresholded (0.5) to obtain a binary image, after which a morphological closing operation with a circular element (R=8 pixels) is used to fill gaps in the U-net prediction. Next, a region splitting algorithm is applied that detects whether a connected component contains more than one ellipse center. If this is the case, the connected component is split into two regions. This procedure is further explained in Section II.B. Then connected-components analysis is used to identify the region in the binary image that most likely corresponds to the CCA. For this purpose, prior information on the expected size of the region is used (approximately 3000 pixels). Note that the expected size is based on the expected cross-sectional area as well as the pixel size, meaning that the size of the region (in pixels) also depends on the ultrasound settings. The remaining regions are discarded and a Sobel operator is used to find the edges of the selected connected-component (a normalized threshold of 0.90 was used to determine the final edge). Alternatively, the find-boundaries function from scikit-image can be used to identify the edge of the connected component directly. Lastly, an ellipse is fit to the detected edge points using a direct least squares approach. More specifically, note that an ellipse is a special case of a general conic that can be described by an implicit second order polynomial:

$$F(\boldsymbol{\theta}, \boldsymbol{X}) = \boldsymbol{\theta} \cdot \boldsymbol{X}$$
$$= Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

which has to satisfy the constraint $B^2 - 4AC < 0$ to be an ellipse. Here, $\boldsymbol{\theta} = [A, B, C, D, E, F]^T$ are the ellipse parameters and $\boldsymbol{X} = [x^2, xy, y^2, x, y, 1]$ are the coordinates of the points on this ellipse. Now, given N edge points $\boldsymbol{x_i} = [x_i, y_i]$, the ellipse that best fits these points minimizes the sum of pointwise distances. As the euclidean distance is difficult to evaluate, typically the algebraic distance, $F(\boldsymbol{\theta}, \boldsymbol{x_i})$, is used instead. As a result, the optimization problem reduces to:

$$\min_{\boldsymbol{\theta}} \sum_{i=0}^{N} F(\boldsymbol{\theta}, \boldsymbol{x_i}). \tag{3}$$

To prevent the trivial solution $A = B = C = D = E = F = 0$, the ellipse parameters are normalized with respect to F by setting $F = 1$ as described in [36]. Lastly, to improve numerical stability of the ellipse fitting method, the mean of the data is removed before fitting such that it is centered on the origin.

For completeness, the conversion from the parameters in the conic equation (A, B, C, D, E) to the standard ellipse parameters (x_0, y_0, a, b, β) are given in the Appendix. These ellipse parameters are the second output of the network ($\hat{Y}_2$) and are then used to generate a mask that corresponds to the interior of the ellipse, which is the final output of the network ($\hat{Y}_3$).

It can now be appreciated that the difference between the original output of the U-net ($\hat{Y}_1$) and the output of the fitted ellipse mask ($\hat{Y}_3$) can be used to quantify how elliptical the original prediction is, without requiring labels or human supervision. Ideally, the original output is a perfect ellipse, such that $\hat{Y}_1 = \hat{Y}_3$, and the difference is zero. Therefore, by minimizing this difference, the network learns to predict a single, completely filled ellipse, which should correspond to the common carotid artery.

2.2. Region splitting

One of the challenges in segmenting the CCA is the neighboring jugular vein, which can also show strong edges. As a result, the network may show activations in both, the artery and the vein, which can result in a single, large connected region as shown in Fig. 3 (c). The ellipse that is fit to this connected region therefore does not correspond well to the CCA. To prevent this, a region splitting algorithm inspired from [37] was implemented that aims to detect whether a connected component contains multiple ellipse centers, and splits the region in two if this is the case. First, in Fig. 3 (a), a Sobel edge detector is used to identify the edges in the binary image. Next, in Fig. 3 (b), the distance transform is used to calculate the distance from each pixel to the closest edge. A template is defined in Fig. 3 (d) that corresponds to the distance transform of a circle. Assuming that vessels can be approximated by an ellipse, the distance map will show high correlation with the circular template in the vessel centers as shown in Fig. 3 (e). A correlation threshold of 0.6 seemed to give reasonable results in most cases. Finally, if a connected component contains multiple centers, the region is split at its narrowest point along the line connecting the two centers, such that one region corresponds to the CCA and one region to the jugular vein (Fig. 3 (f)).

2.3. Labeled data - supervised training

The modified U-net described above was implemented in Tensorflow (v2.7) and can be trained end-to-end. To learn initial features that can be used to detect the CCA, the network is first trained in a supervised way on (mainly) simulated data, of which the data acquisition process is detailed in Section IV. Rather than training only towards the vessel mask using the standard binary cross-entropy loss, knowledge about the elliptical appearance of the vessel is included in training as well by using a custom loss

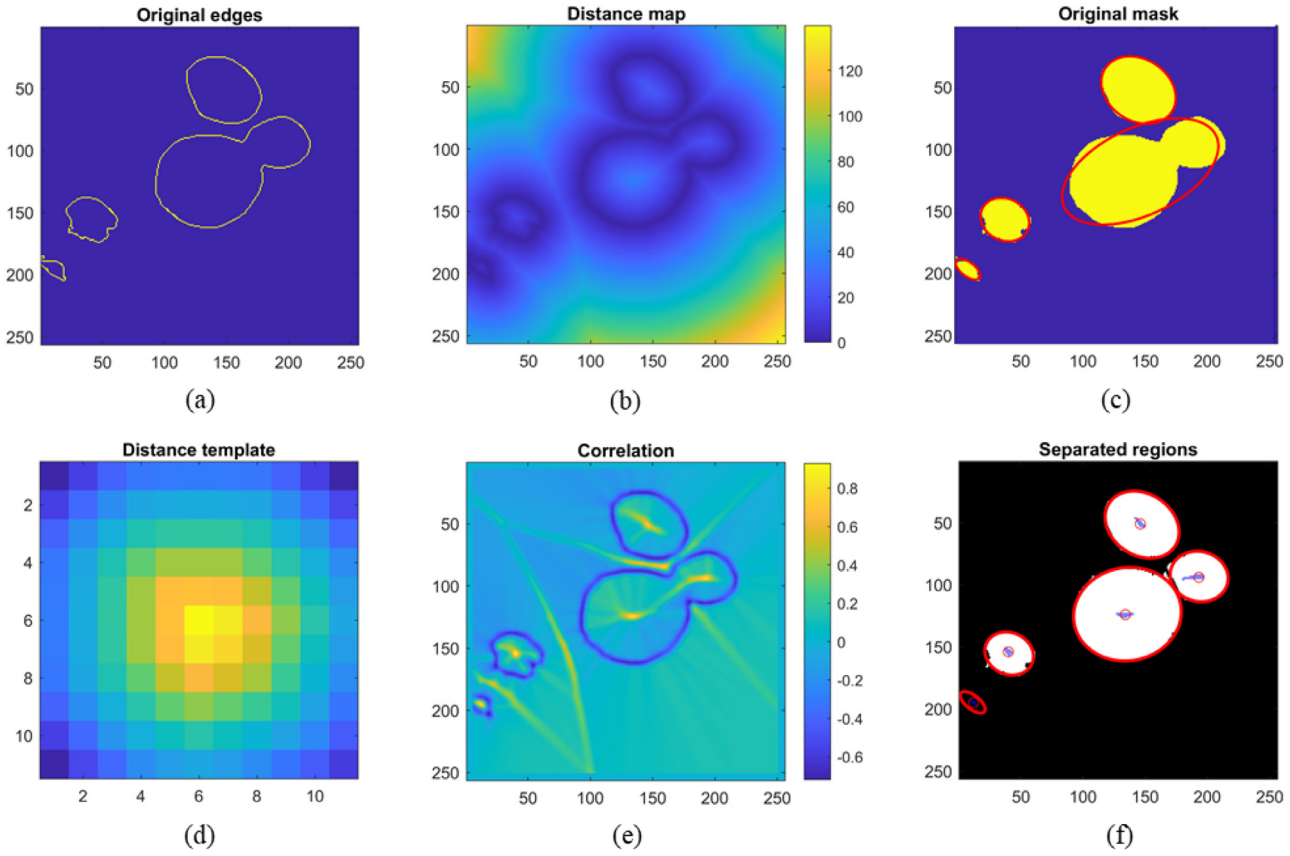

Fig. 3. An example where a connected component consists of two overlapping ellipses (c), which is split in two regions (f). The algorithm steps are edge detection (a), distance transform (b), correlation (e) with a template corresponding to the distance transform of a circle (d) and splitting the region at its narrowest point along the line connecting the two centers (f).

function. The loss function used in this work combines the binary cross-entropy (BCE) loss in $\hat{Y}_1$, Hausdorff distance loss in $\hat{Y}_2$ and soft Dice loss in $\hat{Y}_3$, which are detailed below.

First, a low BCE loss indicates that the network can determine with high certainty whether pixels belong to the CCA or not, and is therefore affected by the overall (global) segmentation performance. In contrast, the Hausdorff distance, d_H, is defined as the maximal shortest distance between two contours, and is therefore solely determined by the largest error. Lastly, the Dice similarity coefficient (DSC) measures the similarity between the interior of the fitted ellipse ($\hat{Y}_3$) and reference mask, enforcing elliptical predictions. More specifically, the DSC can be expressed in terms of pixel-based true positives (TP), false positives (FP) and false negatives (FN):

$$DSC = \frac{2TP}{2TP + FP + FN}. \tag{4}$$

In case of a perfect segmentation, there are no false positives or false negatives and the DSC reduces to one. The soft Dice loss function therefore aims to minimize $1 - DSC$.

Alternatively, the mean squared error (MSE) of the five ellipse parameters can be used to quantify how accurate the ellipse fit is. The effect of removing various elements from this custom loss function, as well as including the MSE, was measured in an ablation study. The various selected combinations are shown in the first column of Table 2.

In addition to the losses defined above, the accuracy (number of correctly classified pixels divided by the total number of pixels) is also measured of outputs $\hat{Y}_1$ and $\hat{Y}_3$, as well as the error in the Doppler angle estimation, computed from (2).

The network is trained for 20,000 epochs using a learning rate of 10^{-3}, where the weights that result in the best Doppler angle estimate are saved during training.

2.4. Unlabeled data - Unsupervised training

After learning initial features from (mainly) simulated data, unsatisfactory performance is still expected when the network is applied to unseen, *in-vivo* data. To adapt to this new domain, we use the prior knowledge on the shape of the prediction to continue training using an unsupervised approach. For this purpose, mixed batches are used consisting of labeled images that were used during the supervised learning stage and unlabeled *in-vivo* data. The new objective then is to minimize the loss of the labeled images (BCE) while at the same time minimizing the difference between the raw prediction ($\hat{Y}_1$) and corresponding ellipse fit ($\hat{Y}_3$) for the unlabeled images. Here, the difference is again quantified using the DSC, which in this case does not require labeled data. As a result, the network will still activate on the initially learned features (edges), while at the same time enforcing consistent, elliptical predictions on *in-vivo* data.

Note that the soft Dice loss can also be small if the number of activated pixels is low. In that case, the majority of the pixels will be classified as true positives even though the prediction does not agree well with the actual CCA segmentation. To prevent this, in addition to the labeled BCE loss and unlabeled soft Dice loss, a penalty is included for unfeasible, small ellipses. This exponential loss function depends on the number of activated pixels in the fitted elliptical mask (N) and therefore promotes region growing

Table 1
Data augmentation parameters.

Augmentation step	Probability	Description
Flip image	0.5	Flips the image (left-right)
Crop image	0.25	Remove outer pixels (0–30 pixels per side) and resize to original size
Adjust brightness	0.5	Reduce or increase brightness through PIL.ImageEnhance.Brightness() (factor between 0.2 and 1.8)
Gaussian noise	0.1	Add Gaussian noise with zero mean and a variance between 1 and 6% of the maximal pixel value
Speckle noise	0.1	Add speckle (multiplicative) noise with a variance between 0.1 and 0.5
Remove pixels	0.1	Randomly set between 0 and 80% of the image pixels to zero
Gaussian blur	0.05	Apply a Gaussian blur filter with a variance between 1 and 3

Table 2
Network performance on the test set of the labeled (mainly simulated) data using various combinations of loss functions during supervised training, consisting of binary cross-entropy loss (BCE), Hausdorff distance loss (HD), mean-squared error (MSE) and Dice similarity coefficient (DSC) loss.

	$\hat{Y}_1$: Raw U-net output		$\hat{Y}_2$: Ellipse parameters			$\hat{Y}_3$: Interior of fit ellipse	
Loss function	BCE loss	Accuracy (%)	d_H (px)	MSE	RMSE Doppler angle (°)	DSC	Accuracy
1. BCE	0.0038	99.84 ± 0.20	1.18 ± 1.84	5.13 ± 28.51	1.67 ± 3.29	0.988 ± 0.014	99.85 ± 0.24
2. BCE + HD	0.0039	99.84 ± 0.18	1.18 ± 1.27	4.29 ± 24.03	1.45 ± 2.79	0.988 ± 0.012	99.85 ± 0.17
3. BCE + MSE	0.0037	99.85 ± 0.15	1.16 ± 1.08	4.33 ± 25.46	1.44 ± 2.73	0.988 ± 0.010	99.86 ± 0.14
4. BCE + DSC	0.0035	99.86 ± 0.11	$\mathbf{1.09 \pm 0.69}$	$\mathbf{2.86 \pm 15.98}$	1.44 ± 2.65	$\mathbf{0.989 \pm 0.009}$	$\mathbf{99.87 \pm 0.10}$
5. BCE + HD + DSC	**0.0034**	$\mathbf{99.86 \pm 0.10}$	1.11 ± 0.68	3.66 ± 21.48	1.45 ± 2.87	$\mathbf{0.989 \pm 0.009}$	99.86 ± 0.10

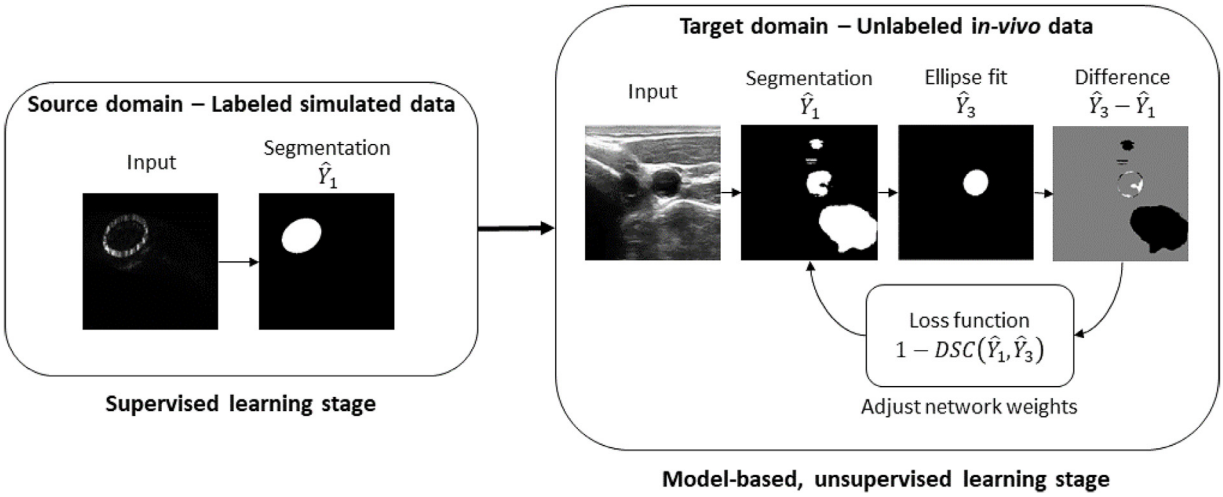

Fig. 4. Overview of the proposed unsupervised domain adaptation method showing images from the training set. First, a network is trained on labeled simulated data (source domain) using conventional supervised learning. Then the network is applied to unlabeled *in-vivo* data (target domain). By minimizing the difference between the output of the network, $\hat{Y}_1$, and the fitted ellipse, $\hat{Y}_3$, (i.e. maximizing the DSC) the network learns to predict a single, consistent ellipse that agrees with the previously learned features (edges).

until the predicted region has a reasonable size (at least 1000 pixels). Note that this penalty term only has the intended effect if the CCA is actually visible in all training images. The total loss function therefore is:

$$\mathcal{L}_{total} = \mathcal{L}_{BCE}(Y_1, \hat{Y}_1) + \mathcal{L}_{DSC}(\hat{Y}_1, \hat{Y}_3) + f(N), \tag{5}$$

where $f(N)$ is an exponential function. This iterative training process is illustrated in Fig. 4. The model is trained for 5000 epochs using early stopping with a patience of 200 epochs and a learning rate of 10^{-5}, where the total validation loss as defined in (5) is monitored for checkpointing and saving the optimal weights during training. Manually labeled data is used to quantify the performance gain in the DSC and Hausdorff distance due to domain adaptation. Statistical significance of the results is calculated using the one-sample Student *t*-test ($p < .05$).

2.5. Adversarial neural network

Lastly, results are compared to a more conventional domain adaptation approach as described in [34], where an adversarial neural network is used to predict from image features to which domain (source or target domain) a presented image belongs. By maximizing the binary cross-entropy loss for this second network, features that are present in images from both domains (i.e. domain-invariant features) are promoted. At the same time, the aim is to minimize the label prediction loss on the labeled images such that the extracted features result in overall good prediction performance on the source domain.

For this purpose, the U-net is modified to include the adversarial neural network. Rather than connecting the adversarial network to the feature extraction layer after the encoder-part of the network, the layer is connected to the output layer of the U-net. As a result, the adversarial network can detect unexpected outputs directly in the output layer, similar to the proposed domain adaptation method. The adversarial network consists of a gradient reversal layer that allows us to maximize instead of minimize the loss of the domain estimator, three convolutional layers, a global max pooling layer and an output layer containing the probability that an image belongs to each domain (2 nodes). The network was trained for 5000 epochs using a learning rate of 10^{-4}.

3. Data acquisition

As described above, two separate datasets are used during this work, where one dataset consists of mainly simulated images of which we have an accurate ground truth mask, and one dataset consists of only unlabeled *in-vivo* images.

3.1. Labeled data

In contrast to *in-vivo* data, simulated data has the advantage that it is easy to generate and has a known ground truth. The labeled dataset (2000 images in total) therefore consists of three parts. The first part (950 images) are augmented ellipses, the second part (950 images) are simulated ultrasound images of a bright vessel wall, and the final part (5% of the dataset) are *in-vitro* and *in-vivo* images (50 images each). The dataset is partitioned into a training, validation and test set using a 60:20:20 split ratio. The simulation steps are detailed below.

Ideally, the vessel wall appears as bright white on the B-mode image due to the strong reflections and the vessel interior is black due to the weak scattering of blood. To simulate this, first the ellipse parameters (x_0, y_0, a, b, β) are randomly generated, within specified bounds, such that the resulting ellipses are feasible and (mostly) within the imaging plane. The normalized ellipse center is uniformly generated between 0.2 and 0.8 both in the lateral and axial direction, the semi-major and semi-minor axis are between 0.1 and 0.3, and the ellipse can have an arbitrary orientation ($-90^\circ <= \beta <= 90^\circ$). These parameters are then used to generate two image masks, where one mask has a radius equal to b and the other mask has a radius equal to $b + d$ (where $2 <= d <= 5$ pixels). As a result, the difference between these two masks corresponds to the vessel wall and the interior of the vessel has a short axis length equal to b. To make the segmentation problem more challenging (and more realistic), Gaussian and speckle (multiplicative) noise are added, random pixels are removed and the images are Gaussian blurred.

Secondly, images are generated using a Verasonics (Verasonics Inc. Kirkland, USA) plane wave imaging simulation, where scatterers in the vessel wall have a high reflectivity (1.0) and scatterers within the vessel have a low reflectivity (0.1). The number of plane waves, transmit frequency, central steering angle and angle compounding range are chosen randomly, resulting in a large variety of simulated ultrasound images.

Lastly, *in-vitro* and *in-vivo* images acquired using a Philips EPIQ 7G ultrasound system (Philips, Eindhoven, NL) and L12-3 linear ultrasound probe (frequency range 3–12 MHz) are included in the dataset, which were used in a previous validation study [11]. These images were labeled manually and aim to facilitate the training process by ensuring that the network has a reasonable initial prediction on the unseen *in-vivo* data. Note that this data only represents a small portion of the dataset (5%) and was acquired using a different ultrasound system and transducer. The model that is trained on the full labeled dataset will be referred to as model A. To evaluate the effect of including some labeled *in-vivo* images during this supervised learning stage, the network is also trained on simulated data only, which will be referred to as model B.

In all cases, both the vessel mask (Y_1 and Y_3) and ellipse parameters (Y_2) are saved. A selection of images from the labeled dataset is shown in Fig. 5 (a).

3.2. Unlabeled data

On the other hand, *in-vivo* data is more challenging to acquire due to strict regulation. Moreover, labels are commonly obtained by manually segmenting the ultrasound images, which is time consuming and introduces variability in the ground truth. In this work, the ultrasound data from [5] is used, where a slow sweep across the subject's neck was performed using a clinical MyLabOne Vascular Ultrasound system (Esaote Europe, Maastricht, NL) with a linear probe SL3323 (fc = 7.5 MHz) to obtain a large variety of CCA images. This study was approved by the ethics committee of the local hospital (Catharina Ziekenhuis Eindhoven, the Netherlands). As the difference between subsequent frames is small, the frames are subsampled by a factor five, resulting in a subset of 2000 images from 67 acquisitions, of which some examples are shown in Fig. 5 (b). Again, a split ratio of 60:20:20 was used to partition the dataset into a training, validation and test set, where it was ensured that images from the same acquisition are only present in one dataset (training, validation or test set). As a result, the performance on the test set reflects the generalizability of the network to other acquisitions or patients.

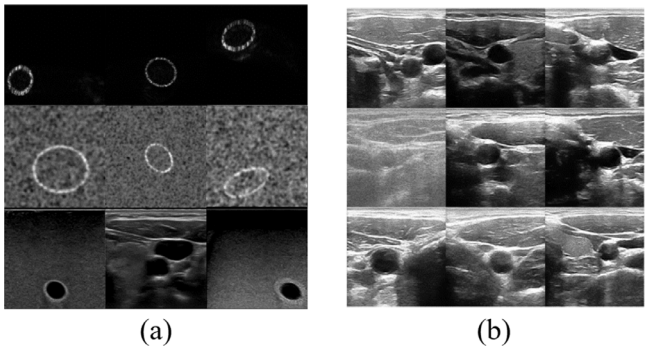

Fig. 5. Example images from both datasets (training set). (a): The labeled dataset consists of 2000 simulated ultrasound images (top row), noisy ellipses (center row) and very few (5% of the total dataset) *in-vitro* and *in-vivo* ultrasound images acquired using a Philips EPIQ 7G (bottom row). (b): The unlabeled dataset consists of *in-vivo* ultrasound images acquired using a MyLabOne Vascular Ultrasound system.

3.3. Data augmentation

Online data augmentation is a commonly used technique to synthetically increase the amount of training data and hence reduce overfitting on models. A generic approach for augmenting image data is to perform geometric augmentations such as cropping, reflecting, rotations and translating the image [38]. Note that we excluded rotations and vertical flipping as ultrasound images are depth-dependent and subject to the position of the probe with respect to the vessel. In addition to these common affine transformations, we included a small probability to modify the brightness, introduce noise (speckle and Gaussian), and/or apply Gaussian blurring. The probabilities and parameters used during augmentation are summarized in Table 1.

4. Results

4.1. Supervised training

The results of the supervised training stage using various loss functions are given in Table 2. The lowest binary cross-entropy loss is achieved when combining three loss function components (binary cross-entropy loss, Hausdorff distance loss and soft Dice loss), which also results in the highest DSC (0.989). The root-mean-square error in Doppler angle estimation is largest (1.67°) when only the binary cross-entropy loss is used during training, although the difference in accuracy and DSC is small compared to the best performing model (0.02% and 0.001, respectively). These minor differences between the different loss functions were found to be not statistically significant ($p > .05$). When only training on simulated

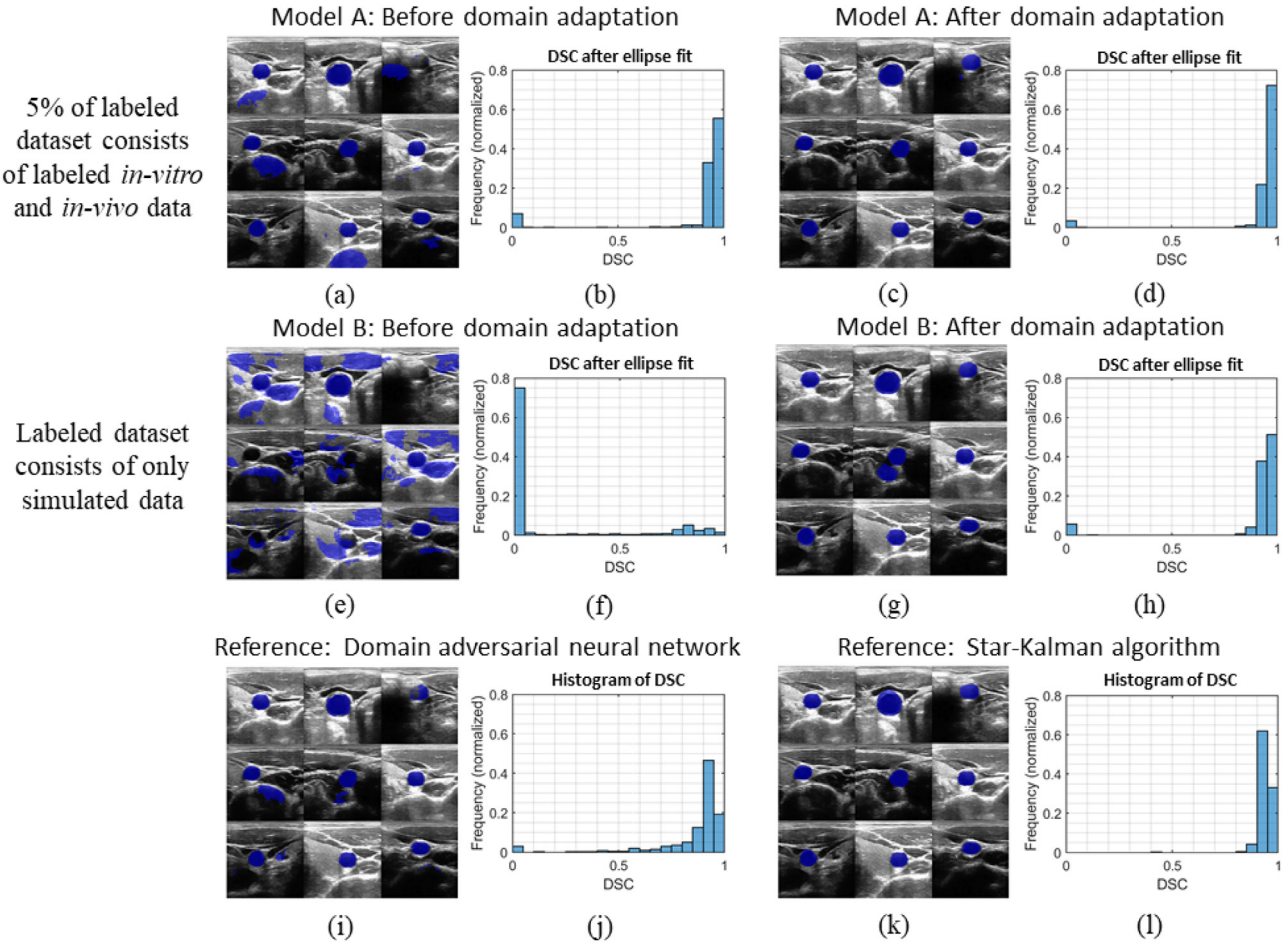

Fig. 6. Example images from the test set that show the improvement due to unsupervised domain adaptation for the network that was originally trained on a dataset containing 5% labeled *in-vitro* and *in-vivo* data (model A, left column) and for the network that was originally only trained on simulated data (model B, second column). (a) and (c) show the raw predictions before domain adaptation, where the distribution of the Dice similarity coefficient (DSC) after ellipse fitting (following the steps from Fig. 2) is shown in (b) and (d). Domain adaptation improves the predictions as shown in (g) and (i), which is also apparent from the distributions in (h) and (j). For comparison, the predictions obtained when using a domain adversarial neural network (trained on the same data as model A) and of the Star-Kalman algorithm [5] are shown in (e) and (f), and (k) and (l), respectively.

data (model B), using the loss function consisting of three components, the binary cross-entropy loss and error in Doppler angle are comparable to the performance of model A after 10,000 epochs (0.0035 and 1.11 degrees, respectively).

4.2. Unsupervised training

After the supervised learning stage, the performance of both models is evaluated using a test set consisting of *in-vivo* data, which was manually annotated by an expert. Note that this test set is therefore different from the test set from Table 2, where the test set consists of mainly simulated data. Examples of these predictions are displayed in Fig. 6 (a) and (e), where the distribution of the DSC is shown in plots (b) and (f). The model that has seen some *in-vivo* data during the supervised training stage (model A) already achieves a high DSC on most of the images (mean 0.877, median 0.953), whereas the elliptical predictions of the model that was trained using only simulated data (model B) generally have a low DSC (mean 0.173, median 0).

The unsupervised training stage stopped after 717 epochs for model A and after 387 epochs for model B. Fig. 7 shows the predictions of model B on an image from the training set at different points in the training process. Initially, the network shows activations in multiple image regions. Then the network predicts a single ellipse that does not respect the vessel edges, and finally the network predicts an elliptical vessel mask that agrees well to the boundaries of the vessel.

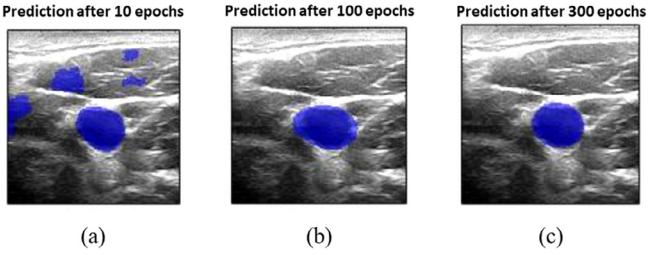

Fig. 7. Visualization of how the segmentation map changes for one input image (from the training set) during training. The predicted segmentation is shown after 10 (a), 100 (b) and 300 (c) epochs.

For both models, the unsupervised domain adaptation has a positive effect on the DSC, of which the predictions are shown in Fig. 6 (c) and (g). The distribution of the DSC is shown in plots (d) and (h). For model A, the median DSC has increased from 0.953 to 0.963, and the IQR decreased from 0.030 to 0.025. Moreover, the mean and median Hausdorff distance decreased from 17.1 to 10.3 and 2.8 to 2.5 pixels, respectively. For model B, the median DSC has increased from 0 to 0.951 with an IQR of 0.035 and median Hausdorff distance of 3.1 pixels. After removing outliers (de-

Table 3
Vessel segmentation performance of the network on the in-vivo test set before and after domain adaptation, compared to two reference methods. Performance is measured using the Dice similarity coefficient (DSC) and Hausdorff distance (d_H) .

Algorithm	DSC (mean ± std)	DSC (median ± IQR)	d_H (px) (mean ± std)	d_H (px) (median ± IQR)
Model A: Before domain adaptation	0.877 ± 0.251	0.953 ± 0.030	17.1 ± 84.3	2.8 ± 2.2
Model A: After domain adaptation	0.923 ± 0.183	$\mathbf{0.963 \pm 0.025}$	10.3 ± 38.9	$\mathbf{2.5 \pm 1.8}$
Model B: Before domain adaptation	0.173 ± 0.326	0 ± 0.050	108.3 ± 71.4	109.4 ± 67.0
Model B: After domain adaptation	0.890 ± 0.225	$\mathbf{0.951 \pm 0.035}$	14.1 ± 44.4	$\mathbf{3.1 \pm 2.1}$
Reference: Star-Kalman + (active contour [5])	0.938 ± 0.034	0.942 ± 0.027	3.8 ± 3.7	3.3 ± 1.6
Reference: Domain adversarial (neural network)	0.857 ± 0.190	0.922 ± 0.072	12.3 ± 34.2	3.8 ± 3.4

fined as predictions with $DSC < 0.5$), statistical significance is assessed by the Student t-test. The improvements compared to both reference methods in DSC for model A and B, as well as the improvement in Hausdorff distance for model A, are statistically significant ($p < .05$). These results are summarized in Table 3, including a comparison to the performance of the adversarial neural network (median 0.922) and the Star-Kalman algorithm described in [5] (median 0.942), of which the predictions and distributions are shown in Fig. 6 (i)-(j) and (k)-(l).

To evaluate the sensitivity of the connected-components analysis, both models were also trained for an expected region size of 2000, 4000 and 6000 pixels. The median DSC in all three cases is 0.960 for Model A and 0.944, 0.949 and 0.949, respectively, for model B. These differences are not statistically significant ($p > .05$).

5. Discussion

In this article, we proposed an unsupervised, model-based domain adaptation method that can be applied to segmentation problems where prior knowledge on the shape of segmentation is available. In our case, it was assumed that the common carotid artery has a circular cross-section, such that the intersection with the ultrasound plane is an ellipse. The parameters of this ellipse can be used to estimate the Doppler angle as in (2) to obtain angle-corrected velocity estimates. The training process consists of two stages, where we first train the network using a supervised approach on labeled data from the source domain, after which we use the elliptical properties of the segmentation map to continue training in an unsupervised way on unlabeled data from the target domain.

The proposed network architecture includes model-based processing operations such as morphological closing, region splitting, connected components analysis, edge detection and ellipse fitting that are appended to the output layer of the network, while the network can still be trained end-to-end. As a result, additional loss functions were used during the supervised training stage besides the traditional binary cross-entropy loss, based on the ellipse parameters (Hausdorff distance and MSE) or ellipse interior (DSC). Although the rate of convergence was similar for all loss functions, it was shown that a combination of three loss functions resulted in the highest DSC (0.9890) and lowest binary cross-entropy loss (0.0034). However, the differences between various combinations of loss functions is small, not statistically significant, and the optimal combination depends on the metric of interest. Due to limited computational resources, only a subset of combinations was tested where each model was trained for 20,000 epochs. To obtain conclusive results, the ablation study should be extended and the number of epochs should be increased to ensure that the loss has fully converged.

To evaluate the effect of using labeled data from the target domain during the supervised training stage, two models were trained. Model A was trained on a dataset containing 1900 simulated images and 100 real ultrasound images (50 *in-vitro* images and 50 *in-vivo* images), where it must be noted that this data was

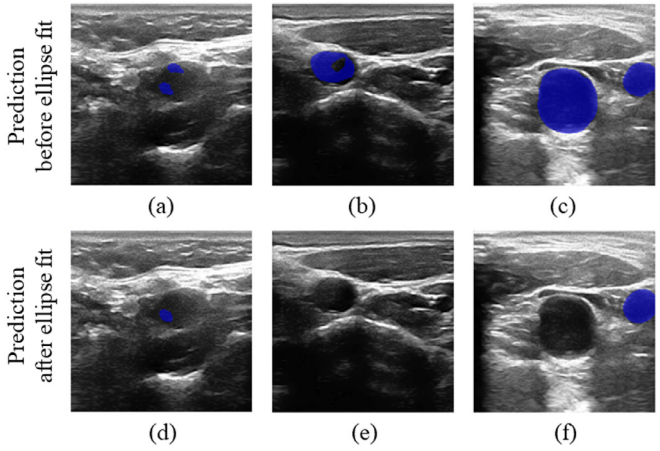

Fig. 8. Examples of images where the DSC is low. Main causes are a poor prediction before ellipse fit (a), a failed ellipse fit (e) while the prediction before ellipse fitting is reasonable (b), or the selection of the wrong connected component in the ellipse fitting process (f).

acquired using a different ultrasound scanner and transducer than the data that was used in the unsupervised training stage. In contrast, model B was trained on only simulated images. As a result, after the supervised learning stage model A showed good performance on unseen *in-vivo* data (mean DSC of 0.877), whereas the ellipse fit from model B was often empty or outside the CCA (mean DSC of 0.173).

From the distributions shown in Fig. 6, it can be concluded that some images are segmented poorly, resulting in a low DSC (< 0.50). The main causes of these outliers are a poor prediction before ellipse fitting, a failed ellipse fit, or selecting the wrong connected component. For example, for model B, 24 out of the 400 predictions have a DSC< 0.50, of which 8 have a poor prediction before ellipse fitting, 13 have a DSC of zero due to a failed ellipse fit, and 3 have a DSC of zero due to the selection of the wrong connected component in the ellipse fitting process. These examples are illustrated in Fig. 8. The fact that the ellipse fit sometimes fails is concerning and most likely caused by instabilities in the region splitting algorithm (which is the most complex step of the ellipse fitting process). Ideally, this step can be skipped during inference, assuming that the network has successfully learned to segment the CCA. On the other hand, the selection of the wrong connected component as in Fig. 8 (c) and (f) is a result of having a fixed expected size for the CCA. Although it was shown that this parameter has a negligible effect on the median DSC when increasing the expected region size from 2000 to 4000 and 6000 pixels, it would be better if this parameter can be learned from data and be changed adaptively as the surface area can strongly vary between datasets. In particular, the area is affected by imaging settings such as maximum depth and imaging orientation (e.g. imaging longitudinally results in a larger surface). Therefore, adaptively changing this parameter would benefit generalizability to a different or ex-

tended dataset. Nevertheless, these outliers have a large effect on the presented values for the mean and standard deviation. Given that the distribution is highly non-normal due to these outliers, it is more appropriate to evaluate the median and interquartile range, rather than the mean and standard deviation.

Due to the proposed unsupervised domain adaptation method, the performance of both models improved, obtaining a median DSC score of 0.963 and 0.951 for models A and B, respectively. The effect of using real ultrasound images in 5% of the labeled dataset therefore results in a better performance before and after domain adaptation. In addition, the training process of model A was more stable, showing good convergence where a lower loss also resulted in a higher DSC. On the other hand, early stopping was used for model B to prevent the model from obtaining a point where a lower loss was achieved whereas the DSC decreased. So although it is feasible to adapt a network that was trained on only simulated images to *in-vivo* data, it is recommended to include at least some *in-vivo* data in the labeled dataset if available for training robustness and improved performance. Alternatively, more realistic simulations can be used, where structures other than the CCA are included as well. This would reduce the domain shift from simulated to *in-vivo* data. For this purpose, a generative adversarial network (GAN) may be used, which is becoming more common and popular in medical image synthesis [39]. Synthetic data were generated using the Verasonics simulator. However, several alternative simulators, such as Field II [40,41] and SIMUS [42], are available which are also suitable for the same purpose.

The Star-Kalman algorithm combined with an active contour model as presented in [5] achieved good segmentation results with a median DSC of 0.942, which is illustrated in Fig. 6 (k) and (l). However, it must be noted that the algorithm was carefully fine-tuned specifically on this dataset. In addition, rather than analyzing images frame by frame, temporal information is included by tracking the ellipse parameters over time. As a result, accurate segmentations can be obtained on a poor image as long as the preceding images were of high quality. In contrast, the presented network analyzed a random subset of the data presented in [5] frame-by-frame, not leveraging temporal information. Still, both models outperform the Star-Kalman algorithm, obtaining a higher median DSC, without fine-tuning or human supervision. However, it must be noted that prior knowledge on the expected surface area of the common carotid artery (surface prior) was utilized in the selection of a connected component and to promote region growing during unsupervised training.

The proposed method was also compared to a more conventional domain adaptation method, where an adversarial neural network was trained that estimates to which domain (source or target) a specific input image belongs. This architecture therefore aims to minimize the loss for the labeled data while simultaneously maximizing the loss of the domain estimator. Ultimately, this adversarial network should result in indistinguishable outputs for images from both domains. However, the median DSC in this case (0.922) is lower than the DSC of the proposed method for both models, while being trained on the full labeled dataset. In addition, there is no suitable metric to monitor during training, given that the loss consists of a term that is being minimized and a second term that is being maximized. Due to this limitation of the adversarial network, we therefore directly monitored the DSC in this case to save the optimal model weights that requires the labels to be known.

The unsupervised training process is strongly based on the assumption that the vessel has a circular cross-section, such that the intersection with the ultrasound plane results in an ellipse. Although this assumption is in general true for major arteries, it does limit the use cases. For example, veins may have an arbitrary cross-sectional shape, which is also affected by the pressure applied to the transducer. In its current form, the framework therefore cannot be used to train a network that segments smaller (non-circular) arteries or even veins. Moreover, the current dataset does not include arteries with the presence of plaque, which would affect the shape of the interior of the vessel. As a result, it remains to be investigated how the network performs in these kind of images.

A limitation in this study was that the ultrasound images from the unlabeled dataset were acquired in the transverse plane (i.e. perpendicular to the vessel), where the Doppler angle is close to 90°. Although this is good for segmentation purposes, velocity estimation requires the Doppler angle to be smaller than 70°. As the vessel appears to be circular in these acquisitions, the ellipse parameters are not well defined, in particular the tilt of the ellipse. As a result, the error in Doppler angle estimation will be large when imaging around a Doppler angle of 90°, which is why the DSC and Hausdorff distance were used to quantify segmentation performance. Ideally, a large *in-vivo* dataset consisting of cross-sectional Doppler ultrasound images would be used, such that the error in Doppler angle can be quantified similar to the supervised training stage.

To have a consistent ground truth, the images were manually labelled at the media-adventitia boundary that results in a strong echo, in contrast to the intima-media boundary, which may not always be visible in B-mode images. Given that common carotid artery intima-media thickness ranges from 0.4 mm to 0.8 mm, whereas the lumen diameter ranges from 4.3 mm to 7.7 mm [43], the choice for which boundary is used in segmentation is relevant and depends on the application. For Doppler angle estimation, it is essential that the fit ellipse represents the (assumed) circular cross-section of the vessel wall, and therefore a well-defined boundary is preferred. The average velocity in this region can then be evaluated with Doppler ultrasound, resulting in an accurate flow estimate.

As discussed above, several improvements can be made to this study. First, the model-based processing should be optimized for inference after training the network to prevent poorly fitted ellipses or even a failed ellipse fit. Second, the labeled dataset can be extended with more realistic simulated ultrasound images to reduce the domain shift to the unlabeled *in-vivo* dataset, further facilitating domain adaptation. Third, the unlabeled dataset should be replaced with ultrasound images where the image orientation is appropriate for cross-sectional Doppler imaging, such that the estimated Doppler angle can be used to angle-correct velocity estimates. Fourth, when analyzing ultrasound loops instead of individual frames, temporal information should be utilized, for example by tracking the ellipse parameters using a Kalman filter as done in [5]. Lastly, other applications for the proposed domain adaptation method may be explored, emphasizing that the method can be applied to any problem where prior knowledge on the shape of the segmentation map is known. Possible examples include fetal head circumference estimation to assess fetal growth [44] and pupil tracking to measure the movement of the eye [45]. Alternatively, instead of using a strong prior such as an ellipse, a more general (polynomial) curve fitting approach can be used to only enforce a smooth contour.

6. Conclusion

The findings presented in this paper suggest that it is feasible to adapt a neural network trained on simulated images to segment the CCA in *in-vivo* ultrasound data without requiring labeled images, overcoming one of the largest limitations in training a neural network. Prior knowledge on the shape of the segmentation mask is utilized by including model-based processing in the neural network, which can then be used to train the network in an unsupervised way, outperforming a domain adversarial neural network

and conventional signal processing methods. However, by including labeled *in-vivo* data in the supervised training stage, training stability is improved and segmentation performance can be further increased. Since the proposed domain adaptation method only requires prior knowledge on the shape of the segmentation mask, performance can be explored in various domains and applications in future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is part of the research programme 17878, which is financed by the Netherlands Organization for Scientific Research (NWO). In addition, this study has been performed in the framework of the e/MTIC-program within the Eindhoven University of Technology in collaboration with Philips Research Eindhoven and the Catharina Hospital Eindhoven.

Appendix

The conversion from the parameters in the conic equation (A, B, C, D, E) to the standard ellipse parameters (x_0, y_0, a, b, β) are [46]:

$$x_0 = \frac{CD\prime - B\prime E\prime}{B\prime^2 - 4AC} \tag{6}$$

$$y_0 = \frac{AE\prime - B\prime D\prime}{B\prime^2 - 4AC} \tag{7}$$

$$a = \sqrt{\frac{2(AE\prime^2 + CD\prime^2 + B\prime^2 - 2B\prime D\prime E\prime - AC)}{(B\prime^2 - AC)(\sqrt{(A-C)^2 + 4B\prime^2} - (A+C))}} \tag{8}$$

$$b = \sqrt{\frac{2(AE\prime^2 + CD\prime^2 + B\prime^2 - 2B\prime D\prime E\prime - AC)}{(B\prime^2 - AC)(-\sqrt{(A-C)^2 + 4B\prime^2} - (A+C))}} \tag{9}$$

$$\beta = \begin{cases} 0 & \text{for } B = 0 \text{ and } A < C \\ \frac{\pi}{2} & \text{for } B = 0 \text{ and } A > C \\ \frac{1}{2}\operatorname{arccot}\left(\frac{A-C}{B}\right) & \text{for } B \neq 0 \text{ and } A < C \\ \frac{\pi}{2} + \frac{1}{2}\operatorname{arccot}\left(\frac{A-C}{B}\right) & \text{for } B \neq 0 \text{ and } A > C \end{cases} \tag{10}$$

Here, the prime symbol denotes a division by two (e.g. $B\prime = B/2$).

References

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

[2] L. Christodoulou, C.P. Loizou, C. Spyrou, T. Kasparis, M. Pantziaris, Full-automated system for the segmentation of the common carotid artery in ultrasound images, 5th Int. Symp. Commun. Control Signal Process. ISCCSP 2012 (2012), doi:10.1109/ISCCSP.2012.6217824.

[3] C.P. Loizou, S. Petroudi, M. Pantziaris, A.N. Nicolaides, C.S. Pattichis, An integrated system for the segmentation of atherosclerotic carotid plaque ultrasound video, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 61 (1) (2014) 86–101, doi:10.1109/TUFFC.2014.6689778.

[4] C. Qian, X. Yang, An integrated method for atherosclerotic carotid plaque segmentation in ultrasound image, Comput. Methods Program. Biomed. 153 (2018) 19–32, doi:10.1016/J.CMPB.2017.10.002.

[5] J. de Ruijter, M. van Sambeek, F. van de Vosse, R. Lopata, Automated 3D geometry segmentation of the healthy and diseased carotid artery in free-hand, probe tracked ultrasound images, Med. Phys. 47 (3) (2020) 1034–1047, doi:10.1002/MP.13960.

[6] E. Ukwatta, J. Awad, A.D. Ward, D. Buchanan, J. Samarabandu, G. Parraga, A. Fenster, Three-dimensional ultrasound of carotid atherosclerosis: semiautomated segmentation using a level set-based method, Med. Phys. 38 (5) (2011) 2479–2493, doi:10.1118/1.3574887.

[7] C.P. Loizou, T. Kasparis, C. Spyrou, M. Pantziaris, Integrated system for the complete segmentation of the common carotid artery bifurcation in ultrasound images, IFIP Adv. Inf. Commun. Technol. 412 (2013) 292–301, doi:10.1007/978-3-642-41142-7_30.

[8] A.C. Rossi, P.J. Brands, A.P. Hoeks, Nonlinear processing in B-mode ultrasound affects carotid diameter assessment, Ultrasound Med. Biol. 35 (5) (2009) 736–747, doi:10.1016/J.ULTRASMEDBIO.2008.10.011.

[9] E. Smistad, F. Lindseth, Real-Time automatic artery segmentation, reconstruction and registration for ultrasound-Guided regional anaesthesia of the femoral nerve, IEEE Trans. Med. Imaging 35 (3) (2016) 752–761, doi:10.1109/TMI.2015.2494160.

[10] U. Weber, N.J. Glassford, G.M. Eastwood, R. Bellomo, A.K. Hilton, A pilot assessment of carotid and brachial artery blood flow estimation using ultrasound doppler in cardiac surgery patients, J. Cardiothorac. Vasc. Anesth. 30 (1) (2016) 141–148, doi:10.1053/J.JVCA.2015.06.025.

[11] L. Van Knippenberg, R.J. Van Sloun, S. Shulepov, R.A. Bouwman, M. Mischi, An angle-Independent cross-Sectional doppler method for flow estimation in the common carotid artery, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 67 (8) (2020) 1513–1524, doi:10.1109/TUFFC.2020.2975315.

[12] J. Wikstrand, Methodological considerations of ultrasound measurement of carotid artery intima media thickness and lumen diameter, Clin Physiol Funct Imaging 27 (2007) 341–345, doi:10.1111/J.1475-097X.2007.00757.X.

[13] C.P. Loizou, A review of ultrasound common carotid artery image and video segmentation techniques, Med. Biol. Eng. Comput. 52 (12) (2014) 1073–1093, doi:10.1007/S11517-014-1203-5.

[14] V. Naik, R.S. Gamad, P.P. Bansod, Carotid artery segmentation in ultrasound images and measurement of intima-media thickness, Biomed Res. Int. 2013 (2013), doi:10.1155/2013/801962.

[15] D.C. Wang, R. Klatzky, B. Wu, G. Weller, A.R. Sampson, G.D. Stetten, Fully automated common carotid artery and internal jugular vein identification and tracking using B-mode ultrasound, IEEE Trans. Biomed. Eng. 56 (6) (2009) 1691–1699, doi:10.1109/TBME.2009.2015576.

[16] J. Guerrero, S.E. Salcudean, J.A. McEwen, B.A. Masri, S. Nicolaou, Real-time vessel segmentation and tracking for ultrasound imaging applications, IEEE Trans. Med. Imaging 26 (8) (2007) 1079–1090, doi:10.1109/TMI.2007.899180.

[17] S. Golemati, J. Stoitsis, E.G. Sifakis, T. Balkizas, K.S. Nikita, Using the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery, Ultrasound Med. Biol. 33 (12) (2007) 1918–1932, doi:10.1016/J.ULTRASMEDBIO.2007.05.021.

[18] J. Stoitsis, S. Golemati, S. Kendros, K.S. Nikita, Automated detection of the carotid artery wall in B-mode ultrasound images using active contours initialized by the hough transform, Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf. 2008 (2008) 3146–3149, doi:10.1109/IEMBS.2008.4649871.

[19] S. Delsanto, F. Molinari, P. Giustetto, W. Liboni, S. Badalamenti, J.S. Suri, Characterization of a completely user-independent algorithm for carotid artery segmentation in 2-D ultrasound images, IEEE Trans. Instrum. Meas. 56 (4) (2007) 1265–1274, doi:10.1109/TIM.2007.900433.

[20] J.D. Gill, H.M. Ladak, D.A. Steinman, A. Fenster, Accuracy and variability assessment of a semiautomatic technique for segmentation of the carotid arteries from three-dimensional ultrasound images, Med. Phys. 27 (6) (2000) 1333–1342, doi:10.1118/1.599014.

[21] A.K. Hamou, S. Osman, M.R. El-Sakka, Carotid ultrasound segmentation using DP active contours, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 4633 LNCS (2007) 961–971, doi:10.1007/978-3-540-74260-9_85.

[22] F. Mao, J. Gill, D. Downey, A. Fenster, Segmentation of carotid artery in ultrasound images: method development and evaluation technique, Med. Phys. 27 (8) (2000) 1961–1970, doi:10.1118/1.1287111.

[23] X. Yang, J. Jin, M. Xu, H. Wu, W. He, M. Yuchi, M. Ding, Ultrasound common carotid artery segmentation based on active shape model, Comput. Math. Methods Med. 2013 (2013), doi:10.1155/2013/345968.

[24] M. Biswas, V. Kuppili, L. Saba, D.R. Edla, H.S. Suri, A. Sharma, E. Cuadrado-Godia, J.R. Laird, A. Nicolaides, J.S. Suri, Deep learning fully convolution network for lumen characterization in diabetic patients using carotid ultrasound: a tool for stroke risk, Med. Biol. Eng. Comput. 57 (2) (2019) 543–564, doi:10.1007/S11517-018-1897-X.

[25] P.K. Jain, S. Gupta, A. Bhavsar, A. Nigam, N. Sharma, Localization of common carotid artery transverse section in B-mode ultrasound images using faster RCNN: a deep learning approach, Med. Biol. Eng. Comput. 2020 583 58 (3) (2020) 471–482, doi:10.1007/S11517-019-02099-3.

[26] J. de Ruijter, J.J. Muijsers, F.N. Van De Vosse, M.R. Van Sambeek, R.G. Lopata, A generalized approach for automatic 3-D geometry assessment of blood vessels in transverse ultrasound images using convolutional neural networks, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 68 (11) (2021) 3326–3335, doi:10.1109/TUFFC.2021.3090461.

[27] E. Smistad, L. Løvstakken, Vessel detection in ultrasound images using deep convolutional neural networks, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10008 LNCS (2016) 30–38, doi:10.1007/978-3-319-46976-8_4.

[28] S. Savaş, N. Topaloğlu, Ö. Kazc, P.N. Koşar, Classification of carotid artery intima

media thickness ultrasound images with deep learning, J. Med. Syst. 2019 438 43 (8) (2019) 1–12, doi:10.1007/S10916-019-1406-2.
[29] M.d.M. Vila, B. Remeseiro, M. Grau, R. Elosua, À. Betriu, E. Fernandez-Giraldez, L. Igual, Semantic segmentation with densenets for carotid artery ultrasound plaque segmentation and CIMT estimation, Artif. Intell. Med. 103 (2020), doi:10.1016/J.ARTMED.2019.101784.
[30] M. Xie, Y. Li, Y. Xue, R. Shafritz, S.A. Rahimi, J.W. Ady, U.W. Roshan, Vessel lumen segmentation in internal carotid artery ultrasounds with deep convolutional neural networks, Proc. - 2019 IEEE Int. Conf. Bioinforma. Biomed. BIBM 2019 (2019) 2393–2398, doi:10.1109/BIBM47256.2019.8982980.
[31] R. Zhou, A. Fenster, Y. Xia, J.D. Spence, M. Ding, Deep learning-based carotid media-adventitia and lumen-intima boundary segmentation from three-dimensional ultrasound images, Med. Phys. 46 (7) (2019) 3180–3193, doi:10.1002/MP.13581.
[32] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 9351 (2015) 234–241, doi:10.1007/978-3-319-24574-4_28.
[33] N. Siddique, P. Sidike, C. Elkin, V. Devabhaktuni, U-Net And its variants for medical image segmentation: theory and applications, IEEE Access (2020), doi:10.1109/ACCESS.2021.3086020.
[34] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, Adv. Comput. Vis. Pattern Recognit. 17 (9783319583464) (2015) 189–209, doi:10.1007/978-3-319-58347-1_10.
[35] E. Tzeng, J. Hoffman, K. Saenko, T. Darrell, Adversarial discriminative domain adaptation, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017 2017-Janua (2017) 2962–2971, doi:10.1109/CVPR.2017.316.
[36] P.L. Rosin, A note on the least squares fitting of ellipses, Pattern Recognit. Lett. 14 (10) (1993) 799–808, doi:10.1016/0167-8655(93)90062-I.
[37] T. Zou, T. Pan, M. Taylor, H. Stern, Recognition of overlapping elliptical objects in a binary image, Pattern Analy. Applic. 24 (2021) 1193–1206, doi:10.1007/S10044-020-00951-Z/FIGURES/11.
[38] L. Perez, J. Wang, The effectiveness of data augmentation in image classification using, Deep Learn. (2017).
[39] X. Yi, E. Walia, P. Babyn, Generative adversarial network in medical imaging: a review, Med. Image Anal. 58 (2019) 101552, doi:10.1016/J.MEDIA.2019.101552.
[40] J. Jensen, N. Svendsen, Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers, IEEE Trans. Ultrason. Ferroelectr. Freq. Contr. 39 (2) (1992) 262–267, doi:10.1109/58.139123. Copyright: 1992 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE
[41] J. Jensen, Field: a program for simulating ultrasound systems, Med. Biolog. Eng. Comput. 34 (sup. 1) (1997) 351–353. 10th Nordic-Baltic Conference on Biomedical Imaging ; Conference date: 09-06-1996 Through 13-06-1996
[42] A. Cigier, F. Varray, D. Garcia, SIMUS: An open-source simulator for medical ultrasound imaging. part II: comparison with four simulators, Comput. Method. Program. Biomed. 220 (2022) 106774, doi:10.1016/j.cmpb.2022.106774.
[43] Y.R. Limbu, G. Gurung, R. Malla, R. Rajbhandari, S.R. Regmi, Assessment of carotid artery dimensions by ultrasound in non-smoker healthy adults of both sexes, Nepal Med. Coll. J. 8 (3) (2006) 200–203.
[44] T.L. van den Heuvel, D. de Bruijn, C.L. de Korte, B. van Ginneken, Automated measurement of fetal head circumference using 2D ultrasound images, PLoS ONE 13 (8) (2018) e0200412, doi:10.1371/JOURNAL.PONE.0200412.
[45] T. Satriya, S. Wibirama, I. Ardiyanto, Robust pupil tracking algorithm based on ellipse fitting, 2016 Int. Symp. Electron. Smart Devices, ISESD 2016 (2017) 253–257, doi:10.1109/ISESD.2016.7886728.
[46] E.W. Weisstein, From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/Ellipse.htmlEllipse.2020https://mathworld.wolfram.com/Ellipse.html.