Vessel lumen segmentation in carotid artery ultrasounds with the U-Net convolutional neural network

1st Meiyan Xie
Department of Computer Science
New Jersey Institute of Technology
Newark, USA
mx42@njit.edu

2nd Yunzhi Li
Department of Computer Science
New Jersey Institute of Technology
Newark, USA
yl722@njit.edu

3rd Yunzhe Xue
Department of Computer Science
New Jersey Institute of Technology
Newark, USA
yx277@njit.edu

4th Lauren Huntress
Vascular and Endovascular Surgery
Robert Wood Johnson University Hospital
Newark, USA
huntrela@rwjms.rutgers.edu

5th William Beckerman
Vascular and Endovascular Surgery
Robert Wood Johnson University Hospital
Newark, USA
beckerwe@rwjms.rutgers.edu

6th Saum Rahimi
Vascular and Endovascular Surgery
Robert Wood Johnson University Hospital
Newark, USA
rahimisa@rwjms.rutgers.edu

7th Justin Ady
Vascular and Endovascular Surgery
Robert Wood Johnson University Hospital
Newark, USA
Jwa60@rwjms.rutgers.edu

8th Usman Roshan
Department of Computer Science
New Jersey Institute of Technology
Newark, USA
usman@njit.edu

***Abstract*—Carotid ultrasound is a screening modality used by physicians to direct treatment in the prevention of ischemic stroke in high-risk patients. It is a time intensive process that requires highly trained technicians and physicians. Evaluation of a carotid ultrasound requires segmentation of the vessel wall, lumen, and plaque of the carotid artery. Convolutional neural networks are state of the art in image segmentation yet there are no previous methods to solve this problem on carotid ultrasounds. We evaluate here a U-net convolutional neural network for lumen segmentation from ultrasound images of the entire carotid system. We obtained de-identified images under IRB approval from 226 patients. We isolated the internal, external, and common carotid artery ultrasound images for these patients giving us a total of 2156 images. We manually segmented the vessel lumen in each image that we then use as ground truth. With our convolutional U-Net we obtained a 10-fold cross-validation accuracy of 94.3%. We see that the U-Net correctly segments the lumen even in the presence of significant plaque, calcified wall, and ultrasound shadowing, all of which make it difficult to outline the vessel. We also see that the common carotid artery vessels are easiest to segment with a 96.6% cross-validation accuracy whereas internal and external carotid are harder both with 92.7% and 91.9% cross-validation accuracies respectively. Our work here represents a first successful step towards the automated segmentation of the vessel lumen in carotid artery ultrasound images and is an important first step in creating a system that can independently evaluate carotid ultrasounds.**

***Index Terms*—vessel segmentation, plaque segmentation, convolutional U-Net, medical AI**

I. Introduction

Stroke is the 5th leading cause of death in the United States [1]. Annually, it is responsible for billions of dollars in lost income and health care costs. For this reason there is significant effort and investment in the prevention of stroke. Ischemic strokes account for 87% of all strokes. Narrowing and deposition of plaque in the carotid arteries due to atherosclerosis is the most common cause of ischemic stroke. Carotid ultrasound is a safe, low-cost procedure that is used as a screening test in patients with risk factors for atherosclerosis [2]. It allows physicians to stratify the stroke risk of a patient and identify those patients that will most benefit from medical therapy or surgical intervention.

During a vascular ultrasound high-frequency sound waves are transmitted into your body. The sound waves are reflected back to the probe when they encounter the boundaries between different tissues in the body. This information is then utilized to create a 2D image of the vessel and surrounding tissue structures. Physicians utilize ultrasound images of the carotid artery in stroke prevention. During their evaluation physicians must first identify the vessel in the image. They then identify any atherosclerotic plaque within the wall and lumen of the vessel and finally they evaluate the physiologic impact of those plaques on the flow of blood within the vessel. This is a time intensive and resource intensive process that requires highly skilled technicians and physicians to perform and interpret the results. As physician workload has increased and healthcare systems investigate ways to streamline processes and cut costs automating the interpretation of vascular ultrasounds has great potential.

Prior work in automated approaches to evaluating carotid ultrasounds is highly limited and there are no prior methods for vessel segmentation in carotid ultrasounds of the entire carotid system. A convolutional U-Net for 2D ultrasounds like ours was explored in previous work [3] but only for internal carotid artery ultrasounds. In this work we explore a U-Net for the entire carotid artery system that includes internal, external, and common carotid arteries. We also explore images containing bifurcations, longitudinal images, and images with ultrasound shadowing, plaque, and gray shading, all of which make vessel segmentation even harder. Thus our work has a much broader scope than the previous study.

Vessel identification in carotid ultrasounds with preprocessing and marker-controlled watershed transform has been explored previously [4]. DeepVesselNet [5] is a deep learning model designed for vessel detection but in 3D magnetic resonance angiography data unlike the 2D ultrasounds that we consider here. A patch-based deep learning solution has also been proposed segmenting and measuring plaque for 3D ultrasounds [6]. Of note, 3D ultrasound is available only in research studies and is not commonly utilized clinically. In contrast in our study is a full end-to-end trainable convolutional network that allows for the segmentation of 2D ultrasounds, the most widely utilized modality.

II. Methods

A. Data collection

We obtained IRB approval from Robert Wood Johnson Medical School to use de-identified images from the Department of Vascular Surgery for this research. We manually downloaded B-mode carotid ultrasound examinations of 226 patients. We utilized an automated script to crop all patient identities from the ultrasound images and manually verified this de-identification.

We then cropped each image to obtain just the ultrasound removing all text and annotations on the image. Each images was resized to 224x224 pixels. This gave us a total of 2156 images that we then manually segmented. Using RectLabel software (https://rectlabel.com/) we manually segmented the vessel lumen for each image to serve as ground truth for training and validation

B. Background

1) Convolutional neural networks: Convolutional neural networks are the current state of the art in machine learning for image recognition [7], [8], including for MRI [9]. They are typically composed of alternating layers for convolution and pooling, followed by a final flattened layer. A convolution layer is specified by a filter size and the number of filters in the layer. Briefly, the convolution layer performs a moving dot product against pixels given by a fixed filter of size $k \times k$ (usually 3×3 or 5×5). The dot product is made non-linear by passing the output to an activation function such as a sigmoid or rectified linear unit (also called relu or hinge) function. Both are differentiable and thus fit into the standard gradient descent framework for optimizing neural networks during training. The output of applying a $k \times k$ convolution against a $p \times p$ image is an image of size $(p-k+1) \times (p-k+1)$. In a CNN, the convolution layers just described are typically alternated with pooling layers. The pooling layers serve to reduce dimensionality, making it easier to train the network.

2) Convolutional U-Net: After applying a series of convolutional filters, the final layer dimension is usually much smaller than that of the input images. For the current problem of determining whether a given pixel in the input image is part of a vessel or plaque, the output must be of the same dimension as the input. This dimensionality problem was initially solved by taking each pixel in the input image and a localized region around it as input to a convolutional neural network instead of the entire image [10].

A more powerful recent solution is the Convolutional U-Net (U-Net) [11]. This has two main features that separate it from traditional CNNs: (a) deconvolution (upsampling) layers to increase image dimensionality, and (b) connections between convolution and deconvolution layers.

3) U-Net for vessel segmentation: We implemented a basic U-Net [11] in the Pytorch library [12] as shown in Figure 1. The U-Net is a popular choice for medical artificial intelligence work and has proven to be a successful baseline that can be built upon. The input to the model is an ultrasound image and output is an image of the same dimensions with 0 and 1 pixel values indicating background and vessel lumen.

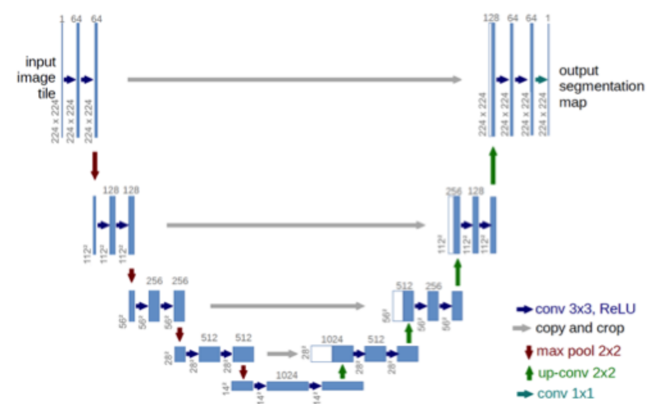


Fig. 1. U-Net architecture [11] that we use in our preliminary work. Shown here are dimensions of our images in each layer and the number of convolutional and transposed convolutions per layer.

Roughly speaking, in this model we first extract features with a series of convolutional kernels and then apply transpose convolutions to increase the dimensionality of the image up to the original. Thus we have an end-to-end network that is much simpler to train than otherwise patch-based approaches that have previously been used for segmentation.

4) Dice loss: The final output from the each of our models is a 2D predicted image of dimensions 224×224. We convert each pixel value into probabilities with softmax [13] and call the resulting image p. The target ground truth r is also of the same dimensions as p and contains a 1 if the pixel is within the vessel lumen and 0 otherwise. We then use the Dice loss to train our model. This is defined to be $1 - D$ where

$$D(p) = \frac{2\sum_i p_i r_i}{\sum_i p_i^2 + \sum_i r_i^2}$$

and p_i and r_i are the i^{th} pixel values of p and r respectively.

C. Implementation, accuracy, and validation

1) Implementation: We implemented our models using Pytorch [12] and ran them on NVIDIA Pascal P100 and NVIDIA Titan RTX GPUs. We trained our models with 20 epochs of stochastic gradient descent [14], a learning rate of 0.03, decay step of 15 (with $\gamma = .1$), and a batch size of 4. We did not perform any normalization on the input images.

2) Post processing: We applied a simple post processing procedure to reduce potential false positives. In the final predicted segmentation we remove all disconnected components except for the largest one that is meant to be the vessel lumen. We found that this improved accuracy by a moderate margin.

3) Measure of accuracy: Dice coefficient: The Dice coefficient is typically used to measure the accuracy of predicted segmentations in medical images [15]. We convert the output image of our network into a binary mask by setting each pixel value to 1 if its softmax output is at least 0.5 and 0 otherwise. Thus we use 0.5 as the probability threshold that a pixel value is part of the vessel lumen or outside it.

Starting with the human binary mask as ground truth, each predicted pixel is determined to be either a true positive (TP, also one in true mask), false positive (FP, predicted as one but zero in the true mask), or false negative (FN, predicted as zero but one in the true mask). The Dice coefficient is formally defined as

$$DICE = \frac{2TP}{2TP + FP + FN} \tag{1}$$

4) 10-fold cross-validation: We performed 10-fold cross-validation experiments on our data. We randomly split our dataset into ten equal parts and selected one part for validation while the remaining nine parts were used to train the model. We then rotated the validation part across the other nine parts giving us a total of 10 pairs of training validation splits. We trained the model on each split and reported the average validation and training accuracy below.

III. Results

We first perform a 10-fold cross-validation on the entire set of images. In Table I below we see that we achieve a high training and validation accuracy of 95.1% and 94.3% respectively. The small difference between our training and validation accuracies indicates our model is not overfitting and instead is generalizing well.

When we train and test on internal (ICA), external (ECA), and the common (CCA) carotid artery ultrasounds alone we see varying degrees of accuracy. Both ICA and ECA images achieve similar and lower train and validation accuracies than CCA which alone has a 96.6% accuracy (Table I).

TABLE I
AVERAGE ACCURACY OF TRAINING AND VALIDATION SPLITS IN OUR 10-FOLD EXPERIMENT.

	Training	Validation
All	95.1%	94.3%
ICA	94.9%	92.7%
ECA	96.2%	91.9%
CCA	97.9%	96.6%

In Figure 2 we see that adding more samples increases both the training and validation accuracy of our model. This is overall encouraging, however the increase in accuracy is by small margins and is plateauing at 95% as we add more samples.

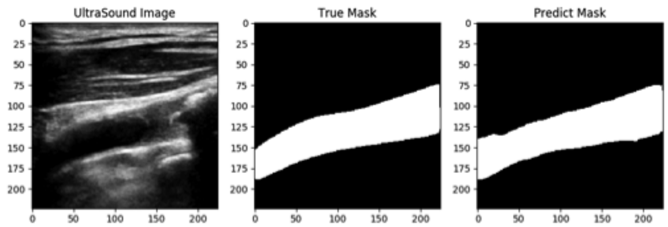
Fig. 2. Training and validation accuracy of our training and validation folds as a function of sample size.

In Figure 3 we see examples of some images with their true and predicted segmentations (also known as masks). Both (a) and (b) show examples with significant plaque and shadowing that could obfuscate the untrained eye but our model gives a highly correct segmentation. In (c) and (d) we have examples of bifurcated and gray shaded vessels that that are also correctly segmented by our model.

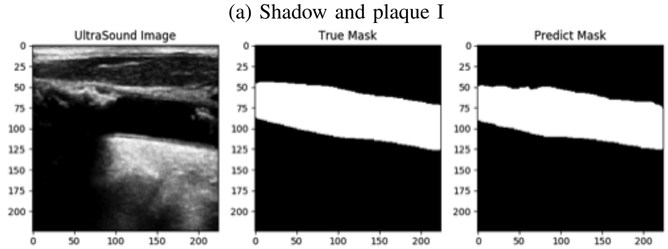
IV. Discussion

Medical imaging has become an essential component in modern medicine. It aids in diagnosis, tracks progression of disease and can be utilized to screen individuals for cancer and for the prevention of stroke. Numerous studies are looking at using deep learning methods to increase accuracy of diagnosis and aid in the interpretation of these studies [16]. As of yet there are few studies that look at utilizing deep learning for vessel identification and evaluation with ultrasound images specifically in the carotid artery system.

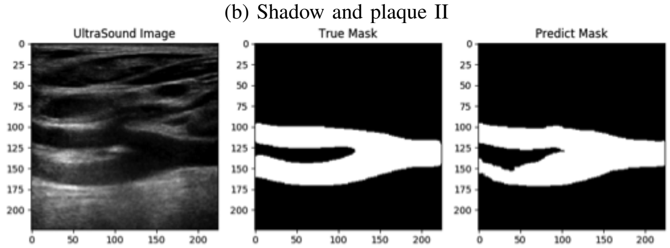
Ultrasound images provide a distinct challenge that is different than other medical imaging modalities. Computed



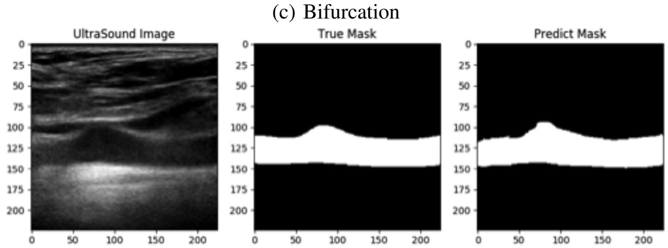
(a) Shadow and plaque I



(b) Shadow and plaque II



(c) Bifurcation

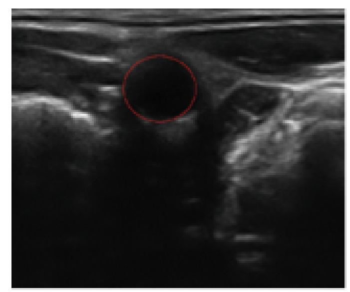


(d) Gray shading

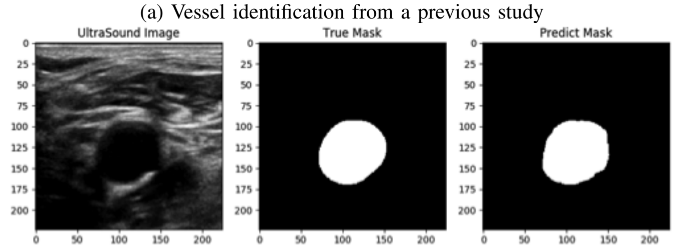
Fig. 3. Non-trivial examples of vessels in ultrasound images

tomography and magnetic resonance imaging (MRI) have set protocols that control formatting and orientation. For example, MRI images are typically aligned to a standard reference brain template such as the Montreal Neurological Institute reference space [17] that makes it easier to compare different MRI images.

A convolutional neural network was previously proposed for vessel detection in ultrasounds of femoral regions and also applied to carotid artery ultrasounds [18]. There are several key differences between our study and this previous one. In the previous study authors evaluate their method on transverse images of the common carotid artery. Specifically, they identify the center of the vessel and outline the vessel with an ellipse that approximates the vessel. To do this they are using a simplified version of the AlexNet [8] convolutional neural network. They reduced it to two convolutional layers, one normalization, two max pooling, and three fully connected. Their modified network outputs the center and two radii of the ellipse enclosing the circular vessel. In contrast, the U-Net that we use outputs a full segmentation of the vessel that can segment both transverse and longitudinal images of the vessel (Figure 4). Their study also only evaluates the common carotid artery, whereas our model can also be used to evaluate the internal and external carotid arteries, which is important because assessment of the carotid bifurcation and internal carotid artery has the most clinical relevance to stroke prevention. The previous study purely helps to identify that a vessel is present, it provides little additional input to aid in the interpretation of the ultrasound.



(a) Vessel identification from a previous study



(b) Our model performs circular segmentation that includes longitudinal images

Fig. 4. Comparison of vessel identification from previous work to vessel segmentation in our work.

V. Conclusion

The work that we present above is entirely novel in scope. It is the first step in attempting to create and implement a neural network that can independently and accurately identify and segment the lumen of the carotid artery in a vascular ultrasound. Further studies will be required to advance this model so that it can handle segmentation of the vessel wall, atherosclerotic plaque and evaluate direction of flow and flow velocity within the lumen before it can provide clinically relevant interpretations. But it has the potential to be the first step in creating a complete end-to-end solution for the evaluation of vascular ultrasound images.

References

[1] Dariush Mozaffarian, Emelia J Benjamin, Alan S Go, Donna K Arnett, Michael J Blaha, Mary Cushman, Sandeep R Das, Sarah De Ferranti, Jean Pierre Després, Heather J Fullerton, et al. Heart disease and stroke statistics-2016 update a report from the american heart association. *Circulation*, 133(4):e38–e48, 2016.

[2] James H Stein, Claudia E Korcarz, R Todd Hurst, Eva Lonn, Christopher B Kendall, Emile R Mohler, Samer S Najjar, Christopher M Rembold, and Wendy S Post. Use of carotid ultrasound to identify subclinical vascular disease and evaluate cardiovascular disease risk: a consensus statement from the american society of echocardiography carotid intima-media thickness task force endorsed by the society for vascular medicine. *Journal of the American Society of Echocardiography*, 21(2):93–111, 2008.

[3] Meiyan Xie, Yunzhu Li, Yunzhe Xue, Randy Shafritz, Saum A Rahimi, Justin W Ady, and Usman W Roshan. Vessel lumen segmentation in internal carotid artery ultrasounds with deep convolutional neural networks. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 2393–2398. IEEE, 2019.

[4] Philip Tamimi-Sarnikowski, Andreas Brink-Kjær, Ramin Moshavegh, and Jørgen Arendt Jensen. Automatic segmentation of vessels in in-vivo ultrasound scans. In *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 10137, page 101371P. International Society for Optics and Photonics, 2017.

[5] Giles Tetteh, Velizar Efremov, Nils D Forkert, Matthias Schneider, Jan Kirschke, Bruno Weber, Claus Zimmer, Marie Piraud, and Bjoern H Menze. Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes. *arXiv preprint arXiv:1803.09340*, 2018.

[6] Ran Zhou, Aaron Fenster, Yujiao Xia, J David Spence, and Mingyue Ding. Deep learning-based carotid media-adventitia and lumen-intima boundary segmentation from three-dimensional ultrasound images. *Medical physics*, 46(7):3180–3193, 2019.

[7] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[9] Jose Bernal, Kaisar Kushibar, Daniel S Asfaw, Sergi Valverde, Arnau Oliver, Robert Martí, and Xavier Lladó. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artificial intelligence in medicine*, 2018.

[10] Dan Ciresan, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in neural information processing systems*, pages 2843–2851, 2012.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[12] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. In *NIPS-W*, 2017.

[13] Ethem Alpaydin. *Machine Learning*. MIT Press, 2004.

[14] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Springer, 2010.

[15] Alex P Zijdenbos, Benoit M Dawant, Richard A Margolin, and Andrew C Palmer. Morphometric analysis of white matter lesions in mr images: method and validation. *IEEE transactions on medical imaging*, 13(4):716–724, 1994.

[16] Alexander Selvikvåg Lundervold and Arvid Lundervold. An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, 29(2):102–127, 2019.

[17] Vladimir Fonov, Alan C Evans, Kelly Botteron, C Robert Almli, Robert C McKinstry, D Louis Collins, Brain Development Cooperative Group, et al. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*, 54(1):313–327, 2011.

[18] Erik Smistad and Lasse Løvstakken. Vessel detection in ultrasound images using deep convolutional neural networks. In *Deep Learning and Data Labeling for Medical Applications*, pages 30–38. Springer, 2016.