Segmentation of apical long axis, four- and two-chamber views using deep neural networks

Erik Smistad
Centre for Innovative Ultrasound Solutions
Norwegian University of Science and Technology
SINTEF Medical Technology
Trondheim, Norway
erik.smistad@ntnu.no

Ivar Mjåland Salte
Sørlandet Hospital
Inst. of Clinical Medicine
University of Oslo
Kristiansand, Norway
ivar.mjaland.salte@sshf.no

Andreas Østvik
Centre for Innovative Ultrasound Solutions
Norwegian University of Science and Technology
SINTEF Medical Technology
Trondheim, Norway
andreas.ostvik@ntnu.no

Sarah Leclerc
CREATIS
Institut national des sciences appliquées (INSA)
Lyon, France
sarah.leclerc@insa-lyon.fr

Olivier Bernard
CREATIS
Institut national des sciences appliquées (INSA)
Lyon, France
olivier.bernard@insa-lyon.fr

Lasse Lovstakken
Centre for Innovative Ultrasound Solutions
Norwegian University of Science and Technology
Trondheim, Norway
lasse.lovstakken@ntnu.no

***Abstract*—It has been shown that deep neural networks can accuratly segment the left ventricle (LV), myocardium and left atrium in apical two and four chamber (A2C and A4C) views. While segmentation of apical long-axis (ALAX) views is quite similar to A2C and A4C, there is one major difference; the left ventricular outflow tract (LVOT) which restricts the myocardium. The objectives of this work were to accurately segment ALAX views, investigate if transfer learning from A2C/A4C improves accuracy, and study how a single network can learn to segment all three views.**

The CAMUS dataset of 500 patients together with an additional dataset of 106 patients with ALAX views were used for training and testing using 10-fold cross-validation. The results showed that by training from scratch the neural network was able to segment ALAX views, but with a lower accuracy to that of A2C/A4C views. Transfer learning only slightly improved mycoardium accuracy (0.77 to 0.78), but was statistically significant (p-value 0.001). Multi-view segmentation with the baseline network showed a reduction in accuracy, resulting in 38 cases of incorrect segmentations in terms of LVOT. The proposed network reduced the number of incorrect segmentations to 8, and achieved the best overall accuracy in terms of dice score where the improvement in myocardium segmentation accuracy (0.776 to 0.786) was statistically significant (p-value 0.005).

***Index Terms*—deep learning, echocardiography, segmentation, apical long axis, transfer learning**

I. Introduction

Automatic segmentation of the left ventricle (LV) in apical two and four chamber (A2C and A4C) ultrasound views has been studied for several decades. In recent years, many have demonstrated that deep neural networks can do the same [1]–[3], and even in real-time [4]. The CAMUS study in 2019 [5] demonstrated the superiority of deep neural networks over the more traditional approaches. To our knowledge, there are no studies so far considering the segmentation of apical long axis (ALAX) views using deep neural networks.

This research was funded by the Research Council of Norway under project 237887.

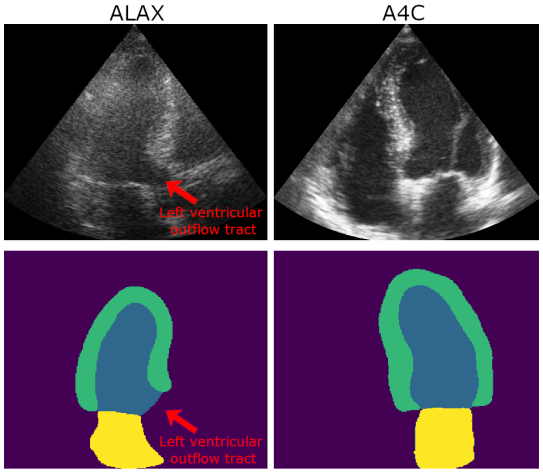

Fig. 1. An example of left ventricular segmentation in apical long axis and apical four chamber ultrasound views. While not present in apical two and four chamber views, the left ventricular outflow tract restricts the myocardium (green) in apical long axis views.

Segmentation of ALAX views is quite similar to A2C/A4C views, except for one major difference: the left ventricular outflow tract (LVOT) in ALAX views, which restricts the myocardium as shown in Fig. 1 The ALAX view is important in order to cover all parts of the LV wall from the apical insonification angle. Together, recordings from the A2C, A4C and ALAX views provide possibilities for quantitative analysis of all myocardial segments. Thus, these three views are typically used for strain measurements such as global longitudinal strain

(GLS) and regional strain [6]. Automatic segmentation of the ALAX view may reduce the high inter-observer variability and time spent on these measurements by removing manual user initialization.

The main objective of this work was to accurately segment ALAX views, and secondly, study how a single network can learn to accurately segment all three apical views.

II. Methods

In this article, four experiments were conducted to find the best approach to segment the LV and myocardium from all three apical views: 1) training an encoder-decoder network from scratch to segment ALAX, 2) using transfer learning from the CAMUS dataset 3) training a network from scratch to segment all three views and 4) using a novel multi-view segmentation network which can segment all three views. In the next sections, the network and training setup for each of these experiments are described.

A. Baseline segmentation network

In the first experiment, the U-net 1 network used to segment A2C and A4C views in the CAMUS study [5] was trained from scratch to segment ALAX views. This network, now referred to as the *baseline network*, is a fully convolutional encoder-decoder network with max pooling in the encoder stage, and 2×2 repeat upsampling in the decoder stage. The network has about 2 million parameters.

B. Transfer learning from A2C/A4C segmentation

The image and the segmentation of A2C/A4C and ALAX views are quite similar, except for the LVOT in ALAX views. Thus, it's reasonable to assume that most of the features used for segmenting A2C/A4C are useful for segmenting ALAX views as well. In this second experiment, the U-net 1 network from the CAMUS study, already trained on A2C/A4C views was used for transfer learning. This pretrained A2C/A4C network was fine-tuned on the ALAX dataset by training with a reduced learning rate of 0.0005 compared to the learning rate 0.001 used to pretrain the network.

C. Multi-view segmentation using baseline network

In this experiment, the baseline segmentation network was trained to segment all three views. Since there were 500 patients for A2C and A4C views, and only 106 patients for ALAX views, the ALAX views were repeated ten times in each epoch to deal with this data balancing issue.

D. Multi-view segmentation using proposed network

When performing segmentation, it is usually known which view is being processed beforehand. The current view can either be specified manually by the operator or automatically using deep neural networks [7]. This inspired the creation of a novel network architecture which has two inputs: 1) the ultrasound image, and 2) a binary scalar indicating whether the image is a A2C/A4C (0) or a ALAX (1) view.

Since encoder-decoder segmentation networks usually only have convolutional layers, and no dense layers, the binary scalar has to be combined with 2D feature maps. To this end, the binary scalar was duplicated to create a 2D image of the same size as the ultrasound image. This binary image is passed on to a layer of N 1×1 convolutions, resulting in a feature map of N channels. This feature map is then multiplied with features from the segmentation network in both the encoder and decoder as shown in Fig. 2. The idea is that the neural network can learn to disable features which are only needed for A2C/A4C or ALAX by using the binary scalar. This is repeated for every level in both the encoder and decoder stage of the U-net structure as displayed in Fig. 2. The proposed network has only 1,920 more parameters than the baseline network, which comes from the added 1×1 convolutions.

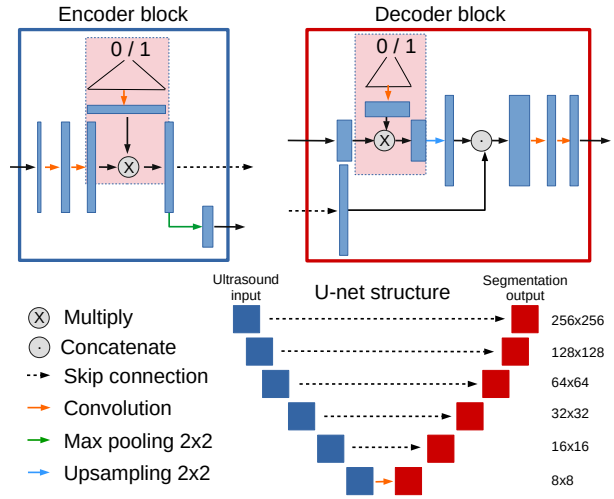

Fig. 2. Proposed neural network for multi-view cardiac ultrasound segmentation. The network has two inputs: the ultrasound image and a binary scalar indicating whether the input is A2C/A4C or ALAX image. In each encoder and decoder block of the U-net structure, the binary scalar is repeated to create a 2D image. N 1×1 convolutions are then applied to this binary 2D image. The resulting tensor is multiplied with the tensors in the U-net before down- and upsampling. The pink areas highlights the parts that were added to the baseline network.

E. Datasets

The publicly available CAMUS dataset of 500 patients was used in this study as a source of annotated A4C and A2C views. For the ALAX views, a new dataset was created consisting of images from 106 patients. These images were segmented using the same protocol as used for making the CAMUS dataset: For each patient, two frames corresponding to the end-diastolic (ED) and end-systolic (ES) time points were selected for segmentation. In these frames, the LV, the myocardium and the left atrium were delineated using a cardinal spline contour defined from a set of points manually selected from the image. The only difference was, that when annotating the ALAX views, the myocardium segmentation was ended at the beginning of the LVOT as shown in Fig. 1.

III. Results

For each experiment, the dice scores of the LV and the myocardium were measured and collected in Table I. The

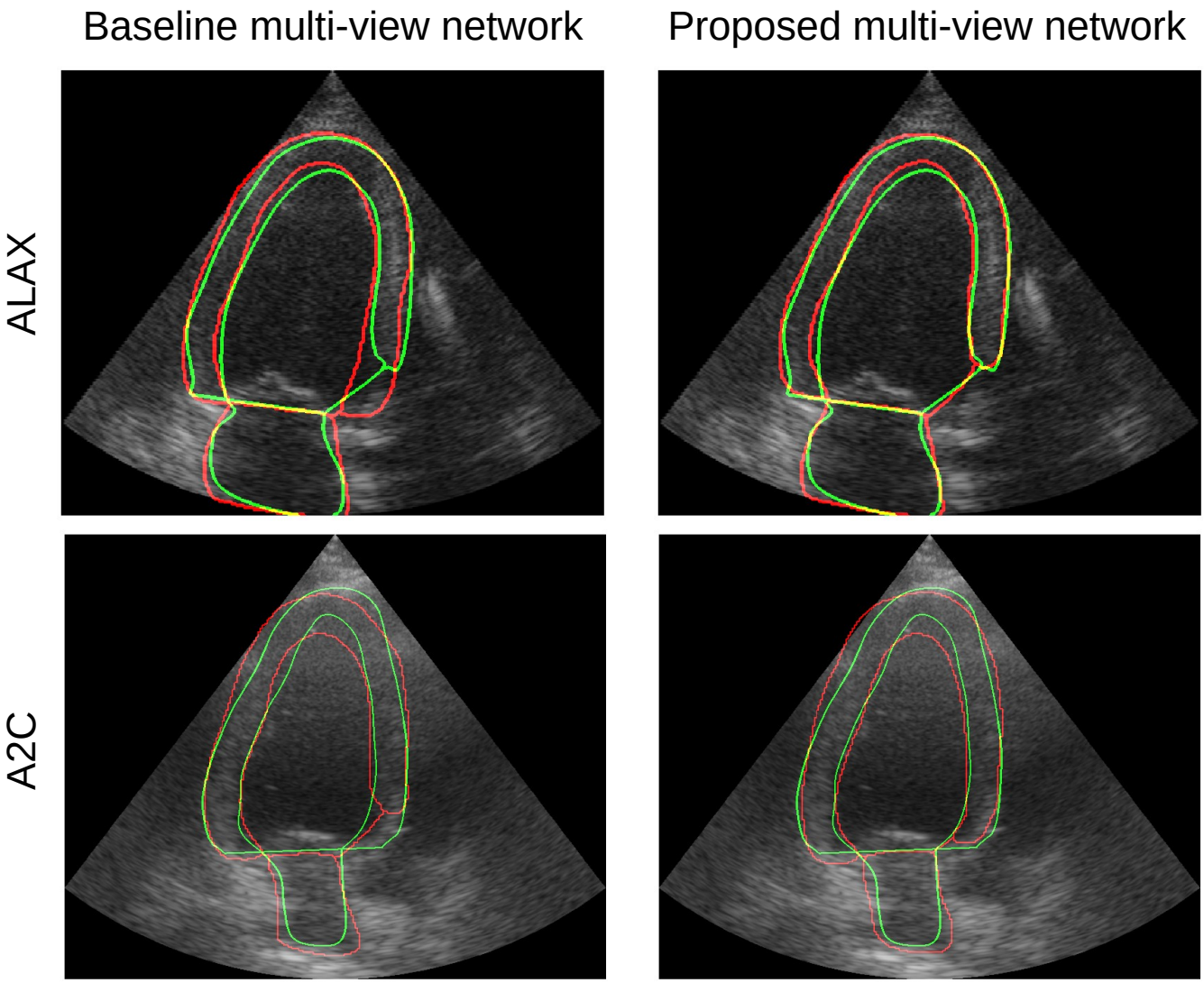

Fig. 3. Two examples that were incorrectly segmented in terms of the LVOT using the baseline multi-view segmentation network, while segmented correctly with the proposed network. The green lines are the expert's delineation, the red lines correspond to edges of the neural network's segmentation output, and yellow means overlap between the two. The top row is a true ALAX image, which is interpreted and segmented as A2C by the baseline network and the LVOT segmented as myocardium. The bottom row is an A2C image, which is interpreted as a ALAX image by the baseline network and therefore incorrect segmented as having an LVOT. The total number of incorrect segmentation's in terms of LVOT were reduced from 38 to 8 using the proposed network.

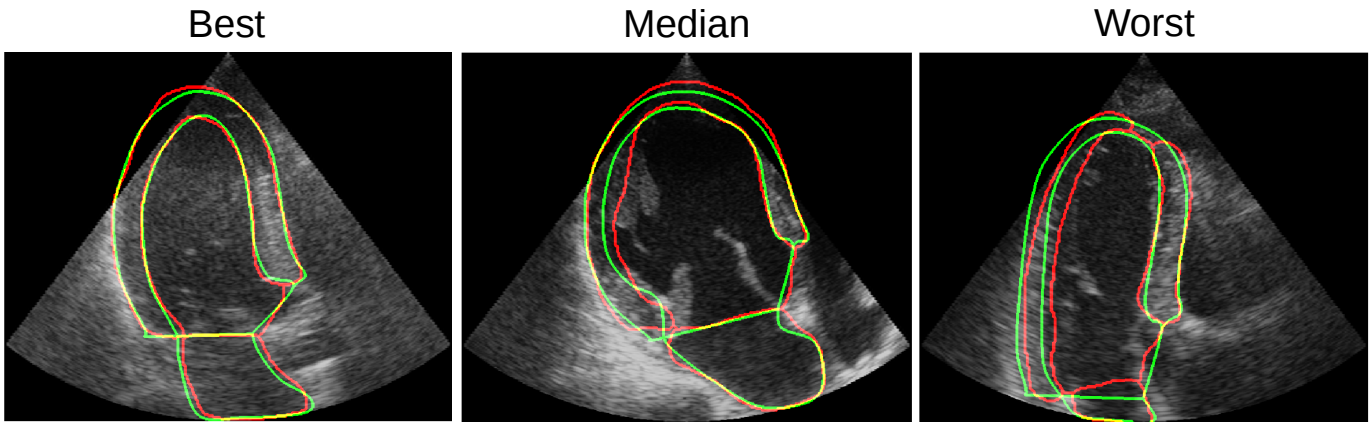

Fig. 4. The best, median and worst case ALAX segmentations in terms of LV and myocardium dice scores using the proposed multi-view segmentation network on the ALAX dataset. The green lines are the expert's delineation, the red lines correspond to edges of the neural network segmentation, and yellow means overlap between the two.

TABLE I
CROSS VALIDATION RESULTS WITH MEAN DICE SCORE AND STANDARD DEVIATION OF ALL 10 FOLDS.

Experiment	Dice Left Ventricle		Dice Myocardium		Dice Left Atrium	
	ALAX	A4C/A2C	ALAX	A4C/A2C	ALAX	A4C/A2C
Baseline network trained with A2/4C only		$\mathbf{0.926} \pm 0.05$		0.861 ± 0.06		$\mathbf{0.894} \pm 0.09$
Baseline network trained with ALAX only	0.917 ± 0.03		0.770 ± 0.07		0.881 ± 0.09	
Transfer learning from CAMUS dataset	0.918 ± 0.03		0.780 ± 0.07		0.890 ± 0.06	
Multi-view segmentation with baseline network	0.920 ± 0.03	0.922 ± 0.05	0.776 ± 0.08	0.856 ± 0.06	0.886 ± 0.08	0.892 ± 0.09
Multi-view segmentation with proposed network	$\mathbf{0.921} \pm 0.03$	0.924 ± 0.04	$\mathbf{0.786} \pm 0.08$	$\mathbf{0.862} \pm 0.05$	$\mathbf{0.892} \pm 0.08$	0.893 ± 0.09

scores were measured on the entire dataset using 10 fold cross-validation. As done in the CAMUS study [5], all neural network segmentation outputs were post-processed by filling any holes in the segmentation and removing any small regions. For comparison, dice scores obtained with the baseline network trained and tested on the A4C/A2C dataset, are also included in the table. For the multi-view segmentation experiments, the number of incorrect segmentations in terms of the LVOT was established as the number of ALAX images segmented *without* an LVOT, and the number of A2C/A4C images segmented *with* an LVOT. With the baseline network, 9 ALAX and 29 A2C/A4C images were segmented incorrectly. With the proposed network, the number of incorrect segmentations decreased to 0 and 8 for ALAX and A2C/A4C respectively as shown in Fig. 3. Fig. 4 contains ALAX examples of the best, median and worst segmentation in terms of LV and myocardium dice score using the proposed multi-view segmentation network.

IV. DISCUSSION

The results in Table I show only a small improvement in using transfer learning from A4C/A2C in terms of segmentation accuracy for ALAX views. A Wilcoxon signed-rank test showed that only the improvement in myocardium segmentation accuracy was statistically significant with a p-value of 0.001. Multi-view segmentation reduces the accuracy of both ALAX and A4C/A2C views using the baseline network, mainly because it is not able to correctly identify the view as either ALAX or A2C/A4C in a total of 38 cases. However, by using the proposed network with an additional input specifying whether the input is an ALAX or A4C/A2C view, this issue is significantly reduced to only 8 cases, thereby achieving the overall highest segmentation accuracy of all experiments. Still, the improvements in the dice scores are very small. Therefore a Wilcoxon signed-rank test was performed, revealing that only the improvement in myocardium segmentation accuracy was significant with a p-value of 0.005. This indicates that it is advantageous for the segmentation network to know which view is being segmented in terms of myocardium segmentation accuracy. This might be because the myocardium is segmented differently in these views, but it might also be due to inter-observer variability since the two datasets with the different views were annotated by two different experts. Another approach to having a network with an additional input value, could be to extend the neural network and have it perform multi-task learning, and learn to perform both view classification and segmentation at the same time. However, this might result in reduced runtime.

V. CONCLUSION

In this work, the left ventricular segmentation of apical long axis views from cardiac ultrasound was investigated. Transfer learning from a neural network trained on apical two and four chamber views improved accuracy slightly, but the best myocardium segmentation accuracy was achieved by a novel multi-view segmentation network which was able to successfully segment all three apical views. The multi-view network achieved this by training on all types of views and integrating a second input in the form of a binary scalar indicating whether the input image is an apical long-axis view.

REFERENCES

[1] E. Smistad, A. Ostvik, B. Haugen, and L. Lovstakken, "2D left ventricle segmentation using deep learning," in *IEEE International Ultrasonics Symposium, IUS*, 2017.

[2] J. Zhang, S. Gajjala, P. Agrawal, G. H. Tison, L. A. Hallock, L. Beussink-Nelson, M. H. Lassen, E. Fan, M. A. Aras, C. Jordan, K. E. Fleischmann, M. Melisko, A. Qasim, S. J. Shah, R. Bajcsy, and R. C. Deo, "Fully Automated Echocardiogram Interpretation in Clinical Practice," *Circulation*, vol. 138, no. 16, pp. 1623–1635, oct 2018.

[3] M. H. Jafari, H. Girgis, Z. Liao, D. Behnami, A. Abdi, H. Vaseli, C. Luong, R. Rohling, K. Gin, T. Tsang, and P. Abolmaesumi, "A Unified Framework Integrating Recurrent Fully-Convolutional Networks and Optical Flow for Segmentation of the Left Ventricle in Echocardiography Data," in *DLMIA*, vol. 10553. Springer International Publishing, 2018, pp. 29–37.

[4] E. Smistad, A. Ostvik, I. Mjal Salte, S. Leclerc, O. Bernard, and L. Lovstakken, "Fully Automatic Real-Time Ejection Fraction and MAPSE Measurements in 2D Echocardiography Using Deep Neural Networks," in *2018 IEEE International Ultrasonics Symposium (IUS)*, vol. 2018-Octob. IEEE, oct 2018, pp. 1–4.

[5] S. Leclerc, E. Smistad, J. Pedrosa, A. Ostvik, F. Cervenansky, F. Espinosa, T. Espeland, E. A. R. Berg, P.-M. Jodoin, T. Grenier, C. Lartizien, J. Drhooge, L. Lovstakken, and O. Bernard, "Deep Learning for Segmentation using an Open Large-Scale Dataset in 2D Echocardiography," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2019.

[6] R. M. Lang, L. P. Badano, V. Mor-Avi, J. Afilalo, A. Armstrong, L. Ernande, F. A. Flachskampf, E. Foster, S. A. Goldstein, T. Kuznetsova, P. Lancellotti, D. Muraru, M. H. Picard, E. R. Rietzschel, L. Rudski, K. T. Spencer, W. Tsang, and J.-U. Voigt, "Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *Journal of the American Society of Echocardiography*, vol. 28, no. 1, pp. 1–39.e14, jan 2015.

[7] A. Østvik, E. Smistad, S. A. Aase, B. O. Haugen, and L. Lovstakken, "Real-Time Standard View Classification in Transthoracic Echocardiography Using Convolutional Neural Networks," *Ultrasound in Medicine & Biology*, vol. 45, no. 2, pp. 374–384, feb 2019.