Fully automatic real-time ejection fraction and MAPSE measurements in 2D echocardiography 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

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

Ivar Mjåland Salte
Sørlandet Hospital
Kristiansand, Norway
ivar.mjaland.salte@sshf.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—Cardiac ultrasound measurements such as left ventricular volume, ejection fraction (EF) and mitral annular plane systolic excursion (MAPSE) are time consuming and highly observer dependent. In this work, we investigate if deep neural networks can be used to fully automate cardiac ultrasound measurements in real-time while scanning. One neural network was used for identifying and separate the cardiac views while a second neural network performed segmentation of the left ventricle. By using TensorFlow, FAST and the highly optimized cuDNN backend real-time runtime of the entire pipeline was achieved with an average frames per second of 43, thus enabling these measurements to be performed while an operator is scanning. The measurement accuracy was evaluated using a Bland-Altmann analysis on a dataset of 75 patients resulting in $(-13.7 \pm 8.6)\%$ for EF and (-0.9 ± 4.6) mm for MAPSE. It is concluded that deep learning can be used to fully automate these measurements, however more work remains to improve the accuracy.

I. Introduction

Automated measurements in echocardiography has the potential of improving workflow by reducing time spent on manual annotation and contouring. It can also reduce interobserver variability known to be significant in echocardiography. In order to fully automate measurements such as volume, ejection fraction (EF) and mitral annular plane systolic excursion (MAPSE) several automated components are needed.

In order to estimate the left ventricle (LV) volume in 2D echocardiography, a typical convention is to acquire two standard views of the heart. The most common views for this task are the apical four-chamber (A4C) and two-chamber (A2C) views. A fully automated system needs to recognize these views, which has shown to be possible with deep neural networks [1], [2]. Another essential step for automating cardiac measurements is image segmentation. This has been an active research area for several decades for both 2D and 3D ultrasound, and recently it has been shown that deep neural networks can also perform this efficiently and accurately [3–6]. In addition to view classification and segmentation, the estimation of end-diastole and end-systole, as well as extraction of apex and base landmarks is needed.

This paper presents methods for automating all of these necessary steps in real-time using deep convolutional neural networks. The accuracy and speed of the methods were measured, and a video posted online demonstrates the software while it is used to perform cardiac measurements real-time by streaming ultrasound images directly from a GE Vivid E95 scanner.

II. Methods

A. View classification

As mentioned, volume and MAPSE measurements are typically performed on apical four- and two-chamber views, thus a detection of views are needed for storing valid data in a buffer. In this work, the view classification network of Østvik et al. [1] was used. It uses inception blocks and a dense connectivity pattern, and was trained with images from about 200 patients, and validated on an independent dataset

of equal size giving an accuracy of approximately 98%. The network can recognize eight different cardiac views: 1) apical four-chamber (A4C), 2) apical two-chamber (A2C), 3) apical long axis, 4) parasternal short axis, 5) parasternal long axis, 6) subcostal four-chamber, 7) subcostal vena cava and 8) unknown view.

B. Segmentation

For every frame recognized as either A4C or A2C, segmentation was also performed. Segmentation was done using a U-net type architecture optimized for both accuracy and speed [3]. This network classifies each pixel into one of four classes: 1) background, 2) left ventricle, 3) myocardium and 4) left atrium. The network was trained with dice loss and the Adam optimizer using a dataset consisting of A4C and A2C images from both ED and ES of 500 patients [6]. Fig. 1 shows an example of segmentations performed by this method.

C. End-diastole and -systole detection

End-diastole (ED) and end-systole (ES) are the two time points in the cardiac cycle corresponding to mitral and aortic valve closure respectively. Both indicate the start of isovolumetric phases, and for volumetric measures, the time points can be approximated to LV volume extrema. The LV area in each view should be roughly proportional to the volume, thus it can be used to determine ED and ES. After segmentation, the LV area is calculated and stored for each frame. The ED and ES frames are then determined by the maximum and minimum LV area respectively.

D. Measurements

Contours of the endocardium were extracted using morphological erosion on the segmentation of the LV at ED and ES. These contours were used to extract three basal landmarks (left, right and mid) and one apex landmark. A contour point was determined to be at the base if any pixels immediately below was segmented as the left atrium. From these base contour points, the left, right and mid base landmarks were extracted. The apex landmark was the contour point furthest away from the base mid landmark.

Using these landmarks, the ED and ES volume was estimated using Simpson's biplane method of 20 discs, as is clinical practice in echocardiography [7]. In this method, the diameter d is calculated at 20 steps perpendicular to the midaxis of the LV. This is done in both the A4C and A2C views.

$$V = \frac{\max(L_{\text{A4C}}, L_{\text{A2C}})}{20} \sum_{i=1}^{20} \pi \frac{d_{\text{A4C}}(i) d_{\text{A2C}}(i)}{4} \qquad (1)$$

From the ED and ES volumes, the ejection fraction was calculated as $100 \cdot (V_{\text{ED}} - V_{\text{ES}})/V_{\text{ED}}$.

MAPSE was calculated in A4C view using the base left and right landmarks $\vec{P}_L, \vec{P}_R$ at ED and ES and the probe origin position $\vec{P}_O$. Septal and lateral MAPSE was calculated

as $|\vec{P}_{L,\text{ED}} - \vec{P}_O| - |\vec{P}_{L,\text{ES}} - \vec{P}_O|$ and $|\vec{P}_{R,\text{ED}} - \vec{P}_O| - |\vec{P}_{R,\text{ES}} - \vec{P}_O|$ respectively.

E. Real-time implementation

All methods described so far, including streaming of ultrasound images was implemented in C++ using the FAST framework [8] which features GPU optimized processing and visualization for real-time applications. The view classification and segmentation neural networks were exported as a TensorFlow graph and stored together with the learned weights in a single file. This file is loaded in FAST which uses the TensorFlow C++ API with the NVIDIA cuDNN backend to perform high performance inference of the neural networks. The data streaming, visualizations and computations happens in three different threads. This enables visualization of the ultrasound data and measurements, while the view classification, segmentation, area calculations, landmark extraction, and measurements are running. Fig. 1 is a screenshot of the software implementation showing the measurement summary for one patient. A video showing this implementation used in real-time while scanning is available online[1].

III. RESULTS

A separate dataset of 75 patients was used to evaluate the measurement accuracy. The dataset consists of images from the A4C and A2C views acquired with a GE Vivid E9 US scanner (GE Vingmed Ultrasound, Horten, Norway) with GE M5S phased-array transducer. The images were manually segmented by another expert than the segmentation training dataset, but using the same protocol. The recordings were streamed through the application as if it were coming from an ultrasound scanner and the measurements were stored for each patient. The segmentation failed in 3 of the 75 patients, and were thus excluded from the evaluation. This was due to movement of the probe during the recording, thereby moving the heart out of the image plane and creating incorrect segmentations. A Bland-Altman analysis of EF and MAPSE was performed on the remaining 72 patients and is depicted in Fig. 2 and Fig. 3. The analysis showed a mean difference of -13.7% and standard deviation of 8.6% for EF and -0.9 mm and 4.6 mm for MAPSE.

An analysis of the ED and ES volumes revealed that the volumes are being underestimated with a mean difference and standard deviation of (16 ± 22) mL and (26 ± 19) mL.

The average number of frames per second was measured to be 43 on an Alienware laptop with an Intel i7-6700 CPU and an NVIDIA Geforce GTX 980M GPU. View classification and segmentation were the most computational intensive tasks with respective average runtimes of 10 and 9 ms. The landmark extraction and volume calculations took about 4 ms on average.

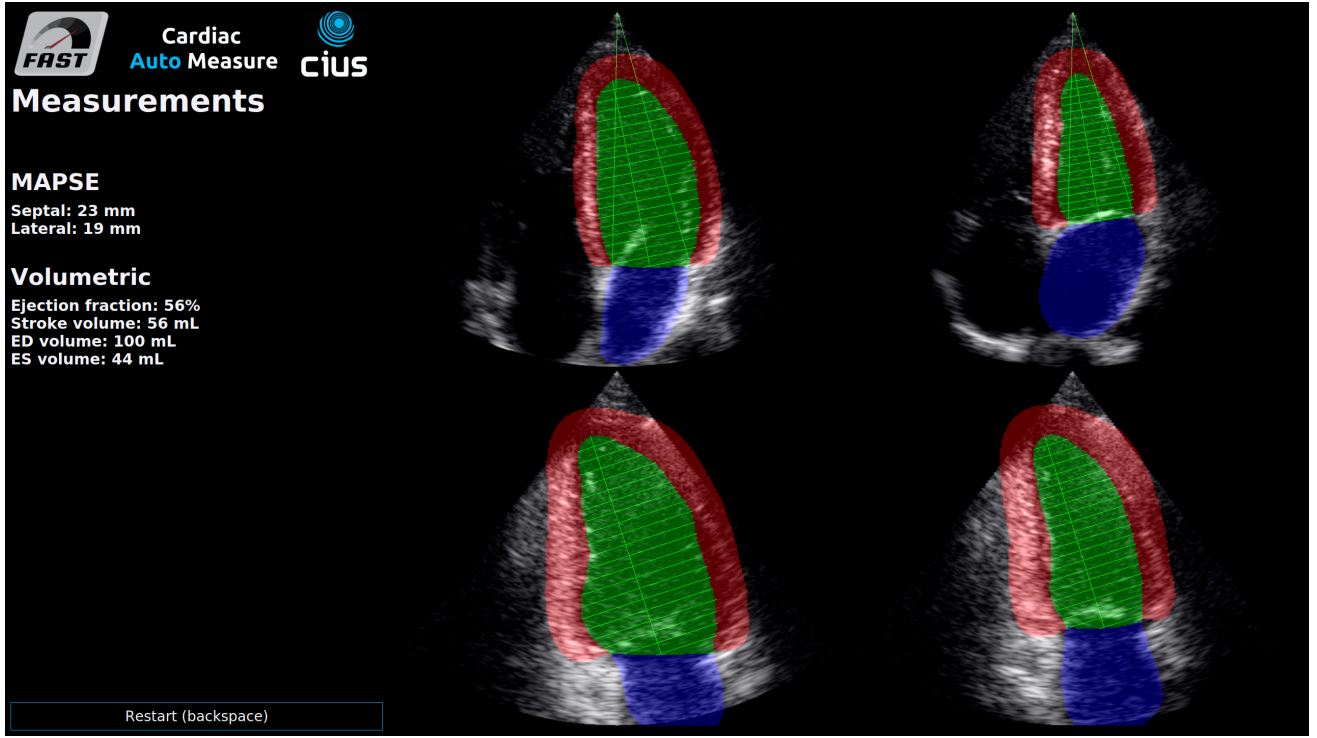[1]https://www.youtube.com/watch?v=EgoTX75KhxE

Fig. 1. Screenshot of the software which uses the proposed methods. The screenshot shows the measurements and ultrasound images with the segmentations on top of ED and ES of A4C (top row) and A2C (bottom row). The horizontal lines are used for Simpson's method of discs, and the vertical lines starting from the probe origin are the MAPSE measurements.
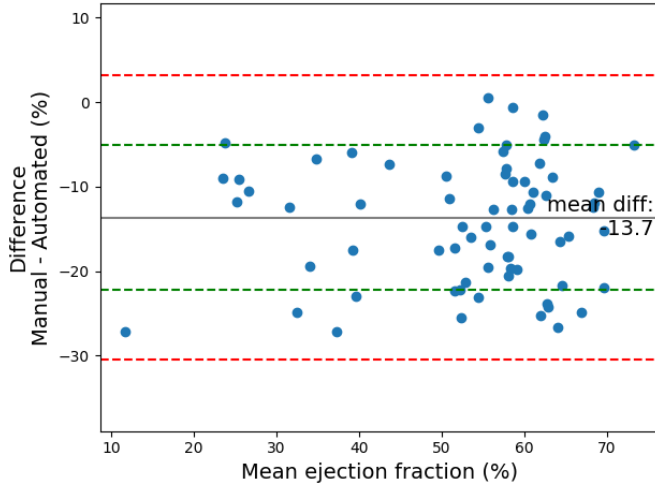

Fig. 2. Bland-Altmann plot of manual versus automated ejection fraction. The green lines represent the standard deviation, while the red lines are the 95% limits of agreement.
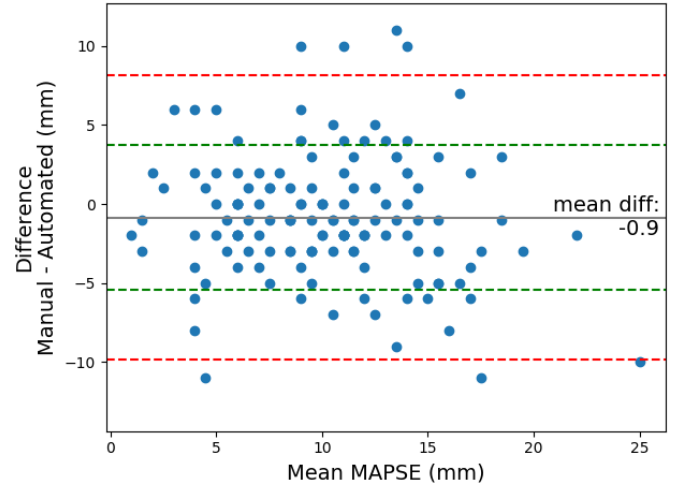

Fig. 3. MAPSE Bland-Altmann plot of manual versus automated MAPSE in millimeters. The green lines represent the standard deviation, while the red lines are the 95% limits of agreement.

IV. DISCUSSION

The results revealed a large bias in the measurements. One possible reason for this bias is that the segmentation sometimes struggles when the mitral valve opens. This can cause the segmentation of the LV to become smaller than it should in the systolic phase. The fact that the mean difference of the ES volume is larger (26 mL) than the ED volume (16 mL) support this. The results is an overestimation of the ejection fraction. Another reason is the interobserver variablity in the annotations, as two different experts were used for the segmentation training and test set.

Simultaneously with this work, Zhang et al. [2] published

a study on automated cardiac measurements such as volume and ejection fraction with deep neural networks using a large dataset. Their result shows a comparable standard deviation to this work on ejection fraction, but with a lower bias. Again, the higher bias in this work may be due to the use of an independent test dataset annotated by another person.

In clinical practice, MAPSE is measured using M-mode scanlines through the mitral annulus. In this work, the septal and lateral base points of the segmentation have been used to calculate MAPSE. These points are not necessarily the same as the annulus. Also, the detected base points may not be consistent over time, and ideally tracking should be used to improve this.

The measured runtime show that automatic real-time ejection fraction and MAPSE measurements are feasible by using deep convolutional neural networks and a modern GPU. This could have major implications on the clinical workflow and might enable novice users to do echocardiography, but quality assurance of the ultrasound images is still necessary to make this happen. Additional future work includes adding support for apical long axis views, which can also be used for volume estimation, improving the segmentation, and ED and ES estimation which currently is only based on the segmentation.

V. Conclusion

Methods for fully automating volume, ejection fraction and MAPSE measurements using deep learning were presented. It was demonstrated that these methods can be used in real-time in a complete software that streams images directly from an ultrasound scanner. The accuracy results are promising, but more work remains to validate the methods on a larger dataset and reduce the large bias in the measurements.

References

[1] 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 and Biology*, 2018.

[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, Deo, and R. C., "Fully Automated Echocardiogram Interpretation in Clinical Practice," *Circulation*, pp. 1–13, 2018.

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

[4] O. Oktay, E. Ferrante, K. Kamnitsas, M. Heinrich, W. Bai, J. Caballero, S. A. Cook, A. de Marvao, T. Dawes, D. P. O'Regan, B. Kainz, B. Glocker, and D. Rueckert, "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 384–395, feb 2018. [Online]. Available: http://arxiv.org/abs/1705.08302http://ieeexplore.ieee.org/document/8051114/

[5] 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. [Online]. Available: http://link.springer.com/10.1007/978-3-319-67558-9http://link.springer.com/10.1007/978-3-030-00889-5{_}4

[6] S. Leclerc, E. Smistad, T. Grenier, A. Østvik, F. Espinosa, P.-M. Jodoin, L. Lovstakken, and O. Bernard, "Deep learning applied to multi-structure segmentation in 2D echocardiography: a preliminary investigation of the required database size," in *IEEE International Ultrasonics Symposium, IUS*, 2018.

[7] 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. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0894731714007457

[8] E. Smistad, M. Bozorgi, and F. Lindseth, "FAST: framework for heterogeneous medical image computing and visualization," *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, no. 11, pp. 1811–1822, 2015. [Online]. Available: http://link.springer.com/10.1007/s11548-015-1158-5