Tracking-based mitral annular plane systolic excursion (MAPSE) measurement using deep learning in B-mode ultrasound

Erik Smistad*†‡, Andreas Østvik*†‡, Jahn Frederik Grue*†, Håvard Dalen*†§, Lasse Løvstakken*†

*Centre for Innovative Ultrasound Solutions (CIUS)
Trondheim, Norway

†Norwegian University of Science and Technology (NTNU)
Dept. of Circulation and Medical Imaging
Trondheim, Norway

‡SINTEF Medical Technology
Trondheim, Norway

§St. Olavs Hospital
Trondheim, Norway

***Abstract*—Mitral annular plane systolic excursion (MAPSE) is an important measure of left ventricular function. Current clinical practice is to measure it manually using M-mode ultrasound imaging which has several disadvantages such as "out-of-line" motion and M-mode angle and operator dependency. In this work, we propose a fully automatic method for measuring MAPSE in B-mode ultrasound using deep learning. The method involves multiple neural networks to detect end-diastolic and end-systolic frames, perform annulus landmark detection, and frame-by-frame tracking. It is also demonstrated how this B-mode based MAPSE can be used to remove radial motion of the annulus from the MAPSE measurement, thereby only measuring longitudinal motion of the annular plane. The landmark detection accuracy in end-diastole was measured to be 3.0 ± 2.5 mm, while the full pipeline gave a MAPSE accuracy of -1.5 ± 2.1 mm on a 72 subject dataset.**

I. Introduction

Mitral annular plane systolic excursion (MAPSE) is a measure of left ventricular (LV) longitudinal function. Current clinical practice is to use M-mode imaging to measure MAPSE along a line from the transducer and through the annulus on both sides of the mitral valve as shown in Fig. 1. This 1-dimensional way of measuring MAPSE has the disadvantages of M-mode angle dependency, "out-of-line" motion of the mitral annulus, operator dependent selection of M-mode direction and manual M-mode image analysis. In this work, the goal was to tackle these disadvantages and fully automate the measurement using B-mode ultrasound imaging and deep learning.

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

The HUNT study (Nord-Trøndelag Health Study) is a collaboration between the HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology), Nord-Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. We thank the Nord-Trøndelag Hospital Trust and for support and for contributing to data collection in this research project.

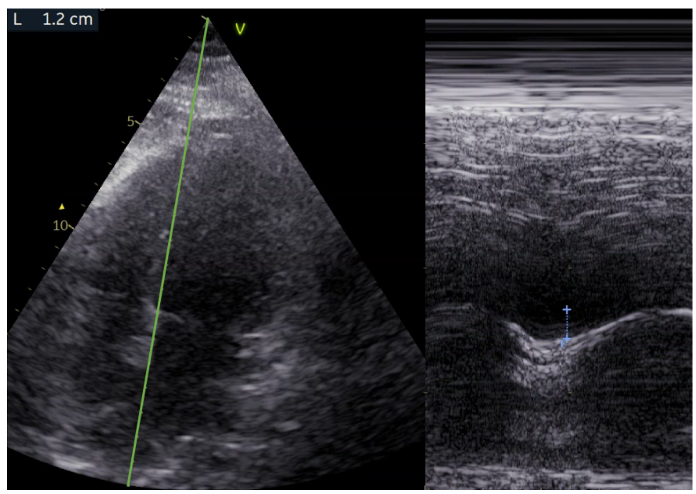

Fig. 1. The conventional way of measuring MAPSE using ultrasound M-mode imaging. The B-mode image to the left shows where the M-mode line (green) is set, and the right image is the corresponding M-mode image used for measuring MAPSE (blue caliper).

Automation of MAPSE measurements in 2-dimensional imaging has been studied before. In 2016, Storve et al. [1] presented a method for apical four chamber views which used an automatic segmentation method to detect the annulus landmarks and then tissue Doppler, also called tissue velocity imaging, to track the landmarks during the heart cycle. The method was accurate with a standard deviation of 2.1 mm, however, it requires both B-mode data, ECG and tissue Doppler. In 2018, Smistad et al. [2] proposed to use an automatic deep learning method to estimate MAPSE directly from B-mode images by segmenting images in end-diastole (ED)

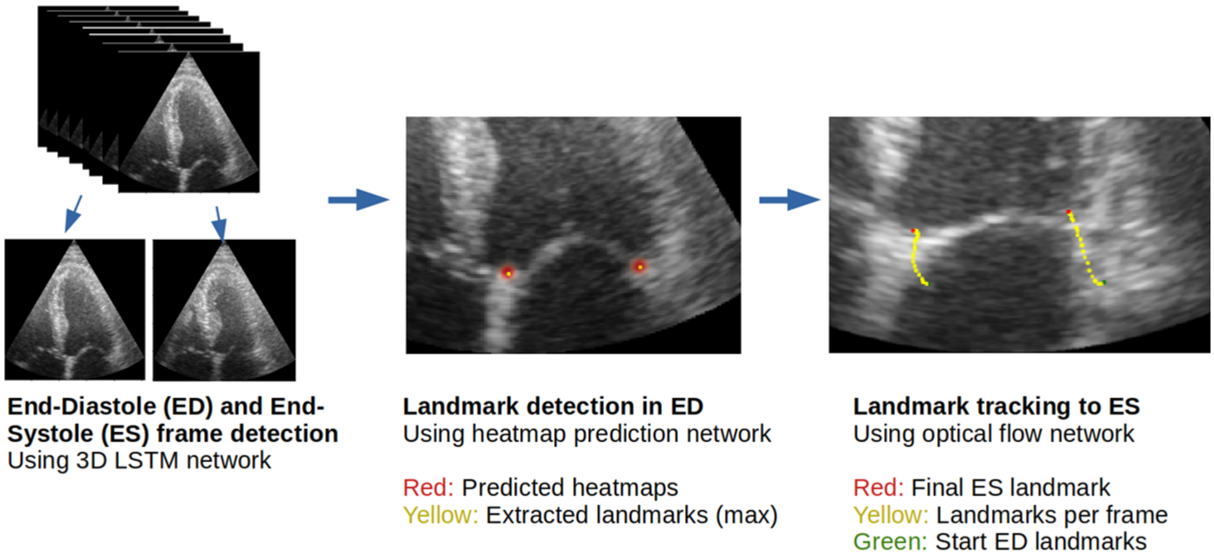

Fig. 2. Proposed pipeline for automatic MAPSE measurement directly from B-mode images using three different neural networks.

and end-systole (ES), and estimating the annulus points from the segmentation. However, this method was inaccurate with a standard deviation of over 4 mm. An automatic method for MAPSE requires an accurate way of identifying the annulus landmarks and an accurate method for tracking the landmarks from ED to ES. Recently, two publications have proposed a deep learning method for accurate motion estimation in B-mode ultrasound images [3], [4]. These methods use a convolutional neural network to estimate the motion in every pixel between two frames.

In this work, we combine the motion estimation network of [3] with an automatic annulus landmark detection method and an ED/ES estimation method to fully automate the MAPSE measurement as shown in Fig. 2.

II. Methods

A. Dataset and annotation

The mitral annulus was annotated by one expert on both sides throughout one whole cardiac cycle in 72 apical four- and two-chamber recordings from the HUNT population dataset. This was done using landmark annotation and manual speckle tracking in the open-source Annotation Web system [5]. Annotation Web enables clinicians to efficiently annotate large amounts of ultrasound recordings in a regular web browser without any manual installation or transfer of data. Since MAPSE calculated using B-mode is not directly comparable to M-mode measurements, this annotation served both as a measurement ground truth and as training data for annulus landmark detection.

B. End-diastolic and end-systolic frame detection

In order to fully automate the MAPSE measurement, we need to estimate which frames correspond to the ED and ES time points in the cardiac cycle. In this work, we used our previously published 3D convolutional LSTM network to estimate ED and ES [6]. This network processes a sequence of frames and outputs a value between 0 and 1 for each frame depending on whether the frame is from systole or diastole of the heart cycle. The time points of ED and ES can then be extracted from the cross-overs between 0-1, and 1-0.

C. Landmark detection

A fully-convolutional encoder-decoder neural network was trained to identify the annulus landmarks in a single 256×256 ultrasound image frame by predicting a 256×256 heatmap for each landmark. The network used was similar to the real-time segmentation network presented in [7] which has six levels and uses max pooling in the encoder and 2×2 repeat upsampling in the decoder. Two 3×3 convolution layers are used at each level, together with ReLU activation. The final layer uses softmax activation. The network has about 2 million parameters. Training heatmaps were generated by creating a Gaussian with standard deviation of 3 pixels at the annotated positions. During inference, landmark positions were extracted from the heatmaps by finding the pixel with maximum value. Intensity, shadow, rotation and JPEG augmentations were applied during training to reduce overfitting.

D. Landmark tracking

In order to calculate MAPSE, the annulus landmarks need to be estimated at both ED and ES. This can be done by using the landmark detection network on the ED and ES frames separately. However, this approach discards valuable temporal information between these frames, and the detected landmarks may not represent the same physical point, even though they are in the annulus region. For a correct MAPSE measure, it is essential to detect the same physical point in ED as in ES. This can be achieved using an accurate tracking method. Tracking in cardiac ultrasound images is usually performed using speckle tracking methods, and in this study we have used

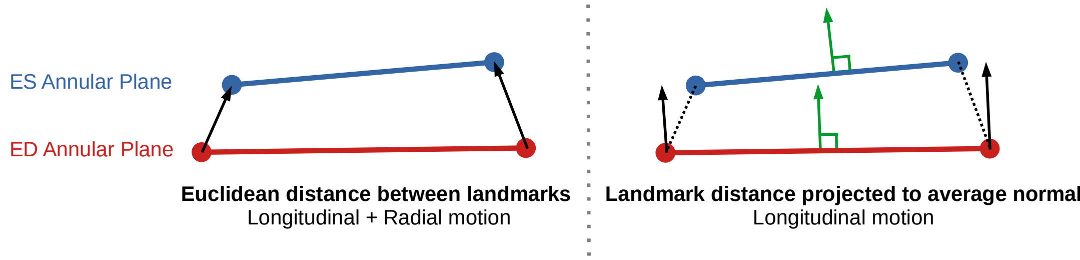

Fig. 3. Two different approaches to calculate MAPSE using detected landmarks in B-mode images. The simplest approach is to measure the distance between the landmarks respectively (left, black arrows). But this will include both longitudinal as well as any radial motion. The other approach (right) involves projecting the landmark displacement (black arrows) to the longitudinal direction defined by the average of the two annular plane normals (green arrows).

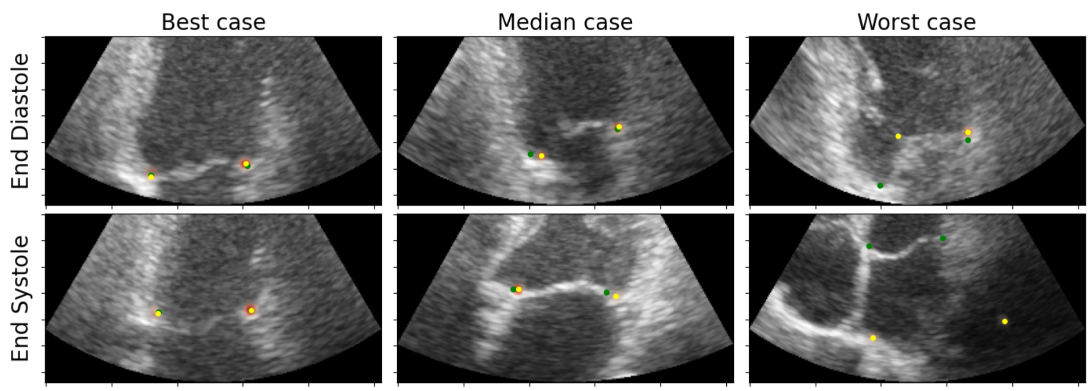

Fig. 4. Best, median and worst case landmark detection results in end-diastole and end-systole. Green points are the manually annotated landmark points, while yellow are the landmarks extracted from the heatmap shown in red.

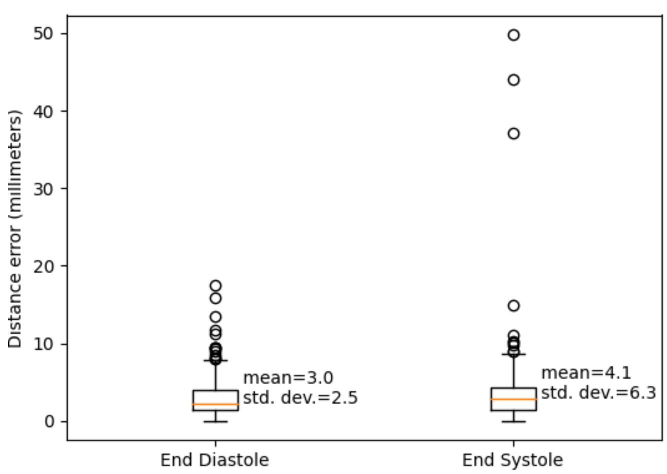

Fig. 5. Boxplot of landmark detection errors in ED and ES.

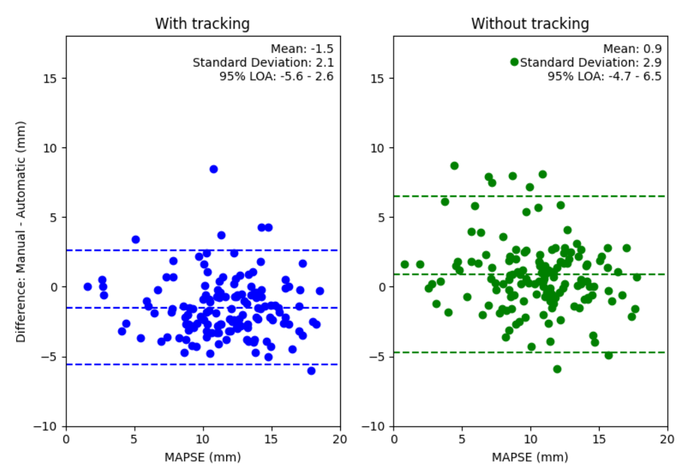

Fig. 6. Bland-Altman plot of MAPSE difference between manual B-mode measurements and proposed automatic method using with and without tracking.

our previously published Echo-PWC-Net optical flow neural network [3] which predicts a displacement field for each pair of two consecutive frames. Fig. 2 shows how the ED/ES frame detection, landmark detection and landmark tracking methods are connected to form an automatic pipeline for MAPSE.

E. MAPSE calculation

After establishing the position of the two annulus landmarks $\vec{L}$ in ED and ES, MAPSE can be calculated simply as the euclidean distance of the landmark positions in ED and ES:

$$\text{MAPSE} = |\vec{L}_{\text{ED}} - \vec{L}_{\text{ES}}| \tag{1}$$

Anatomically speaking, MAPSE should represent the displacement of the mitral annular plane. However, measuring the euclidean distance for each landmark will include any change in annulus size or global transversal movement during systole as shown in Fig. 3. For instance, one study has shown that the mitral annulus consistently enlarges in the antero-posterior direction during systole [8]. When measuring MAPSE in B-mode one can however remove any radial motion by measuring landmark displacement in the direction normal of the mitral annular plane (see Fig. 3). This longitudinal direction may also be estimated using the apex of the left ventricle. However, the apex can be hard to detect, especially when image quality is poor. Thus, we believe using the normal of the mitral plane is more robust. We calculate this MAPSE′ by calculating the mitral annular plane normal $\vec{N}$ as the average between the normal of the plane in ED and ES, and then use orthogonal projection onto the line spanned by this normal:

$$\vec{N} = \frac{\vec{N}_{\text{ED}} + \vec{N}_{\text{ES}}}{2} \tag{2}$$

$$\text{MAPSE}' = (\vec{L}_{\text{ES}} - \vec{L}_{\text{ED}}) \cdot \vec{N} \tag{3}$$

III. Results

The method was evaluated using 10-fold cross-validation. For each fold, the landmark detection and the MAPSE measurement accuracy were measured. The mean and standard deviation distance error of the annulus landmark detection network were 3.0 ± 2.5 mm in ED and 4.1 ± 6.3 mm in ES. The boxplots in Fig. 5 reveal that there are several outliers. Fig. 4 shows best, median and worst case examples of the landmark detection in ED and ES using the distance error. The proposed automatic method with landmark tracking was able to calculate MAPSE in all recordings with an accuracy and precision of -1.5 ± 2.1 mm compared to the manual B-mode reference. For comparison, we also tried to detect the annulus landmarks directly in both ED and ES without tracking which gave a MAPSE accuracy and precision of 0.9 ± 2.9 mm. This indicates that frame-by-frame tracking of the annulus improves precision. Fig. 6 shows Bland-Altman plots of the MAPSE measurement difference with and without tracking.

IV. Discussion

Performing the MAPSE measurement in B-mode using 2D tracking has the potential of eliminating the disadvantages of M-mode based MAPSE such as "out-of-line" motion and M-mode angle dependency. In this work, we have demonstrated that this is possible with an accuracy of -1.5 ± 2.1 mm using deep learning methods for annulus landmark detection and tracking. This is similar to the accuracy reported in the study of Storve et al. [1] (-0.6 ± 2.1 mm) which used tissue Doppler for tracking.

By fully automating the measurement, inter- and intra-observer variability and time consumption may also be reduced. Automation of MAPSE can also provide measurements over multiple heart cycles, even in real-time. Still, a limitation of this study is the small dataset used for both training and testing, and a bigger study is required to investigate this method further.

As seen in the boxplots in Fig. 5, the landmark detection network has several outliers, thus there is room for improvement. Fig. 4 depicts some of these outliers, in which the worst (Fig. 4, bottom far right) is from an apical four-chamber view, not focused on the left ventricle, but with a scan depth that reaches past the posterior atrial walls. This type of view was underrepresented in the training data. The landmark detection can probably be improved with more training data, and by using more advanced landmark detection networks which uses more than a single frame as input. Also, the tracking network used in this study is computationally expensive as it performs tracking in the entire image although MAPSE only requires tracking of the annulus. Thus, there is room for improving runtime speed using a tracking method that only tracks the annulus.

V. Conclusion

A fully automatic deep learning method for measuring MAPSE directly from B-mode images was proposed. The method showed a promising MAPSE accuracy of -1.5 ± 2.1 mm. Still, a study with a bigger dataset is required to validate and develop the method further.

References

[1] S. Storve, J. F. Grue, S. Samstad, H. Dalen, B. O. Haugen, and H. Torp, "Realtime Automatic Assessment of Cardiac Function in Echocardiography," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 3, pp. 358–368, 2016.

[2] 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. [Online]. Available: https://ieeexplore.ieee.org/document/8579886/

[3] A. Ostvik, I. M. Salte, E. Smistad, T. M. Nguyen, D. Melichova, H. Brunvand, K. Haugaa, T. Edvardsen, B. Grenne, and L. Lovstakken, "Myocardial Function Imaging in Echocardiography Using Deep Learning," *IEEE Transactions on Medical Imaging*, vol. 40, no. 5, pp. 1340–1351, may 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9335592/

[4] E. Evain, Y. Sun, K. Faraz, D. Garcia, E. Saloux, B. L. Gerber, M. De Craene, and O. Bernard, "Motion Estimation by Deep Learning in 2D Echocardiography: Synthetic Dataset and Validation," *IEEE Transactions on Medical Imaging*, vol. 41, no. 8, pp. 1911–1924, 2022.

[5] E. Smistad, A. Østvik, and L. Lovstakken, "Annotation Web - An open-source web-based annotation tool for ultrasound images," in *IEEE International Ultrasonics Symposium, IUS*, 2021.

[6] A. M. Fiorito, A. Østvik, E. Smistad, S. Leclerc, O. Bernard, and L. Løvstakken, "Detection of Cardiac Events in Echocardiography using 3D Convolutional Recurrent Neural Networks," in *2018 IEEE International Ultrasonics Symposium (IUS)*, 2018.

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

[8] J. Kwan, M. J. Jeon, D. H. Kim, K. S. Park, and W. H. Lee, "Does the mitral annulus shrink or enlarge during systole? A real-time 3d echocardiography study," *Journal of Korean Medical Science*, vol. 24, no. 2, pp. 203–208, 2009.