EigenBoundaries for temporally regularized segmentation of echocardiographic images

Georgios Tziritas

Computer Science Department, University of Crete
tziritas@csd.uoc.gr

Abstract. Automatic segmentation of echocardiography videos is critical for assessing various cardiac functions and improving the diagnosis of cardiac diseases. Convolutional Neural Networks (CNNs) have recently demonstrated their ability to segment 2D cardiac ultrasound images, but they suffer from temporal inconsistency. The main approaches for more reliable spatio-temporal analysis include 3D deep learning methods, recurrent segmentation algorithms, and post-processing of 2D segmentations. This paper presents a new efficient method for temporally regularized segmentation of cardiac ultrasound images. To address this problem, we propose a post-processing procedure based on the description of left ventricular, endocardial and epicardial, boundaries by their central distance signatures, which emerges as a powerful representation with excellent temporal coherence. Principal component analysis of the boundary signature is shown to provide a concise model for its representation. In particular, the coefficient of the first component has an interesting functional interpretation. Temporal smoothing of the obtained parametric representation enforces the temporal consistency of the segmentation. Results are given on the TED dataset [11] to illustrate the regularized segmentation and to measure some anatomical cardiac features.

Keywords: echocardiography · segmentation · shape signature · principal component analysis · regularization · anatomical cardiac features.

1 Introduction

Echocardiographic image analysis is an essential diagnostic tool for assessing cardiac function [8], [13]. Correctly performing automated segmentation of the heart's regions of interest also allows measurement of important parameters, such as left ventricle (LV) ejection fraction, LV area, LV length and myocardial area. Deep learning in medical image analysis, including echocardiographic image segmentation, has been the focus of the scientific community in recent years, especially when incorporating temporal along with spatial information. An evaluation of the clinical applicability of deep learning methods for dynamic cardiac imaging using spatiotemporal image information was presented in [3].

Following the exploration of 2D deep learning approaches, several methods have been developed in recent years to take into account the temporal aspect.

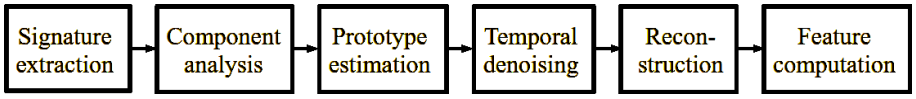


Fig. 1. Flow chart of the proposed post-processing method.

In [7] different methods are compared and it is shown that a generic 3D nnUnet CNN for the entire sequence gives better results on a private data set. In [14], a co-learning mechanism is proposed to explore the mutual benefits of cardiac segmentation and myocardium tracking at the appearance and shape levels in a semi-supervised framework. A proxy- and kernel-based video segmentation method for left ventricle localization has also been proposed in a semi-supervised context [15]. In the context of self-supervised segmentation, multi-level semantic adaptation was proposed for few-shot segmentation on cardiac image sequences [2]. A post-processing framework for learning the 2D+time apical long-axis shape of the heart to improve the accuracy and consistency of segmented echocardiographic sequences has also been proposed [11].

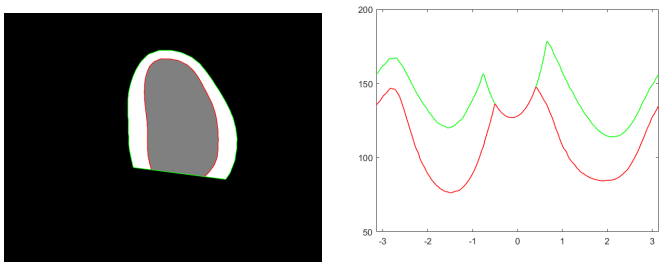
This article presents a new post-processing method for temporal regularization of cardiac ultrasound images. To address this issue, we propose to use the left ventricular boundaries and model their central distance signature using Principal Component Analysis (PCA), assuming that the ultrasound images are aligned. The flow chart of the proposed method is shown in Fig. 1. The PCA is based on the ground truth segmentation maps and is presented in Section 2. Since the first component is dominant, a prototype for its temporal evolution is estimated (Section 3). The temporal inconsistency appears clearly on the first component time sequence. Therefore, denoising the sequence of the first component results in a temporally consistent segmentation (Section 4). After reconstructing temporally smooth segmentation maps, anatomical features can be estimated (Section 5).

We evaluated our approach by testing it on the predictions of some deep CNN architectures [5], [6], [9], [10] and [12], resulting in improved left ventricular endocardial and epicardial segmentation. Finally our method provides a simple and efficient approach to achieve an accurate and temporally consistent segmentation of ultrasound images of the heart. Results are shown on the TED dataset [11][1].

2 Representation basis computation

The shape of any region can be described by its boundary. A one-dimensional periodic signal can then be used to represent the resulting contour.Among the various approaches used to represent shapes, the central distance signature has interesting properties for nearly convex closed curves. In particular, it is covariant with respect to scale and rotation transformations and invariant with respect to translation. Therefore, the central distance signature is suitable for modeling

[1] https://humanheart-project.creatis.insa-lyon.fr/databases.html



(a) (b)

Fig. 2. Endocardium and epicardium boundaries (a) and their signatures (b).

the shape of the left ventricular, endocardial and epicardial boundaries, as the main transformations during the cardiac cycle are translation, rotation and scale variation.

Let $(x(\theta), y(\theta))$ be the boundary coordinates, with an appropriate location parameter θ. The central distance is defined as

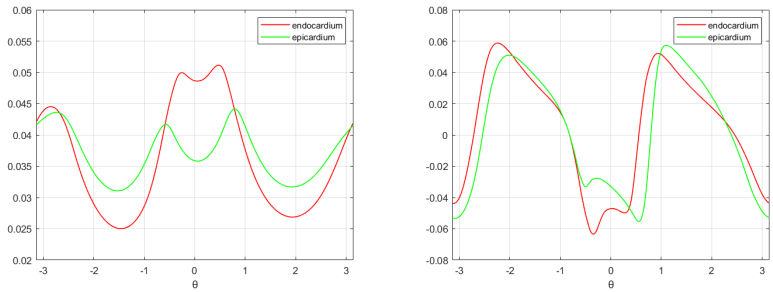
$$\rho(\theta) = \sqrt{(x(\theta) - x_c)^2 + (y(\theta) - y_c)^2}, \tag{1}$$

where (x_c, y_c) is a central shape point. Since we are considering a representation basis for an entire set of boundaries, the position parameter θ is defined by the angular position, which is determined in exactly the same way over the entire set. Therefore, the entire set of signatures is computed at the same angular positions with a fixed angular interval step. Furthermore, for both the endocardial and epicardial boundaries of the left ventricle, the center is extracted from the epicardial boundary. Then, adaptive representations can be computed separately for endocardium and epicardium.

The basis of the representation is determined from the ground truth segmentation to be the most expressive and accurate. An example of a ground truth segmentation and corresponding signatures of endocardium ('red') and epicardium ('green') is shown in Fig. 2. The angular interval step is set to 1/2 degree, so we have $K = 720$ angular positions, from $-\pi$ to π, to calculate the distance of the boundary points from the origin (x_c, y_c). Distance is given in pixels in Fig. 2(b), and the pixel size is $[0.308, 0.308]$.

Let Y be the $N \times K$ matrix containing the set of N known signatures from the ground truth segmentations. The basis of the representation is determined by the eigenvectors of the matrix $C = Y^t Y$. We will call these eigenvectors the *eigenboundaries*. Since the number of significant eigenvalues is limited, the basis can be restricted to the eigenvectors corresponding to the most significant eigenvalues. Exactly, the cumulated sum of the first seven eigenvalues are

$$\begin{array}{ccccccc} 0.988 & 0.992 & 0.996 & 0.998 & 0.998 & 0.999 & 0.999 \\ 0.996 & 0.998 & 0.999 & 0.999 & 1.000 & 1.000 & 1.000 \end{array}$$



(a) (b)

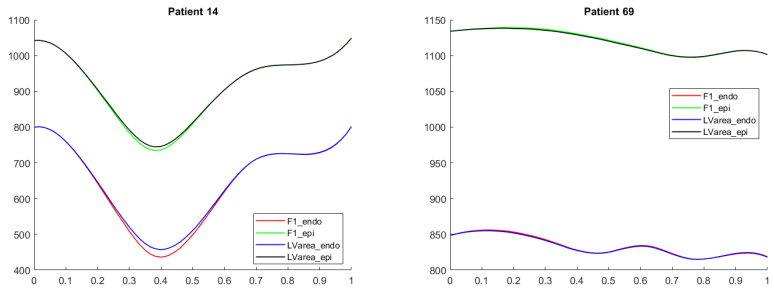
Fig. 3. The two most important eigenboundaries for the representation of the endocardium and epicardium boundaries.

The first row shows the coefficients for the endocardium and the second row shows the coefficients for the epicardium. Note that only the first component explains approximately 0.99 of the total. Based on these observations, our experiments retain 36 components for endocardium and epicardium. This number is more than sufficient for the ground truth, and at the same time limiting the number of components to 5% results in noise reduction at the boundaries of real machine segmentations.

The two main eigenboundaries for the endocardium ('red') are shown in Fig 3. Note that the two peaks near the angle origin in Fig. 3(a), where the first eigenboundary is shown, correspond to the two endpoints of the valve, while the peak near the angle $-\pi$ gives the approximate position of the apex. In the same figure are shown in 'green' the two first eigenboundaries for the epicardium, with the same observations regarding the characteristic points.

Comparing the typical LV signatures shown in Fig. 2 with the first eigenboundaries, it can be concluded that they are quite similar, the similarity being measured by the first eigenvalue mentioned above. On the ground truth dataset, the standardised signature is always close to the first eigenboundary.

Having the basis of the representation, we can represent each boundary with the corresponding set of expansion coefficients. Then, for each patient and for the corresponding boundary sequence, we can have the sequence of coefficients of the representation. Fig. 4 shows, for two patients, the sequences of coefficients for the principal eigenboundary, both for the endocardium ('red') and for the epicardium ('green'), for one cardiac cycle. It can be seen that the shape of the cardiac function follows a systolic phase with a decreasing coefficient and a diastolic phase with an increasing coefficient. In fact, the minimum position of the sequence of coefficient values gives the approximate position of the end of the systolic phase. An exception is patient number 69, in whom the ejection fraction was extremely low, equal to 13, while the ejection fraction of patient 14 is equal to 75.



(a) (b)

Fig. 4. The coefficient of the most important eigenvector for the endocardium and the epicardium during one cardiac cycle versus the square root of the corresponding area.

The same figure shows the corresponding sequences of the square root of the left ventricular area and the area defined by the epicardium. Note the extremely high correlation between the most important coefficient and the area of the respective regions for the ground truth. Indeed, the first coefficient is

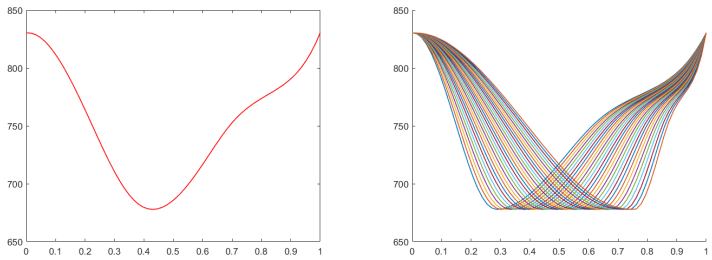
$$\sum_{k=1}^{K} \rho(\theta_k)x(\theta_k) \approx \sqrt{\sum_{k=1}^{K} \rho^2(\theta_k)} \approx \sqrt{\frac{KA}{\pi}}, \tag{2}$$

where $\{x(\theta_k), 1 \leq k \leq K\}$ is the first eigenboundary and A is the area of the corresponding region. The distance between the epicardial coefficient and the endocardial coefficient is an approximation of the thickness of the myocardium.

3 Model construction for the eigenvector coefficients

Since the most important eigenvector explains about 99% of the energy of the boundary representation, it is interesting to try to construct a model of the sequence of coefficients of this eigenvector, as this will help to make the segmentation sequence temporally consistent. To construct a generic model, the average coefficient sequence over the entire ground truth set is calculated and used to estimate an initial model for each individual sequence. The average sequence for endocardium is shown in Fig 5(a). Corresponding average epicardial sequence is similar. We observe that the systolic phase lasts on average about 45% of the time of the cardiac cycle and the diastolic phase about 55%.

However, for individual patients the variation in the division of the total cycle time into systolic and diastolic phases is very large. In order to create a model suitable for each patient, both the position of the end of the systolic phase and the amplitude of the variation for the two phases should be adjusted. If we consider different time division positions for the two phases, we obtain



(a) (b)

Fig. 5. (a) Average coefficient of the first eigenboundary for the endocardium and (b) adaptation to the individual duration of the two phases of the cardiac cycle.

appropriate models, as shown in Figure 5(b) for the endocardium. The duration of the systolic phase ranges from 30% to 75% of the total cardiac cycle. Similar results are obtained for the epicardium.

For each individual sequence, it is necessary to adjust the separation of the two phases and the amplitude of the change from end diastole to end systole. Three main parameters determine the pattern of the individual sequence: the time division of the two cardiac phases, the mean size of the ventricular area, and the ratio of the area between the end of the diastolic phase and the end of the systolic phase. Therefore, a linear model can be assumed, given the division into systolic and diastolic phases

$$c(t) = b_1 + b_2 f_k(t), \tag{3}$$

where b_1 corresponds to the mean size of the ventricular area, b_2 corresponds to the total change in area, and k indicates the relative duration of the two phases in the heart cycle. $f_k(t)$ is the prototype model shown in Fig. 5(b).

The linear model for the sequence $c(t)$ should be estimated in a robust way, since the data are noisy and have abrupt changes. Robust iteratively reweighted least-squares regression is used with the Cauchy fitting weight function [4]. With respect to the time division of the two phases of the heart cycle, the search is exhaustive. The model is checked for all possible values of k and the total absolute error $e(k)$ is calculated over the sequence,

$$e(k) = \sum_t |c(t) - \hat{b}_1 - \hat{b}_2 f_k(t)|. \tag{4}$$

The time position k that minimizes the error $e(k)$ is chosen. In some cases, it is preferable to estimate the parameters b_1 and b_2 separately in the two phases. The calculation method remains the same, but the number of parameters is doubled and the two estimates must be linearly combined.

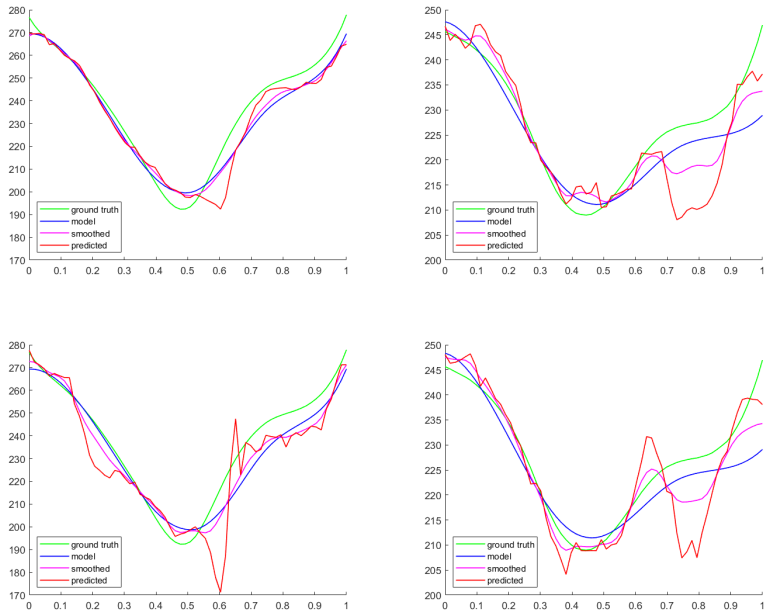
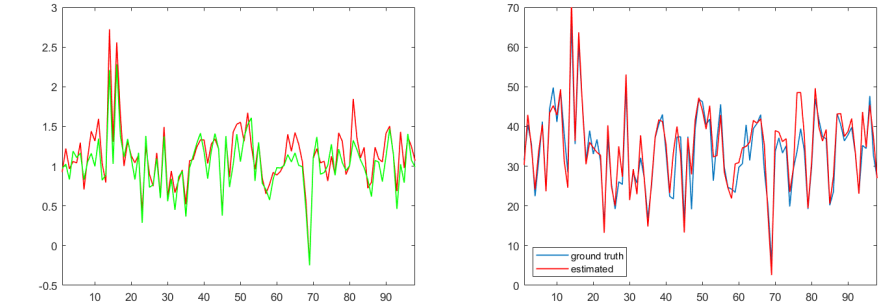


Fig. 6. Coefficient sequence for the segmentation ('red'), the ground truth ('green'), the model ('blue') and after smoothing ('magenta') for the endocardium of two patients.

Results for the endocardium for two patients are shown in Figure 6, where the GUDU segmentation method [12] was used for the plots in the first row, the ENet (resp. U-Net) predictor was used for the second row on the left (resp. right). The coefficient is given in mm. The sequence computed from the initial segmentation is shown in ('red'). For the patient on the left side, the segmentation was noisy, while on the right side there is noise and unexpected discontinuities. The prototype of the first coefficient could therefore be used for the detection of temporal inconsistencies in the segmentation results.

A consequence of the model analysis is the finding of a significant correlation between the parameter b_2 and the ejection fraction, specifically the correlation coefficient is 0.7 for the set of experimental data. The values of the estimated parameter $\hat{b}_2$, for the GUDU segmentation method, are shown in Fig. 7(a) for both the endocardium ('red' line) and the epicardium ('green' line). We see that it takes a maximum value for the patient with the maximum ejection fraction (75) and a minimum value for the sequence corresponding to the minimum ejection fraction (13). For the last sequence, where the parameter $\hat{b}_2$ takes a negative value, the above model is not used, but only a smoothed sequence is computed in the next algorithmic step, since it is clear that this sequence does not fit the generic model.



(a) (b)

Fig. 7. (a) Estimation of parameter b_2. (b) Ejection fraction after temporal smoothing.

4 Temporal regularization of the segmented images

The representation on the basis of eigenvectors allows to obtain temporally smoothed segments, since it is sufficient to smooth the values of the coefficients of the representation. Therefore, smoothing the latent representation yields the regularized segmentation.

Let $c(t)$ be the first coefficient data sequence and $s(t)$ the desired smoothed sequence. The following criterion is proposed to be minimized

$$\Omega(\mathbf{s}) = \|W(\mathbf{c} - \mathbf{s})\|^2 + \mu\|D\mathbf{s}\|^2, \tag{5}$$

where $\mathbf{c}$ is the vector of all values of the data sequence and $\mathbf{s}$ is the vector of the smoothed sequence. The matrix W is diagonal and gives weights for each time point to reduce the weight of outliers [1]. The matrix D models the second derivative to obtain a degree of smoothing μ and gives the filter d of the recurrence relation given below. The direct solution to the above optimization problem is

$$\widehat{\mathbf{s}} = (W'W + \mu D'D)^{-1}W'W\mathbf{c}. \tag{6}$$

It is proposed to implement the smoothing process by an iterative method for the coefficient corresponding to the maximum eigenvalue. The matrix W is re-estimated after each computation, while the initial estimate $f(t)$ is the model described in the previous section. The iterations, starting from $n = 1$, are formulated as follows, $*$ being the symbol of convolution,

$$s^{(n)}(t) = s^{(n-1)}(t) - \beta\left(\epsilon^{(n-1)}(t) + \mu(d * s^{(n-1)}(t))\right), \tag{7}$$

where $\epsilon^{(n-1)}(t) = w^{(n-1)}(t)(s^{(n-1)}(t) - c(t))$ and with $s^{(0)}(t) = f(t)$. The factor $w^{(n-1)}(t)$ is calculated iteratively in order to limit the consequences of the outliers. The other expansion coefficients are simply smoothed using a linear filter, without any *a priori* model.

Table 1. Accuracy of segmentation methods after (before) regularization.

	U-Net	DeepLab	LUNet	ENet	GUDU
Dice endo	0.943(0.939)	0.938(0.934)	0.940(0.934)	0.936(0.931)	0.947(0.944)
Dice epi	0.965(0.963)	0.962(0.959)	0.963(0.961)	0.957(0.954)	0.966(0.965)
Dice myo	0.877(0.870)	0.871(0.864)	0.876(0.871)	0.858(0.852)	0.889(0.885)
Hausdorff endo	3.87(4.13)	4.21(4.62)	4.07(4.38)	4.37(4.77)	3.70(3.97)
Hausdorff epi	3.85(4.16)	4.21(4.79)	4.06(4.58)	4.70(5.18)	3.73(3.90)
Hausdorff myo	4.32(4.82)	4.77(5.57)	4.65(5.32)	5.18(6.05)	4.20(4.61)
EF MAE	2.65(2.74)	2.65(2.85)	2.85(3.08)	2.87(3.23)	2.65(2.71)
EF correlation	0.943(0.934)	0.946(0.938)	0.937(0.932)	0.942(0.926)	0.939(0.940)

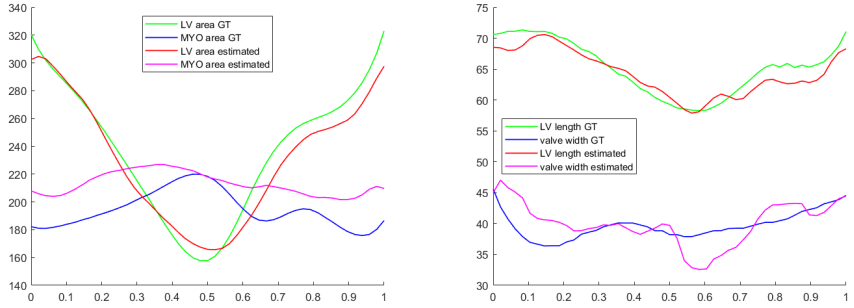
Five deep learning segmentation methods are evaluated using statistical and clinical indices on the TED dataset: U-Net [5], LUnet [6], DeepLab [9], ENet [10] and GUDU [12]. The results after regularization are given in Table 1 in comparison to the results before regularization, which are given in parentheses. The Hausdorff distance is measured in mm. In addition, Fig. 6 shows the smoothed coefficient sequence for two patients for qualitative evaluation, compared to the initial sequence, the prototype and the ground truth. Comparing the regularized result in the first row with that in the second row, we see that quite different segmentation results provided by different predictors yield very similar temporally consistent segmentations.

Although the original results showed significant differences between the algorithms, after smoothing the methods have similar performances. Among the indices considered, an improvement in the Hausdorff distance is obtained for all methods. The number of iterations in Eq. (7) was equal to 30 in almost all cases, with $\beta = 0.05$. A higher number of iterations results in a better Dice coefficient and Hausdorff distance, but the EF measurement is worse.

The 2D ejection fraction (EF) is included in the results as it can be calculated after constructing the smoothed segmentations given the end-systolic and diastolic positions. Mean Absolute Error (MAE) and correlation to ground truth are measured for 2D EF. A slight improvement is observed after temporal regularization. Fig. 7(b) compares the GUDU segmentation calculated values ('red') with those of the ground truth ('blue').

5 Anatomical feature computation

Anatomical features can be calculated from the reconstructed boundary signatures and segmentation maps after smoothing. For comparison with the true values, it is also necessary to compute them from the ground truth segmentation. The measured anatomical features are those considered in [11]: left ventricular (LV) area, myocardial (MYO) area, long axis length, valve width, left ventricular orientation and epicardial (EPI) center position. For the latter, the global displacement due to cardiac function is of interest.



(a) (b)

Fig. 8. Anatomical results for one patient : (a) LV and MYO area. (b) LV length and valve width.

The two areas, the orientation and the centroid are computed in MATLAB® using the function *regionprops*. The entire region of the left ventricle, including the myocardium, is used to calculate the orientation and the centroid. The apex of the left ventricle is localized as the top point of the corresponding segment, since the left ventricle is oriented vertically, as shown in Fig. 2(a). Therefore, to locate the apex, the endocardial border is used. Long axis length is measured as the distance from the center of the base to the apex. The extreme points of the valve are determined directly from the segmentation map for the ground truth and original segmentations.

For the regularized segmentations the extremities of the base are determined by the two vertices of the signature located around the angle origin, as shown in Fig. 2(b) for a ground truth segment, to measure the valve width. The fact that the boundaries of the endocardial and epicardial regions coincide at the valve position is also taken into account for the predicted segment. Therefore, the valve is localized differently for the original segmentation and the reconstructed one after smoothing.The comparison of the corresponding measures may be slightly affected.

Fig. 8 shows the results for the GUDU segmentation method for one patient. In (a), the estimated LV and MYO areas are plotted in square mm compared to the ground truth. In (b), the estimated LV length and valve width in mm are shown, also compared to the ground truth.

Table 2 shows the average error in the calculation of anatomical features for the five segmentation algorithms. The absolute mean error is computed for orientation, in degrees. The relative error is measured for all other features

$$\frac{\sum_t |\phi_s(t) - \phi_g(t)|}{\sum_t \phi_g(t)}, \tag{8}$$

Table 2. Errors in the estimation of anatomical attributes.

	U-Net	DeepLab	LUNet	ENet	GUDU
LV area	0.051(0.057)	0.060(0.071)	0.067(0.078)	0.056(0.065)	0.054(0.059)
Valve width	0.075(0.070)	0.085(0.083)	0.083(0.081)	0.081(0.083)	0.072(0.063)
Long axis length	0.025(0.028)	0.027(0.032)	0.024(0.027)	0.030(0.035)	0.026(0.029)
Orientation	2.443(2.571)	2.647(2.800)	2.252(2.392)	2.366(2.583)	2.218(2.320)
MYO area	0.057(0.064)	0.077(0.082)	0.100(0.106)	0.091(0.098)	0.086(0.086)
EPI hor center	0.009(0.009)	0.010(0.010)	0.010(0.010)	0.012(0.012)	0.008(0.008)
EPI ver center	0.015(0.014)	0.018(0.018)	0.018(0.018)	0.018(0.018)	0.014(0.014)

where $\phi_s(t)$ is the segmentation-estimated feature value after smoothing and $\phi_g(t)$ is the feature value as computed from the ground truth segments. We observe that the results after regularization are almost equivalent for all methods. Compared to the measurements before regularization, an improvement is observed for the LV and MYO areas, the long axis length and the orientation. The errors in estimating the center are small anyway. The difference in valve width estimation error should be mainly due to the different estimation method.

6 Discussion and conclusion

All the algorithmic modules are implemented in MATLAB® and tested on an HP Pavilion laptop (Intel® i7 CPU at 2.60GHz). The computation time per patient was about 14 seconds. The first modules of signature computation, coefficient extraction and prototype fitting, using *robustfit* function, take about 10.5 seconds due to the exhaustive search for the model. The last modules of regularization, reconstruction, feature estimation and evaluation take about 3.5 seconds.

We have introduced a generic model for representing the left ventricular boundaries, endocardium and epicardium, in long-axis apical echocardiographic images. Based on principal component analysis of the boundary signatures, we have shown that a concise set of parameters describes the boundaries. In particular, the coefficient of the first eigenboundary has strong correlation with the left ventricular area and its temporal variation is related to the cardiac cycle. We proposed an iterative least squares method to smooth this coefficient over time. The reconstructed segmentation maps are then nearly time consistent and are used to estimate anatomical features. The main improvements obtained on five different segmentation results on the TED dataset are in a global temporal consistency and in the measurement of the Hausdorff distance and the anatomical features.

Acknowledgments

The author would like to thank N. Painchaud *et al.* [11] for providing the predictions of their U-Net, DeepLab, LUNet and ENet methods on the full cycle

ultrasound sequences to evaluate the proposed approach, and C. Sfakianakis and G. Simantiris for the predictions of the GUDU method. The segmentations used in this work are the result of the method proposed in [12] with a richer set of data augmentations.

References

1. Barron, J., A general and adaptive robust loss function, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
2. Guo, S., *et al.*, Multi-level semantic adaptation for few-shot segmentation on cardiac image sequences, *Medical Image Analysis*, Vol. 73, 2021.
3. Hernandez, K., Rienmüller, T., Baumgartner, D., Baumgartner, C., Deep learning in spatiotemporal cardiac imaging: A review of methodologies and clinical usability, *Computers in Biology and Medicine*, Vol. 130, 2021.
4. Holland, P., Welsch R., Robust regression using iteratively reweighted least-squares, *Communications in Statistics: Theory and Methods*, pp. 813-827, 1977.
5. Leclerc S., Smistad E., Pedrosa J., *et al.*, Deep learning for segmentation using an open large-scale dataset in 2D echocardiography, *IEEE Transactions on Medical Imaging*, Vol. 38, pp. 2198-2210, 2019.
6. Leclerc S., Smistad E., *et al.*, LU-Net: a multistage attention network to improve the robustness of segmentation of left ventricular structures in 2-D echocardiography, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 67, pp. 2519-2530, 2020.
7. Ling, H.J. *et al.*, Extraction of volumetric indices from echocardiography: which deep learning solution for clinical use?, *Functional Imaging and Modeling of the Heart*, 2023.
8. McDonagh, T. A., *et al.*, 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure, *European Heart Journal*, Vol. 42, pp. 3599–3726, 2021.
9. Ouyang D., He B., Ghorbani A., *et al.*, Video-based AI for beat-to-beat assessment of cardiac function, *Nature*, Vol. 580, pp. 252–256, 2020.
10. Painchaud, N., Skandarani, Y., Judge, T. *et al.*, Cardiac segmentation with strong anatomical guarantees, *IEEE Transactions on Medical Imaging*, Vol. 39, pp. 3703-3713, 2020.
11. Painchaud, N., Duchateau, N., Bernard, O., Jodoin, P.-M., Echocardiography segmentation with enforced temporal consistency, *IEEE Transactions on Medical Imaging*, Vol. 41, pp. 2867–2878, 2022.
12. Sfakianakis, C., Simantiris, G., Tziritas, G., GUDU: Geometrically-constrained ultrasound data augmentation in U-Net for echocardiography semantic segmentation, *Biomedical Signal Processing and Control*, Vol. 82, 2023.
13. Smiseth O., Torp H., Opdahl A., Haugaa K., Urheim S., Myocardial strain imaging: How useful is it in clinical decision making, *European Heart Journal*, Vol. 37, pp. 1196–1207, 2016.
14. Wei H., *et al.*, Co-learning of appearance and shape for precise ejection fraction estimation from echocardiographic sequences, *Medical Image Analysis*, Vol. 84, 2023.
15. Wu, H., Lin, J., Xie, W., Qin, J., Super-efficient echocardiography video segmentation via proxy- and kernel-based semi-supervised learning, *37th AAAI Conference on Artificial Intelligence*, 2023.