Motion Estimation for Automatic Measurement of Left Ventricular Strain in Echocardiography

Sofia Ferraz[1,2]
sofia.a.ferraz@inesctec.pt

Miguel Coimbra[2,3]
mcoimbra@fc.up.pt

João Pedrosa[1,2]
joao.m.pedrosa@inesctec.pt

[1] Faculty of Engineering of the University of Porto (FEUP)

[2] Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)

[3] Faculty of Sciences of the University of Porto (FCUP)

Abstract

Motion estimation in echocardiography is critical when assessing heart function and calculating myocardial deformation indices. Nevertheless, there are limitations in clinical practice, particularly with regard to the accuracy and reliability of measurements retrieved from images. In this study, deep learning-based motion estimation architectures were used to determine the left ventricular longitudinal strain in echocardiography. Three motion estimation approaches, PWC-Net, RAFT and FlowFormer, were applied to a simulated echocardiographic dataset, achieving an average end point error of 0.24, 0.22 and 0.21 mm per frame, respectively. Thus, optical flow-based motion estimation has the potential to facilitate the use of strain imaging in clinical practice.

1 Introduction

Cardiac motion estimation from ultrasound (US) images is an essential tool for the diagnosis of cardiovascular disease [5]. Currently, speckle tracking echocardiography (STE) is frequently used to non-invasively image heart displacement and strain with sufficient temporal resolution. Despite being considered the standard, there is a significant clinical concern related to the reported low reproducibility between the solutions offered by various vendors [1]. Additionally, these strategies face a number of obstacles due to fundamental limitations of US acquisitions, such as dropouts, shadows, out-of-plane motion, drift sensitivity and foreshortening. The US speckle pattern is affected by several of these distortions, which makes tracking more difficult [8].

The use of deep learning based motion estimation in US, and especially in echocardiography, is still limited. In [4], the authors proposed a modified version of Pyramidal processing, Warping, and Cost volume Network (PWC-Net) and generated a synthetic dataset to measure global longitudinal strain (GLS) in echocardiograms. When compared to the GLS produced through manual segmentation, the approach showed promise, with a mean absolute error of 2.5 ± 2.1% and a correlation of 0.77. Furthermore, [8] developed a novel DL-based framework for motion estimation in echocardiography, based on a PWC-Net architecture, with the goal of fully automating myocardial function imaging. Using simulated data from an open access database, the motion estimator obtained an average end point error of (0.06 ± 0.04) mm per frame. Similarly, [2] sought to fully automate measurements of GLS, producing comparable results to a commercially available semi-automated speckle-tracking method. Motion estimation was once again based on a modified PWC-Net, trained using synthetic echocardiography images where the true motion was known. All steps in the artificial intelligence (AI) pipeline were performed in $<$ 15 s, eliminating the need for time-consuming manual input. Nevertheless, the documented low reproducibility across the solutions provided by different manufacturers is still a serious clinical concern. The intrinsic limitations of US acquisitions, including as dropouts, shadows, out-of-plane motion, drift sensitivity, and foreshortening, also provide a variety of challenges for these approaches [8].

The main purpose of this work is to explore the viability of applying cutting-edge computer vision techniques for motion estimates to 2D echocardiography, with the ultimate goal of fully automatic strain analysis. The utilization of optical flow (OF) for motion estimation served as the main premise behind this research.

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within projects THOR, with reference DSAIPA/AI/0083/2020, and LA/P/0063/2020. João Pedrosa is supported by Fundação para a Ciência e Tecnologia (FCT) through application reference 2022.06138.CEECIND.

2 Methods

2.1 Datasets

Two publicly available datasets commonly used for training and benchmarking of OF methods, FlyingChairs2D and FlyingThings3D, were used in the models pre-trained weigths. The datasets consist of image pairs and a corresponding dense displacement map.

For testing, the publicly available dataset of simulated echocardiography images [1], consisting of 105 sequences or 6,165 frames was used. The data is created with a complex biomechanical model for comparison of speck tracking imaging algorithms. In total, the data is composed of templates from 7 different vendors, 3 echocardiographic views and 5 motion patterns from the biomechanical model, including one healthy and four pathologies. A set of 180 points divided among five longitudinal lines and 6 segments is provided by the authors for each frame [1].

2.2 Automated Strain Analysis using Deep Learning

An automated strain quantification method using OF-based DL networks was developed. The proposed workflow for this work is summarized in the following steps:

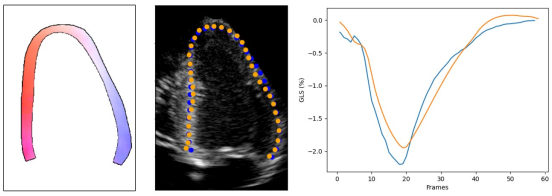
1) Data Pre-processing: To reduce the image size, the simulated US images were cropped around the left ventricle. This allowed for less computational effort, while maintaining the image original resolution. The size of the input images varies across the seven different vendors.

2) Motion Estimation: OF-based motion estimation networks were used to estimate each pixel movement field of the heart from two consecutive images (I_t, $I_t + 1$). In this work, three different variants (PWC-net [6], RAFT [7] and FlowFormer [3]) were implemented to identify the optimum basic architecture. PWC-Net [6] estimates OF in a coarse-to-fine way with several pyramid levels using CNNs, and constructs a cost volume with a feature map from the source image and the warped feature map from the target image based on the current flow. Nevertheless, PWC-Net inherently suffered from missing small fast-motion objects in coarse stage. To remedy this issue, [7] proposed RAFT, which performs OF estimation in a coarse-and-fine (i.e. multi-scale search window in each iteration) and recurrent manner. On the other hand, FlowFormer [3] adapts transformer architectures to effectively process cost volumes. Currently, FlowFormer is currently ranked within the top 10 models on the Sintel benchmark. All the methods produce a dense displacement map of velocity components v(x,y) = v_x, v_y between I_t and $I_t + 1$ (Figure 1.a). These output dense displacement maps are later used to update the position of left ventricle mid-centerline.

3) Centerline Extraction: The centerline C of the myocardium on the initial image frame was used as a starting point for tracking. The centerline is defined as the mid-point between two nearest endo- and epicardial points on the line perpendicular to the longitudinal. Furthermore, the mid, base and apex points are defined as the points furthest away from the LV lumen centroid, in left bottom, right bottom and top direction respectively.

4) Tracking Update: In this step, the centerline coordenates are updated by propagating the points with the displacement field, i.e. C(t + 1) = C(t)+ f(t), and its arc-length representing the longitudinal length of LV myocardium is calculated at each timestep/frame (Figure 1.b.

5) Strain Measurements: Following, the calculation of longitudinal ventricular length, the global longitudinal strain is determined for each timestep (Figure 1.c): $GLS = \frac{L(t)-L(0)}{L(0)}$.. The reference length L_0 is measured at the initial frame. The peak-systolic strain was used for GLS estimation, where the peak was defined as the minima of the strain values.



(a) (b) (c)

Figure 1: (a) Example of predicted flow patterns for the FlowFormer method within the myocardium. The hue values, color and saturation indicate direction and magnitude respectively. (b) Velocity vector comparison between the ground truth (GT) and the FlowFormer method. Blue dots represent the GT, while orange dots represent predictions. (c) Strain measurements derived from the tracking results.

2.3 Evaluation

The end point error (EPE), which is a common metric for benchmarking motion estimation performance, was used to evaluated the models. It is defined as the Euclidean distance between the GT velocity and the predictions, i.e. EPE = $||v_G T - v_p red||$ [8].

Moreover, strain measurements are used to describe the degree of deformation of the myocardium, providing quantitative evaluations of motion estimates. Correlation metrics, such as regression slope α and correlation coefficient ρ, as well as bias μ, are reported for the strain measurements [1].

3 Results and Discussion

The average EPE with corresponding standard deviation can be seen in Table 1. The smallest EPE for each model are 0.0782, 0.0916 and 0.0781 for PWC-Net, RAFT and FlowFormer, respectivelly. FlowFormer had lower EPE in comparison to PWC-net and RAFT, and PWC-Net had a better performance than RAFT.

Vendor	PWC-Net	RAFT	FlowFormer
Hitachi	0.0886 ($\pm$ 0.0045)	0.1006 ($\pm$ 0.0083)	**0.0789 ($\pm$ 0.0066)**
Toshiba	0.0924 ($\pm$ 0.0053)	0.1029 ($\pm$ 0.0101)	**0.0836 ($\pm$ 0.0076)**
ESAOTE	0.1109 ($\pm$ 0.0065)	0.1088 ($\pm$ 0.0083)	**0.0956 ($\pm$ 0.095)**
Samsung	**0.1043 ($\pm$ 0.0096)**	0.1373 ($\pm$ 0.0185)	0.1073($\pm$ 0.0177)
Siemens	0.0782 ($\pm$ 0.0052)	0.0916 ($\pm$ 0.0097)	**0.0781 ($\pm$ 0.0103)**
Philips	0.0952 ($\pm$ 0.0056)	0.1093 ($\pm$ 0.0109)	**0.0876 ($\pm$ 0.0097)**
GE	0.1269 ($\pm$ 0.0067)	0.1356 ($\pm$ 0.0119)	**0.1057 ($\pm$ 0.0127)**

Table 1: Results on simulated ultrasound data. Average EPE for every vendor. Units given in mm per timestep/frame.

Since the FlowFormer had the best performance regarding motion tracking, only the results of this method were used to assess the strain correlation. The GLS was calculated across all frames for all sequences. Of note, there is a variability in strain accuracy between vendors. Overall, the tested algorithm seemed to perform better on simulated Toshiba images (α closest to 1, highest lowest ρ and low μ) and slightly worse on simulated Samsung images (lowest α AND lowest ρ; Table 2). This is a direct consequence of the high variability between image appearance from different vendors and shows the difficulty in developing techniques designed to work on generic systems.

All three types of networks perform well at estimating motion, however it is unknown which is best for echocardiography. The applications of PWC-Net in myocardial movement, which presented a refinement method that optimizes flow prediction based on the pyramidal layers, have been detailed by [8], [2] and [4]. Recent results, however, imply that the spatial pyramid architecture may disregard small, quick movements. This design might not be ideal for this application because the myocardium has small displacement magnitudes between frames. In this work, RAFT and FlowFormer were compared to PWC-net, and ultimately the FlowFormer was shown to have the best performance in motion estimation (Table 1).

Although the results are promising, there are a few points that should be mentioned as areas that need more work. Common datasets utilized in OF research, such as FlyingChairs2D and FlyingThings3D, have motion magnitudes that are often substantially larger than the normal echocardiogram data's frame-to-frame displacement. Therefore, to get better outcomes, it should be investigated whether training FlowFormer using simulated US data is a viable option. Additionally, the capacity of DL-based approaches to adapt to the data representation used during training is one of the primary reasons for researching their application. This makes the employment of augmentation procedures that imitate common image artifacts particularly enticing since it may also be able to solve some of the major problems with conventional techniques.

Vendor	α	ρ	μ
Hitachi	0,3755	0,6947	2,8634
Toshiba	0,8359	0,8991	0,697
ESAOTE	0,7468	0,7931	-0,224
Samsung	0,2972	0,6332	2,7467
Siemens	0,7761	0,8531	-0,0645
Philips	0,8541	0,8722	0,6807
GE	0,5241	0,7944	2,2313

Table 2: Results from the FlowFormer Method Averaged Over the Different Vendors in the Simulated US Data.

4 Conclusion

Estimating cardiac motion is crucial for the diagnosis of cardiovascular diseases. The development of effective motion estimate techniques is still a challenge due to the limited annotated clinical data, US image noise and the differences among different vendors. The experimental findings demonstrated that the most recent model, FlowFormer, which is presently one of the highest-ranked networks on the MPI Sintel benchmark, performed better than the other evaluated techniques. The obtained results could be further improved by training in the simulated US dataset and use of data augmentation to increased data volume, and resemblance to clinical echocardiography. The presented pipeline has the capacity to automatically measure longitudinal strain in a prospective manner.

References

[1] Martino Alessandrini et al. Realistic vendor-specific synthetic ultrasound data for quality assurance of 2-d speckle tracking echocardiography: Simulation pipeline and open access database. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 65(3): 411–422, 2018. doi: 10.1109/TUFFC.2017.2786300.

[2] Ivar Mjåland Salte et al. Artificial intelligence for automatic measurement of left ventricular strain in echocardiography. *JACC. Cardiovascular imaging*, 2021.

[3] Zhaoyang Huang et al. Flowformer: A transformer architecture for optical flow, 2022.

[4] Ewan Evain et al. Motion estimation by deep learning in 2d echocardiography: Synthetic dataset and validation. *IEEE Transactions on Medical Imaging*, 41(8):1911–1924, 2022.

[5] Nora Ouziret al. Motion estimation in echocardiography using sparse representation and dictionary learning. *IEEE Transactions on Image Processing*, 27(1):64–77, 2018. doi: 10.1109/TIP.2017.2753406.

[6] Deqing Sun et al. Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume, 2017.

[7] Zachary Teed and Jia Deng. Raft: Recurrent all-pairs field transforms for optical flow, 2020.

[8] Andreas Østvik et al. Myocardial function imaging in echocardiography using deep learning. *IEEE Transactions on Medical Imaging*, 40(5):1340–1351, 2021. doi: 10.1109/TMI.2021.3054566.