DL-LVEF: Deep-Learning Measurement of the Left Ventricular Ejection Fraction from Echocardiographic Images

Agnese Sbrollini, MHD Jafar Mortada, Selene Tomassini, Haidar Anbar, Micaela Morettini, Laura Burattini

[1]Department of Information Engineering, Università Politecnica delle Marche, Ancona, Italy

Abstract

Left ventricular ejection fraction (LVEF) is a commonly used index of cardiac functionality. Thus, accuracy in its measurement is fundamental. LVEF measure is usually manually performed by clinicians from echocardiographic images. Use of automatic algorithms could make LVEF measurement more objective. Thus, the aim of the present work is to present DL-LVEF, a new automatic algorithm for LVEF measurement based on deep learning identification and segmentation of the left ventricular endocardium performed by combining the YOLOv7 algorithm and a U-Net. To this aim, the CAMUS database was used, which includes 1800 echocardiographic images acquired from 450 patients with annotated LVEF values and manual segmentation of the left ventricular endocardium. The database was divided into training dataset (70%) and testing dataset (30%). In both datasets, measured and annotated LVEF values (%) were found to be highly correlated (ρ=0.96 and ρ=0.89, respectively) and not statistically different (52.6% vs. 52.6% and 54.6% vs. 53.9%, respectively); mean absolute error was 4% and 5%, respectively. Thus, DL-LVEF provided objective and accurate LVEF measurement. Future DL-LVEF evolutions will also provide segmentation of other cardiac anatomical structures and, thus, will allow measurement of other clinically relevant cardiac indexes.

1. Introduction

Left ventricular ejection fraction (LVEF) is defined as the fraction of ventricular volume ejected in systole (stroke volume) in relation to the volume in the ventricle at the end of diastole [1].

As reported by the American College of Cardiology (ACC) [2], standard clinical classification of cardiac dysfunctions according to LVEF values (Table 1) is a cornerstone in cardiology, considering the LVEF a powerful predictor of cardiac mortality. Indeed, in the clinical practice, it is used as a reliable index of mortality in heart-failure patients [3-4], as an important parameter to monitor left ventricular functionality during cardiotoxic chemotherapy [5], as a primary criterion for defibrillator placement [6-7], and as a basilar indicator to determine the management of patients affected by valvular diseases [8].

The most common methodology to measure LVEF is based on the echocardiographic test. Echocardiography provides videos of the heart during all cardiac cycle phases, with the most informative frames being those acquired at end of systole and end of diastole. Nowadays, echocardiographic exam is still manually performed by clinicians who optimize the image acquisition, detect the cardiac chambers and segment anatomical structures. Thus, echocardiography and, consequently, LVEF measurement are still clinician dependent [9].

To reduce subjectiveness, automatic algorithms are desirable, even though their design is not free of challenges. Ultra-sound images are usually characterized by low signal-to-noise ratio [10]. Additionally, location and dimension of cardiac anatomical structures may act as confounders due to the intrasubject variability. Consequently, application of conventional image processing methods may face many technical issues [11], while deep learning (DL) methods may result being more efficient [12].

Thus, aim of the present work is to present DL-LVEF, a new automatic algorithm for LVEF measurement based on DL segmentation of the left ventricular endocardium.

Table 1. Standard clinical classification of cardiac dysfunctions according to LVEF values, as reported by the American College of Cardiology (ACC) [2].

Clinical Classification	LVEF value
Hyperdynamic	Higher than 70%
Normal	From 50% to 69%
Mild dysfunction	From 40% to 49%
Moderate dysfunction	From 30% to 39%
Severe dysfunction	Less than 30%

2. Materials and Method

2.1. Data

The data used in this work are from the CAMUS database [13], made open in 2019 by the University Hospital of St Etienne (France) together with the regulation set by the local ethical committee of the hospital. The database contains 1800 echocardiographic images from 450 patients acquired through the Vivid E95 ultrasound scanner (GE Vingmed Ultrasound) with a GE M5S probe (GE Healthcare, US). For each patient, manual annotations of left ventricular volume at end of diastole (LV_{ED}, ml), left ventricular volume at end of systole (LV_{ES}, ml) and LVEF (%) were collected. Moreover, two echocardiographic frame sequences showing the apical two-chamber and four-chamber views were acquired. According with the standard dimension criteria [14], end-of-systole and end-of-diastole frames were determined, and, from them, the left ventricular endocardium was manually segmented and reviewed by three independent cardiologists. Annotations of the LVEF (aLVEF) and of the manual segmentation of the left ventricular endocardium were considered here as gold standard, and thus used as reference values to evaluate performance of DL-LVEF.

The CAMUS database was divided into training dataset (70% of the entire database, 1260 images from 315 patients), also including the validation dataset (14% of the training dataset; 180 images from 45 patients), and testing dataset (30% of the entire database; 540 images from 135 patients).

2.2. The DL-LVEF Algorithm

The DL-LVEF algorithm includes two computational phases, which are 1) deep-learning identification and segmentation of the left ventricular endocardium and, 2) LVEF estimation. DL-LVEF was trained on the training dataset; the validation dataset was used to define the early stopping point. Its implementation was performed on Google Colab Pro, a cloud service allowing the possibility to select high system RAM (32 GB) and GPU hardware acceleration (NVIDIA Tesla P100 with 16GB of video RAM) settings.

Deep Learning Identification and Segmentation of the Left Ventricular Endocardium. The DL-LVEF computational phase finalized to identify and segment the left ventricular endocardium was implemented by combining the YOLOv7 [15] algorithm and a U-Net. The YOLOv7 algorithm was used to identify the anatomical structure of interest. Its architecture was maintained unchanged with respect to what proposed in [15]. In particular, initial learning rate and number of epochs were set at 0.01 and 100, respectively. Instead, the U-Net was used to segment the identified anatomical structure. U-Net architecture was proposed by ourself and consists of: an encoder composed of 5 stages providing a feature map of 20x20x512 pixels; and a decoder, also composed of 5 stages that use transpose layers to perform up-sampling. The output of this first DL-LVEF computational stage is the area (in pixels) of the left ventricular endocardium.

LVEF Computation. This DL-LVEF computational phase provides measurement of the LVEF (mLVEF) starting from the values of the left-ventricular-endocardium area previously measured on the end-of-diastole and end-of-systole echocardiographic frames. According to Simpson's rule [16,17], area and volume (in ml) of the left ventricular endocardium correlate. Thus, left ventricular volumes at the end of diastole $\widehat{LV_{ED}}$ and at the end of systole $\widehat{LV_{ES}}$ were estimated after computation of the conversion factor obtained using the annotated values ventricular area and volumes [16,17]. Finally, mLVEF was estimated as:

$$\text{mLVEF}=100\times\frac{\widehat{LV_{ED}}-\widehat{LV_{ES}}}{LV_{ED}}. \quad (1)$$

2.3. Statistical Analysis

In order to evaluate DL-LVEF performance in varying technical settings, patients were stratified according to the number of chamber views (two or four) in their images and according to the quality of their images (poor, medium and good). Additionally, in order to evaluate DL-LVEF performance in varying clinical scenarios (robustness to pathology), patients were stratified according to their aLVEF into hyperdynamic (Hyp; aLVEF≥70%), normal (Nor; 50%≤aLVEF≤69%), mild dysfunction (Mil; 40%≤aLVEF≤49%), moderate dysfunction (Mod; 30%≤aLVEF≤39%), and severe dysfunction (Sev: aLVEF< 30%).

Normality of LVEF distributions was evaluated using the Lilliefors test; not normal distributions were reported in terms of 50^{th} (median) [25^{th};75^{th}] percentiles and compared using the paired Wilcoxon ranksum test. Comparison of mLVEF and aLVEF values was performed by computing the mean absolute error (MAE, %) and Spearman correlation coefficient (ρ). Statistical level (P value) was set to 0.05 for all statistical analyses.

3. Results

Table 2 reports mLVEF and aLVEF distributions and MAE and ρ values for all considered cases, whereas Figure 1 displays the scatterplot of the distribution of mLVEF values vs the distribution of aLVEF values for the training and testing datasets. No statistically significant difference was ever observed between mLVEF and aLVEF. MAE was, on average 4% and 5% in the training and testing datasets, respectively. Its maximum value was 14%

obtained for the patients with severe dysfunction in the testing dataset. Mean ρ value was 0.96 and 0.89 ($P<0.05$) in the training and testing datasets, respectively. Its minimum value was 0.45 ($P<0.05$) obtained for the patients with moderate dysfunction in the testing dataset.

4. Discussion

In the present study we presented DL-LVEF, a new automatic algorithm LVEF measurement based on a tested deep learning segmentation algorithm of the left ventricular endocardium [18]. The left ventricular endocardium is indeed usually manually segmented by clinicians in order to measure LVEF from echocardiography.

DL-LVEF was set up and tested on the CAMUS database [13], which contains anamnestic data of patients, annotations of their LVEF values, and echocardiographic images of different chamber views with annotations of their manual segmentation and classification based on their quality. The annotated LVEF values were used here as gold standard to evaluate goodness of the LVEF measurements provided by DL-LVEF. Satisfactory results (no statistical difference, high correlation and low mean absolute error between gold standard and measured LVEF values; Table 2) were obtained in both training and testing datasets, indicating that DL-LVEF is a useful tool to accurately measure LVEF from echocardiographic frames of the end of systole and end of diastole. Good results in both training and testing datasets also indicate the generalization ability of DL-LVEF.

In order to evaluate the technical and clinical reliability of DL-LVEF, analysis of accordance between gold standard and measured LVEF values was also performed by stratifying for chamber views, image quality and clinical interpretation (Table 2). Satisfactory results were obtained in all cases, demonstrating the robustness of DL-LVEF to variable technical and clinical conditions. The relatively low values (0.45 and 0.52) of the correlation coefficient obtained in the patients with mild and moderate dysfunction are indeed due to the narrow variability range of LVEF values, that makes the correlation coefficient not particularly suitable to evaluate agreement between gold standard and measured LVEF values, as conformed by the low values of the mean absolute error (≤10%).

Future DL-LVEF evolutions will also provide segmentation of other cardiac anatomical structures in addition to the left ventricular endocardium and, thus, will allow measurement of other clinically relevant cardiac indexes.

Table 2. Distributions of measured (mLVEF) and annotated (aLVEF) LVEF values, together with their mean absolute error (MAE) and correlation coefficient (ρ) in all considered cases (NP=number of patients; NI=number of used images).

		Training dataset					Testing Dataset				
		NP/ NI	aLVEF (%)	mLVEF (%)	MAE (%)	ρ (adi)	NP/ NI	aLVEF (%)	mLVEF (%)	MAE (%)	ρ (adi)
Chamber Views	Two	315/ 630	52.6 [44.4;60.6]	52.9 [43.8;61.8]	4 [2; 7]	0.97§	135/ 270	54.6 [45.1;60.1]	53.7 [44.4;62.3]	7 [3;12]	0.89§
	Four	315/ 630	52.6 [44.4;60.1]	51.9 [44.0;60.3]	3 [2;7]	0.96§	135/ 270	54.6 [45.1;60.1]	54.4 [44.3;60.0]	5 [2;9]	0.90§
Image Quality	Poor	36/ 144	47.6 [40.2;59.9]	48.2 [37.9; 59.7]	5 [2;10]	0.94§	14/ 56	48.1 [36.1;56.2]	52.4 [36.7;57.8]	9 [4;15]	0.88§
	Medium	123/ 492	52.2 [43.6;60.4]	51.7 [43.1;60.6]	4 [2;8]	0.96§	47/ 188	54.1 [42.1;57.8]	53.7 [41.3;59.8]	5 [3;11]	0.88§
	Good	156/ 624	54.4 [46.6;61.2]	54.5 [46.3;61.1]	3 [1;5]	0.96§	74/ 296	56.9 [46.9;62.0]	54.4 [45.8;62.6]	5 [2;8]	0.89§
Clinical Classification	Hyp	17/ 68	72.5 [71.1;75.1]	72.7 [70.7;77.0]	2 [1;6]	0.60§	3/ 12	71.0 [70.4;81.9]	71.6 [65.2;79.3]	4 [3;7]	0.96§
	Nor	170/ 680	58.5 [54.1;62.3]	58.3 [54.0;62.6]	3 [1;5]	0.89§	84/ 336	57.7 [54.9;62.2]	58.8 [54.1;62.6]	5 [2;9]	0.70§
	Mil	76/ 304	46.0 [43.1;47.8]	45.9 [41.8;48.6]	4 [2;8]	0.69§	27/ 108	45.5 [42.1;46.9]	45.0 [41.3;49.0]	6 [3;10]	0.52§
	Mod	37/ 148	36.3 [33.1;38.3]	35.9 [33.0;38.5]	5 [3;11]	0.45§	15/ 60	35.5 [31.7;38.4]	33.8 [30.1;37.0]	10 [5;17]	0.45§
	Sev	15/ 60	26.8 [20.0;28.5]	26.1 [21.6;28.7]	8 [5;15]	0.68§	6/ 24	27.6 [23.7;28;3]	27.8 [23.2;32.4]	14 [5;20]	0.69§
Overall		315/ 1260	52.6 [44.3;60.6]	52.6 [43.8;60.8]	4 [2;7]	0.96§	135/ 540	54.6 [44.9;60.2]	53.9 [44.2;61.1]	5 [3;11]	0.89§

*,§P-value < 0.05 when comparing mLVEF and a distributions by paired Wilcoxon ranksum test and Spearman correlation analysis, respectively.

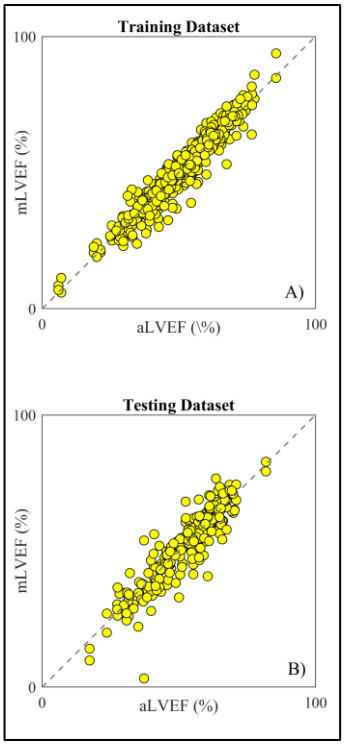

Figure 1. Scatter plot of measured (mLVEF) and annotated (aLVEF) LVEF values in both training (panel A) and testing (panel B) datasets. Dot line represent the mLVEF=aLVEF line.

5. Conclusion

Our deep DL-LVEF algorithm, based on the DL identification and segmentation of the left ventricular endocardium, proved to be an objective and accurate tool to measure LVEF from echocardiographic images. Future studies will evaluate the real-time ability of the tool in standard clinical scenarios.

References

[1] A. Kosaraju, A. Goyal, Y. Grigorova, and A. N. Makaryus, "Left ventricular ejection fraction," In: StatPearls. Treasure Island (FL): StatPearls Publishing; Jan. 2023.

[2] R. M. Lang et al. "Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *J Am Soc Echocardiogr*, vol. 28, no. 1, pp. 1-39, Jan. 2015.

[3] A. Yoshihisa et al. "Comprehensive clinical characteristics of hospitalized patients with mid-range left ventricular ejection fraction," *Eur J Prev Cardiol.* vol. 27, no. 19, pp. 2084-2088, Dec. 2020.

[4] B. Alushi et al. "Impella versus IABP in acute myocardial infarction complicated by cardiogenic shock," *Open Heart*, vol. 6, no. 1, Art. no. e000987, May 2019.

[5] J. C. Plana et al. "Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *Eur Heart J Cardiovasc Imaging.*, vol. 15, no. 10, pp. 1063-1093, Oct. 2014.

[6] E. R. Duque, A. Briasoulis, and P. A. Alvarez. "Heart failure with preserved ejection fraction in the elderly: pathophysiology, diagnostic and therapeutic approach," *J Geriatr Cardiol.*, vol. 16, no. 5, pp. 421-429, May 2019.

[7] J. C. Craig et al. "Central and peripheral factors mechanistically linked to exercise intolerance in heart failure with reduced ejection fraction," *Am J Physiol Heart Circ Physiol.*. vol. 317, no. 2, pp. H434-444, Aug. 2019.

[8] R. O. Bonow at al. "Appropriate use criteria task force. ACC/AATS/AHA/ASE/EACTS/HVS/SCA/SCAI/ SCCT/SCMR/STS 2017 Appropriate use criteria for the treatment of patients with severe aortic stenosis," *Eur J Cardiothorac Surg.*, vol. 53, no. 2, pp. 306-308, Feb. 2018.

[9] A. Pinto et al., "Sources of error in emergency ultrasonography.," *Crit. Ultrasound J.*, vol. 5 Suppl 1, no. Suppl 1, Art. no. S1, Jul. 2013.

[10] J. M. Sanches, A. F. Laine, and J. S. Suri, "Ultrasound imaging". Springer, 2012.

[11] S. Mazaheri et al., "Echocardiography image segmentation: A survey," *Proc. - 2013 Int. Conf. Adv. Comput. Sci. Appl. Technol. ACSAT 2013*, pp. 327–332, Dec. 2013.

[12] S. Minaee et al, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, Jul 2022.

[13] B. Xie at al. "SePiCo: Semantic-guided pixel contrast for domain Adaptive semantic segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–17, Jan 2023.

[14] Y. Lei et al., "Echocardiographic image multi-structure segmentation using Cardiac-SegNet," *Med. Phys.*, vol. 48, no. 5, pp. 2426–2437, *MON* 2021.

[15] C. Y. Wang, A. Bochkovskiy, and H. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors" *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* pp. 7464-7475, Jun 2023.

[16] C. Sfakianakis, G Simantiris, and G. Tziritas, "GUDU: Geometrically constrained ultrasound data augmentation in U-Net for echocardiography semantic segmentation," *Biomed Signal Process Contr*, vol. 82, Art. no. 104557, Apr 2023.

[17] E. Folland, A. Parisi, P. Moynihan, D. R. Jones, C. L. Feldman, and D. Tow. "Assessment of left ventricular ejection fraction and volumes by real-time, two-dimensional echocardiography. A comparison of cineangiographic and radionuclide techniques," *Circulation*, vol. 60, no. 4, pp. 760-766, Oct 1979.

[18] M. J. Mortada, S. Tomassini, H. Anbar, M. Morettini, L. Burattini, and A. Sbrollini, "Segmentation of anatomical structures of the left heart from echocardiographic images using deep learning," *Diagnostics*, vol. 13, no. 1683, May 2023.

Address for correspondence:

Laura Burattini.
Università Politecnica delle Marche,
Department of Information Engineering,
Via Brecce Bianche, 60131 Ancona, Italy.
E-mail address. l.burattini@univpm.it.