Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tciv20

Automatic cine-based detection of patients at high risk of heart failure with reduced ejection fraction in echocardiograms

Delaram Behnami, Christina Luong, Hooman Vaseli, Hany Girgis, Amir Abdi, Dale Hawley, Ken Gin, Robert Rohling, Purang Abolmaesumi & Teresa Tsang

To cite this article: Delaram Behnami, Christina Luong, Hooman Vaseli, Hany Girgis, Amir Abdi, Dale Hawley, Ken Gin, Robert Rohling, Purang Abolmaesumi & Teresa Tsang (2020) Automatic cine-based detection of patients at high risk of heart failure with reduced ejection fraction in echocardiograms, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 8:5, 502-508, DOI: 10.1080/21681163.2019.1650398

To link to this article: https://doi.org/10.1080/21681163.2019.1650398

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING: IMAGING & VISUALIZATION
2020, VOL. 8, NO. 5, 502–508
https://doi.org/10.1080/21681163.2019.1650398

Taylor & Francis
Taylor & Francis Group

Automatic cine-based detection of patients at high risk of heart failure with reduced ejection fraction in echocardiograms

Delaram Behnami[a], Christina Luong[b,c]*, Hooman Vaseli[a], Hany Girgis[b,c], Amir Abdi[a], Dale Hawley[b], Ken Gin[b,c], Robert Rohling[a], Purang Abolmaesumi[a]† and Teresa Tsang[b,c]†

[a]Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada; [b]Vancouver Coastal Health, Vancouver, Canada; [c]Department of Cardiology, University of British Columbia, Vancouver, Canada

ABSTRACT

Heart failure with reduced ejection fraction (HFrEF) is associated with high mortality rates. The identification of patients with HFrEF in urgent and emergent situations can be delayed in the absence of clinicians skilled in acquiring and interpreting echocardiography (echo) images. Although the standard-of-care for measuring ejection fraction (EF) involves segmentation of the left ventricle in echo, experienced interpreters often opt for visual estimation of EF in clinical settings. In this paper, we mimic this process using dual-channel deep neural networks for segmentation-free classification of echo cine loops: low-risk ($40\% < \text{EF} \leq 75\%$) vs. high-risk (reduced EF, *i.e.* $\text{EF} \leq 40\%$). The proposed architecture utilises densely connected networks for extraction of unsupervised spatial features. Temporal embedding is then achieved by aggregating these frame-level feature vectors and feeding them through recurrent blocks. We use 949 and 237 clinical echo exams to train and validate the proposed method, respectively. We demonstrate that direct cine-based detection of patients at high-risk of HFrEF is feasible using densely connected convolutional and recurrent-based architectures.

ARTICLE HISTORY
Received 21 November 2018
Accepted 23 July 2019

KEYWORDS
Echocardiography; recurrent; convolutional; ejection fraction

1 Introduction

Heart failure (HF) is a complex clinical syndrome, which refers to the diminished cardiac performance for pumping blood in and out of chambers (Savarese and Lund 2017). HF is a epidemic that affects the lives of approximately 26 million people globally (Savarese and Lund 2017). Heart failure with reduced ejection fraction (HFrEF) is associated with the majority of HF-related mortality (Bytyçi and Bajraktari 2015). Given the reciprocal relationship between left ventricular ejection fraction (EF) and six month-mortality in patients with HFrEF, timely diagnosis may have a critical impact on outcomes. EF is a measurement of the percentage of blood that leaves the heart with each contraction. It is the most-commonly measured index for quantitative evaluation of the cardiac function and is routinely measured in 2D echocardiography (echo). Echo is real time, non-ionising, affordable and portable (Cameli et al. 2016), making it an ideal modality for EF assessment.

The standard method for quantification of EF in 2D echo is Simpson's biplane method of discs. For a given patient, this method helps estimate the left ventricle (LV) volume via delineation of the endocardial border at the end-diastolic (ED) and end-systolic (ES) frames, in apical two-chamber (A2C) and apical four-chamber (A4C) views. Many works have focused on automating LV segmentation for measuring EF in echo images. Several machine learning-based solutions have been recently proposed for automatic LV segmentation Carneiro and Nascimento (2013); Dong et al. (2016); Deo et al. (2017); Leclerc et al. (2017); Ngo et al. (2017); Zhang et al. (2017); Jafari et al. (2018). EF estimation is challenging due to its dependence on accurate (1) LV segmentation; (2) detection of the main LV axis; and (3) extraction of ED and ES frames. Even amongst cardiologists with echo training, studies reveal high intra- and inter-user variability, particularly amongst those with less experience (Bresser et al. 2015; Cole et al. 2015).

An alternative and more workflow-efficient technique for EF evaluation is direct visual assessment of echo cine loops (Figure 1). This rapid method of EF estimation, more commonly employed by experienced cardiologists (Kim et al. 2017), has been shown to correlate well with the volumetric segmentation-based Organization WH et al. (2014) approach (Gudmundsson et al. 2005). Nevertheless, direct EF estimation relies heavily on the interpreter experience, and may not be appropriate for novice clinicians who have not undergone sufficient training (Cameli et al. 2016; Kim et al. 2017). Several segmentation-free methods have been proposed for estimation of LV volume and EF in cardiac magnetic resonance (MR) (Gu et al. (2018); Kabani and El-Sakka (2017); Xue et al. (2017); Zhen et al. 2016). These works have shown the feasibility of machine learning techniques, especially deep neural networks for EF analysis. Kabani and El-Sakka (2016) used convolutional neural networks (CNNs) to pre-localise and LV segment for volume estimation and subsequently EF calculation. Xue et al. (2017). similarly relied on a deep CNN to extract features and RNNs to aggregate them for dynamic embedding of the LV motion, etc. Direct cine-based EF estimation in echo is a challenging problem

CONTACT Delaram Behnami delaramb@ece.ubc.ca University of British Columbia, Vancouver V6G3B1, Canada
*Joint first authors
†Joint senior authors

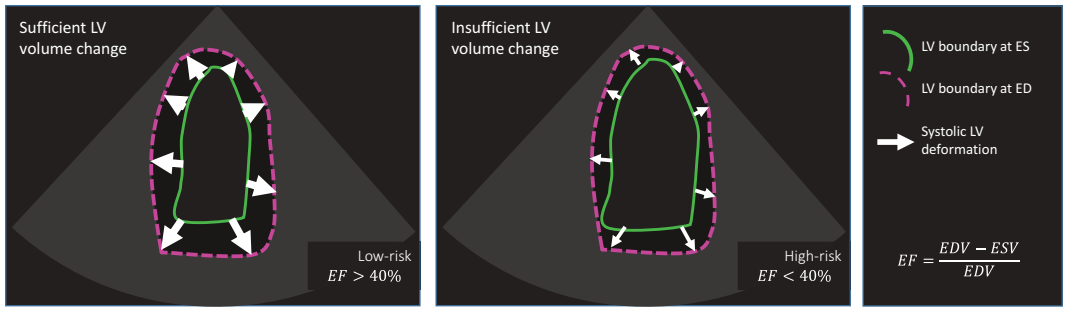


Figure 1. EF can be evaluated by studying the dynamics of the heart throughout the cardiac cycle. Low-risk (left) and high-risk (right) systolic performance are associated with sufficient and insufficient changes in the LV volume, respectively. The solid contours depict the ES volume (ESV) and the dashed contours show the ED volume (EDV).

due to: 1) inherent noisiness of ultrasound, which leads to blurry chamber boundaries; 2) variable quality of 2D cross-sectional echo images due to dependence on the experience level of the echo operators; and 3) the complexity of the cardiac anatomy and function, especially in the apical views.Silva et al. (2018) used residual networks (ResNets) and 3D convolutions to classify EF in A4C cine loops. We recently investigated the feasibility of automatic segmentation-free binary classification of EF using several different architectures for spatio-temporal feature learning of echo cine loops (Behnami et al. 2018). We showed that the proposed dual-channel can be effective in direct EF estimation. This two-step approach involves (1) extraction of frame-level spatial features, and (2) aggregation of these feature maps using recurrent networks, which enable sequential learning. In order to incorporate both A2C and A4C cines in decision making, we synchronise the two cines based on the cardiac phase. In this paper, we build on our previous work with a focus on densely-connected convolutional networks (DenseNet) (Huang et al. 2017) for spatial feature learning and bidirectional gated recurrent units (GRU) (Chung et al. 2014) for temporal encoding. We further provide details for the network architecture, implementation and training. We report the accuracy of the proposed classification framework for various EF ranges and discuss the sources of misclassification. Finally, we conclude the paper and provide some of our future strategies for improving the proposed direct EF estimation framework.

2. Material and methods

2.1. Clinical data

Clinical echo data was accessed through Vancouver Coastal Health (VCH), British Columbia, Canada. Ethics approval was obtained to use this data for research. A total of 1,186 exams and corresponding EF labels were used in this study: 645 low-risk and 541 high-risk cases. EF labels used in this study were recorded by expert echocardiographers using the biplane Simpson's method. A2C and A4C cine loops of the exams were extracted using the echo view classifier proposed by (Van Woudenberg et al. 2018). All exams were acquired using Philips iE33 ultrasound machines and had a frame resolution of 800×600 pixels. Frame rate (FR) and heart rate (HR) vary across patients; FR of2566*Hz* and HR of 60100 beats per minute. 80% of the data was used for training, and the remainder for testing.

2.2. Binary cine-based classification of EF

We formulate the detection of patients at a high risk of HFrEF as a binary classification problem for evaluating $\mathbf{Y}_{High-risk}$, where:

$$\mathbf{Y}_{High-risk} = 1 - \left[\frac{EF}{40}\right]. \tag{1}$$

EF is the ground truth EF label for each exam calculated using the Simpson's method, as recorded in the corresponding clinical report. Here, $\mathbf{Y}_{High-risk} \in \{\{0\},\{1\}\}$. ([.] operator takes the integer part of its input). By definition, $0\% \leq EF \leq 100\%$. However, only samples with $EF \leq 75\%$ were included in this study due to lack of hyperdynamic ($EF > 75\%$) cases in the database.

2.3. Image data preparation

Direct EF estimation relies on dynamic information available in both A2C and A4C cine loops. From a technical standpoint, synchronisation enables reduction in the dimensionality of the data by concatenating in-phase view-specific spatial features of different views, hence, achieving more efficient dual-channel sequential learning. In order to synchronise the two cine loops, cardiac cycle is extracted by detecting the first two consecutive R-peaks in the electrocardiogram (ECG) signal burnt onto the echo images. Each cine loop AxC, $\text{AxC} \in \{\text{A2C}, \text{A4C}\}$, is then trimmed between two R-peaks R_1^{AxC} to R_2^{AxC}. A fixed number of frames $F = 25$ are sampled between R_1^{AxC} to R_2^{AxC} for each cine. This step effectively synchronises the two trimmed cine loops based on their phase relative to the cardiac cycle, even if FR and heart rate (HR) vary.

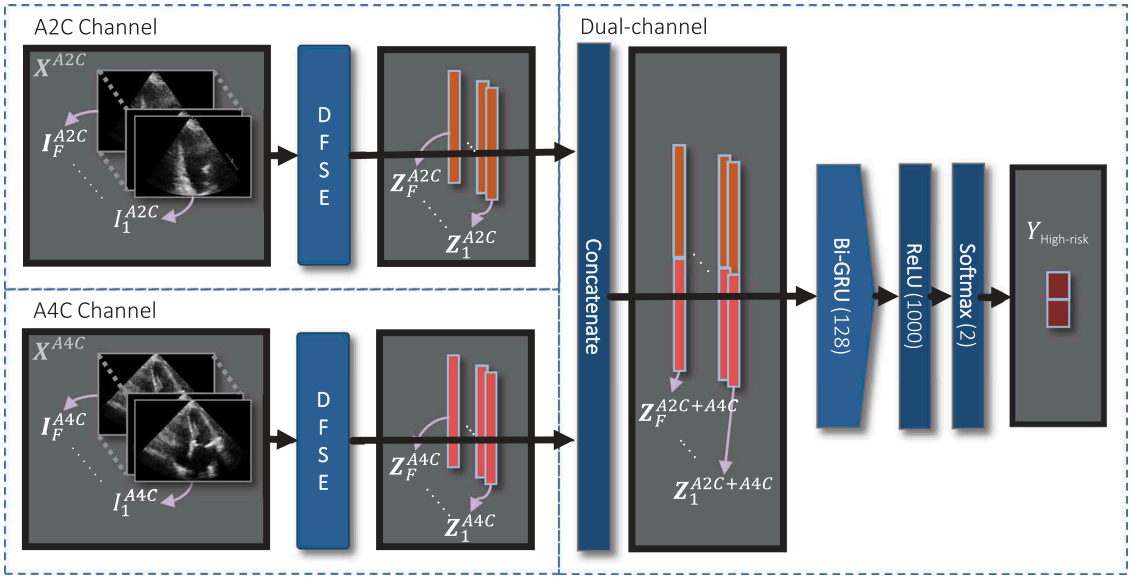
The images are cropped down to the ultrasound beam. A fixed beam-shaped binary mask is applied to the frames to remove image annotations. The frames are then scaled down to 128×128 pixels. If $\mathbf{I}_i$ denotes the i-th processed frame of the 2D greyscale echo cine clip, any cine in AxC view can hence be described as a 3D tensor $\mathbf{X}^{\text{AxC}} = [\mathbf{I}_1^{\text{AxC}} : \mathbf{I}_{CL^{\text{AxC}}}^{\text{AxC}}]$; where CL^{AxC} is the cycle length, and $CL^{\text{AxC}} = R_2^{\text{AxC}} - R_1^{\text{AxC}} + 1$.

2.4. Network architecture

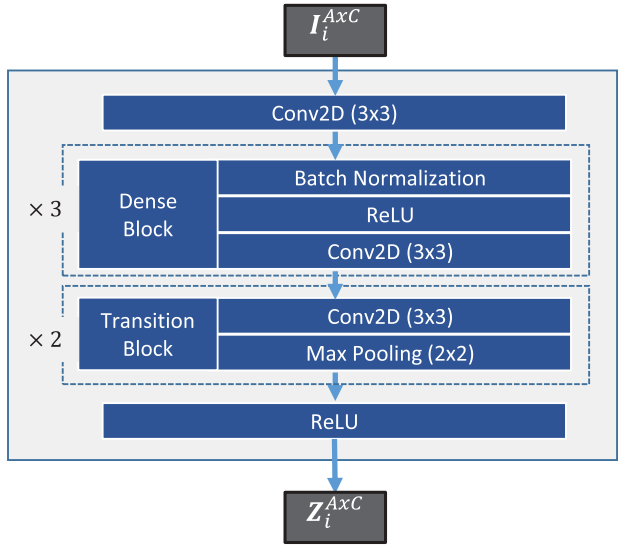
The architecture of the proposed network for binary classification of EF is shown in Figure 2. The main components of this network, a deep spatial feature extraction (DSFE) block and bidirectional GRU (bi-GRU), are discussed below.

2.4.1. Dense spatial feature learning

In order to obtain a compact representation of the content of the frame images $\mathbf{I}_i^{AxC}$, we use a DenseNet-like (Huang et al. 2017) architecture depicted in Figure 2(b). The i-th frame $\mathbf{I}_i^{AxC}$ is fed through a 2D convolutional layer, followed by a cascade of three interconnected dense blocks and two transition blocks. The dense block consists of a batch normalisation layer, a rectified linear unit (ReLU), and a 2D convolutional layer. Batch normalisation reduces the variance of the input, therefore, improving the stability of training. ReLU activation, described as $f(x) = max(0, x)$, introduces and promotes non-linearity in the network. The main characteristic of the DenseNet architecture is that each layer is interconnected with all those that precede it. This provides multiple alternate paths for the gradients to flow back through the network during training, which can lead to more effective learning, especially when dealing with smaller datasets (Huang et al. 2017). Three dense blocks are used, with eight convolutional channels in the first block and a growth rate of 12. In order to contain the network size and parameters, with the network complexity increasing with extra connections, a transition block is designated after the first two dense blocks. The transition block consists of



(a)



(b)

Figure 2. Architecture of the proposed dual-view network for risk-based classification of EF. Frame-level image features are extracted for each channel, then synchronously concatenated and fed through a bi-GRU for temporal embedding (a). The architecture used in the DSFE block consists of 2D convolutional layers, three interconnected dense blocks, two transition blocks and ReLU nonlinearity (b).

a convolutional layer, followed by a pooling layer. A compression factor of 0.5 is used in the transition layers, as suggested by Huang et al. (2017). If $\mathbf{Z}_i^{\text{AxC}}$ denotes the output feature vector of the DSFE block for frame $\mathbf{I}^{\text{AxC}_i}$, the input cine $\mathbf{X}^{\text{AxC}}$ can be compactly described as $[\mathbf{Z}_1^{\text{AxC}} : \mathbf{Z}_{CL^{\text{AxC}}}^{\text{AxC}}]$.

2.4.2. Recurrent block for temporal embedding

The other main component used to analyse the image sequences is a bi-GRU. GRU is a modified and improved version of the standard of recurrent neural networks (RNNs). In the standard RNN, at time step *i*, the *i*-th step of the sequence and the outputs from all previous steps are fed through a dense layer. This effectively allows the system to remember information from past time steps, and hence encodes the sequence. A major drawback of RNNs is that as the sequence length increases, the gradient from earlier time steps start to vanish, leading to loss of information from the earlier phase of the sequence. One way to alleviate this issue is by using Gated Recurrent Units (GRU). In addition to the main architecture of and RNN unit, a GRU also contains an update and a reset gate Chung et al. (2014), which help selectively choose whether features should be kept or discarded. This mechanism enables GRU to remember relevant information from the past time steps. The choice of the bidirectional GRU (bi-GRU) is rooted in the fact that in bidirectional RNNs, the output of each time step is calculated based on both past and future steps. This gives the network more context about the sequence at hand and is hence applicable for EF analysis as our sequential data (echo cine) is periodic. The extracted frame-level features $[\mathbf{Z}_1^{\text{AxC}} : \mathbf{Z}_{CL^{\text{AxC}}}^{\text{AxC}}]$ are concatenated synchronously and fed through a bi-GRU with 128 hidden cells. The spatiotemporal feature maps, *i.e.* the output of bi-GRU, are then passed through a dense layer with 1000 neurons with ReLU nonlinearity. Finally, a dense layer with two neurons and softmax activation, *i.e.* $f(x) = \frac{1}{1+e^{-x}}$, is used to map these dense features to one of the two possible classes, high risk and low risk.

2.4.3. Implementation and training

The proposed framework is implemented in Keras with TensorFlow backend (Chollet et al. 2015). The entire network was trained from scratch with random weight initialisation. The network was trained in an end-to-end fashion on an Nvidia GeForce GTX 980 Ti, with stochastic gradient descent (SGD) parameter optimisation. Stratified batch sampling was used to compensate for the class imbalance. The *TimeDistributed* wrapper was used to apply the DSFE operations to all frames. The weights of the DFSE block are updated separately for A2C and A4C views. However, the parameters of the DFSE block are shared between *F* frames of each view as the image content does not change significantly throughout a cine loop. The start point of the synchronised input cines was shuffled in training within the range of R_1^{AxC} to R_2^{AxC}. We also augment our data on the fly by standard rotation ($\pm 12^\circ$), zoom ($\pm 20\%$), and gamma intensity transformation (0.41.4). For each cine, these transforms are generated randomly within the aforementioned ranges and are applied on the whole cine at hand.

Table 1. Classification accuracy obtained based on A2C, A4C and A2C+A4C cine loops using the proposed DenseNet-based DSFE and bi-GRU.

EF range	$(0, 40]$	$(40, 75]$
$Y_{High-risk}$	1	0
A2C	68.5%	73.8%
A4C	77.5%	81.7%
A2C+A4C	81.1%	84.9%

3. Results and discussion

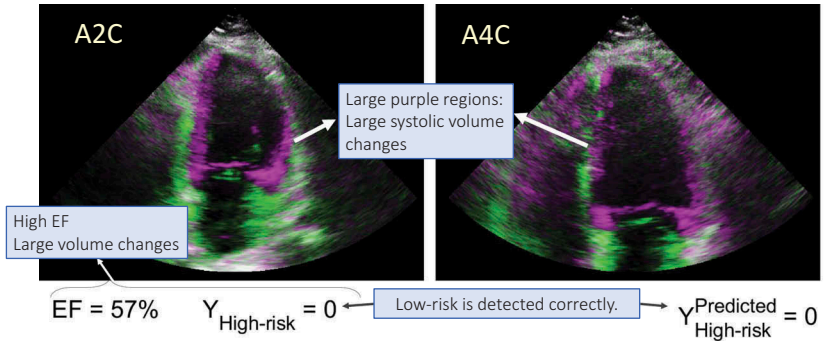
Using the proposed dual-channel method, an overall accuracy of 83.15%, precision of 82.6% and recall of 81.1% were achieved. The test accuracy for binary classification based on A2C, A4C, and synchronised A2C+A4C is provided in Table 1. Using the same experimental set-up, we observed that A4C is a more reliable view for direct evaluation of EF, compared to A2C. This is most likely because: (1) A2C is a generally harder view to capture, which often leads to suboptimal endocardial border definition as compared to A4C; (2) the LV is more likely to be foreshortened in A2C images compared to A4C (Nosir et al. 1997); and (3) A4C might benefit from more context-relevant information from four chambers. We observed that the better performance was obtained when the network sees both A2C and A4C views. Table 2 shows the results broken down to four ranges, each described by $(EF_{min}, EF_{max}]$ intervals. The number of total and test samples are provided as well. The larger number of samples in the low-risk range compared to high-risk may have contributed to the difference in success rates. Our quantitative results suggest misclassification occurs more frequently at the hard threshold of $EF = 40\%$. Figure 3 shows several pairs of A2C and A4C samples. As illustration of videos is challenging, the ES and ED frames (the extreme frames in terms of the appearance of LV) are shown as overlaid image pairs. Green (ED) and purple (ES) areas suggest motion and absolute changes in the position of the imaged anatomy. In the LV region, large purple regions correspond to large LV volume changes in the systolic phase (low-risk EF). As depicted by these images, we observed that such systolic volume changes of the LV were more directly apparent in the A4C view compared to A2C. This observation is consistent with the quantitative results. We also found that the quality of the captured plane, especially A2C, and the accuracy of the ECG-based synchronisation are determining factors for the performance of the proposed model.

4. Conclusion and future work

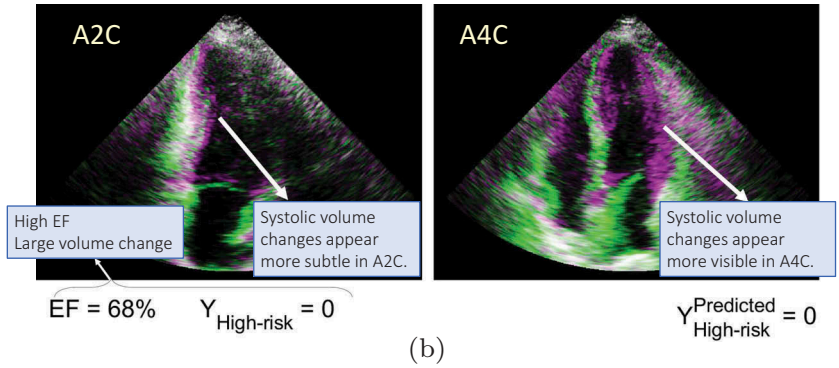
We presented a framework for direct cine-based, segmentation-free prediction of patients at a high risk of HFrEF. The proposed approach enables use of information lying in the dynamics of the heart in two orthogonal and complimentary views. It eliminates the reliance on accurate LV segmentation and detection of ED and

Table 2. A breakdown of the EF classification accuracy based on EF range and number of samples.

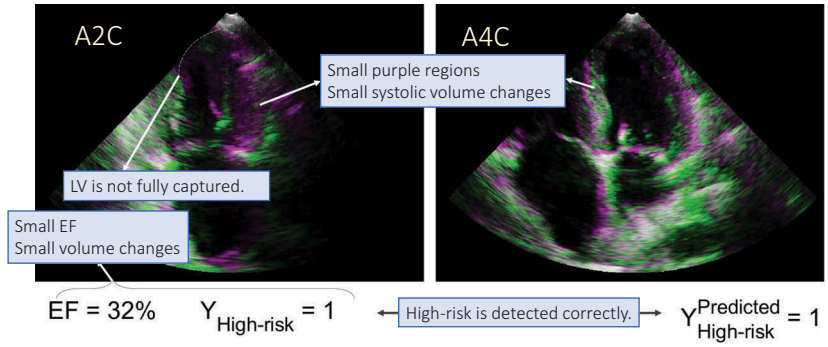
EF range	$(0, 25]$	$(25, 40]$	$(40, 55]$	$(55, 75]$
A2C+A4C	84.9%	78.4%	80.5%	86.5%
Total samples	194	330	408	254
Test samples	40	71	75	51



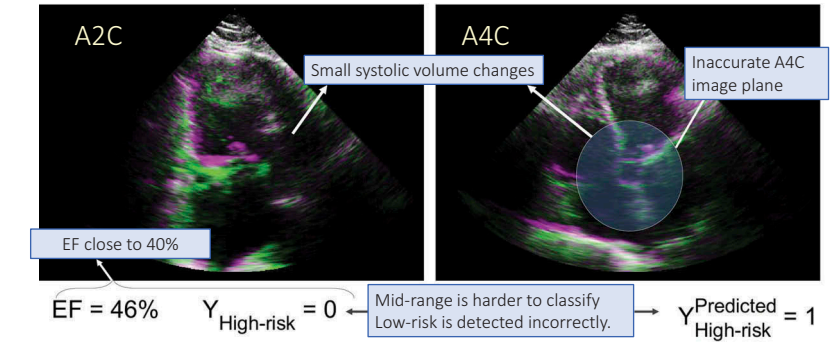
(a)



(b)



(c)



(d)

Figure 3. Example results of dual-channel segmentation-free classification of high-risk vs. low-risk EF using DenseNets and bi-GRU. A2C and A4C views are shown on the left and right of each pair, respectively. The image overlays show the ED (green) and ES (purple) frames of the cycle for better visualisation. $Y^{Predicted}_{High-risk}$ shows the prediction of the network for the given test sample. The large purple regions correspond with a large volume difference between ES and ED frames (ESV and EDV) and low-risk EF (a). In many cases, volume changes appear more subtle in the A2C views (b), which may explain the lower classification accuracy when only A2C is used. High-risk EF corresponds with small LV volume changes, hence smaller purple regions in (c). Quality of the acquired planes and the closeness to the sharp boundary of 40% have impacts on the performance of the proposed solution (d).

ES frames, by learning spatio-temporal features that are associated with EF. Binary classification of left ventricular systolic function is a key step towards direct and automatic estimation of EF. A framework for classification of EF can assist front-line clinicians detect patients at a high-risk of HFrEF, and can allow for expedited referral to the echocardiography lab for comprehensive imaging. While the presented binary accuracy may seem low, it is worth noting that the ground truth *EF* labels used in this study may suffer from some uncertainties and inaccuracies. As mentioned above, studies suggest EF estimation is highly user-dependant. This was evident in our database as well. Analysis of the clinical report database revealed significant discrepancies between the qualitative and quantitative EF reported, with a correlation coefficient of only 0.71 between *EF* labels derived from cine loops (directly estimated values) and those acquired via volumetric analysis (Simpson's method). That being said, we believe the proposed solution could be used to achieve better performance, should more data be used. Given the limited number of samples with consistent *EF* labels, we plan to utilise a semi-supervised learning approach to first train an auto-encoder on a larger unlabelled set of A2C and A4C cine loops, and subsequently train a supervised network on the labelled *EF* data. Furthermore, incorporating more data in training may allow us to move towards a regression-based formulation, eliminating the issue of having a sharp threshold at $EF = 40\%$. Direct estimation of EF is a very complex problem due to the noisiness of echo, complex anatomy and motion, especially in the presence of pathologies, and dependence on image quality.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Canadian Institutes of Health Research [x].

Notes on contributors

Delaram Behnami is a PhD candidate in Electrical and Computer Engineering at UBC.

Dr. Christina Luong is a cardiologist and clinical assistant professor in the Cardiology Division at UBC.

Hooman Vaseli is an undergraduate student in the Department of Electrical and Computer Engineering (ECE) at UBC.

Hany Girgis is a cardiology researcher at Vancouver General Hospital (VGH), and an Associate Professor of Cardiology Fayoum University, Egypt.

Amir Abdi is a PhD candidate and machine learning scientist at ECE, UBC.

Dale Hawley is a Senior Systems Analyst /Designer, Information Management/Information Technology Services (IMITS) for Provincial Health Services Authority (PHSA), Providence Health Care (PHC), and Vancouver Coastal Health (VCH).

Ken Gin is a Clinical Professor, Department of Medicine at UBC, Head of the Division of Cardiology VGH Medical Manager of Cardiology Programs Vancouver Acute, Past Director of Atrial Fibrillation Clinic and CCU VGH, Associate Director Coronary Care Unit and Echo at VGH.

Robert Rohling is a Professor with a joint appointment in Electrical and Computer Engineering & Mechanical Engineering at UBC. Dr. Rohling's research is in the field of biomedical engineering with a specialization in medical ultrasound. Dr. Rohling is also the Director of the Institute for Computing, Information and Cognitive Systems (ICICS).

Dr. Purang Abolmaesumi is a Professor of Medical Imaging, Image Guided Therapy and Applied Machine Learning at the University of British Columbia.

Dr. Teresa Tsang is a cardiologist, Director of Echocardiography at Vancouver General Hospital and UBC Hospital, and Professor of Medicine, Associate Head of Research for the Department of Medicine.

References

Behnami D, Luong C, Vaseli H, Abdi A, Girgis H, Hawley D, Rohling R, Gin K, Abolmaesumi P, Tsang T. 2018. Automatic detection of patients with a high risk of systolic cardiac failure in echocardiography. In: Stoyanov D. et al. editors. Deep learning in medical image analysis and multimodal learning for clinical decision support. DLMIA 2018, ML-CDS 2018. Lecture Notes in Computer Science, vol 11045. Cham: Springer; p. 65–73.

Bresser P, De Beer J, De Wet Y. 2015. A study investigating variability of left ventricular ejection fraction using manual and automatic processing modes in a single setting. Radiography. 21(1):e41–e44.

Bytyçi I, Bajraktari G. 2015. Mortality in heart failure patients. Anatolian J Cardiol. 15(1):63.

Cameli M, Mondillo S, Solari M, Righini FM, Andrei V, Contaldi C, De Marco E, Di Mauro M, Esposito R, Gallina S, et al. 2016. Echocardiographic assessment of left ventricular systolic function: from ejection fraction to torsion. Heart Fail Rev. 21(1):77–94.

Carneiro G, Nascimento JC. 2013. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. IEEE Trans Pattern Anal Mach Intell. 99(1):1.

Chollet F, et al. 2015. Keras. https://keras.io/getting-started/faq/.

Chung J, Gulcehre C, Cho K, Bengio Y. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv Preprint arXiv:14123555.

Cole GD, Dhutia NM, Shun-Shin MJ, Willson K, Harrison J, Raphael CE, Zolgharni M, Mayet J, Francis DP. 2015. Defining the real-world reproducibility of visual grading of left ventricular function and visual estimation of left ventricular ejection fraction: impact of image quality, experience and accreditation. Int J Cardiovasc Imaging. 31(7):1303–1314.

Deo RC, Zhang J, Hallock LA, Gajjala S, Nelson L, Fan E, Aras MA, Jordan C, Fleischmann KE, Melisko M, et al. 2017. An end-to-end computer vision pipeline for automated cardiac function assessment by echocardiography. CoRR. https://ui.adsabs.harvard.edu/abs/2014arXiv1412.3555C/abstract.

Dong S, Luo G, Sun G, Wang K, Zhang H. 2016. A left ventricular segmentation method on 3d echocardiography using deep learning and snake. In: Computing in Cardiology Conference (CinC). IEEE; p. 473–476. https://ieeexplore.ieee.org/abstract/document/7868782/references#references

Gu B, Shan Y, Sheng VS, Zheng Y, Li S. 2018. Sparse regression with output correlation for cardiac ejection fraction estimation. Inf Sci (Ny). 423:303–312.

Gudmundsson P, Rydberg E, Winter R, Willenheimer R. 2005. Visually estimated left ventricular ejection fraction by echocardiography is closely correlated with formal quantitative methods. Int J Cardiol. 101 (2):209–212.

Huang G, Liu Z, van der Maaten L, Weinberger KQ. 2017. Densely connected convolutional networks. In: 2017 IEEE conference on Computer Vision and Pattern Recognition (CVPR). July. p. 2261–2269.

Jafari MH, Girgis H, Liao Z, Behnami D, Abdi A, Vaseli H, Luong C, Rohling R, Gin K, Tsang T, et al. 2018. A unified framework integrating recurrent fully-convolutional networks and optical flow for segmentation of the left ventricle in echocardiography data. In: Stoyanov D. et al. editors. Deep learning in medical image analysis and multimodal learning for clinical decision support. DLMIA 2018, ML-CDS 2018. Lecture Notes in Computer Science, vol 11045. Cham: Springer; p. 29–37.

Kabani A, El-Sakka MR. 2016. Estimating ejection fraction and left ventricle volume using deep convolutional networks. In: Campilho A., Karray F.

editors. Image analysis and recognition. ICIAR 2016. Lecture Notes in Computer Science, vol 9730. Cham: Springer; p. 678–686.

Kabani A, El-Sakka MR. 2017. Ejection fraction estimation using a wide convolutional neural network. In: Karray F., Campilho A., Cheriet F. editors. Image analysis and recognition. ICIAR 2017. Lecture Notes in Computer Science, vol 10317. Cham: Springer; p. 87–96.

Kim C, Hur J, Kang BS, Choi HJ, Shin JH, Kim TH, Chung JH. 2017. Can an offsite expert remotely evaluate the visual estimation of ejection fraction via a social network video call? J Digit Imaging. 30(6):718–725.

Leclerc S, Grenier T, Espinosa F. 2017. A fully automatic and multi-structural segmentation of the left ventricle and the myocardium on highly heterogeneous 2d echocardiographic data. 2017 IEEE International Ultrasonics symposium (IUS), 2017, Washington, DC. IEEE; p. 1–4.

Ngo TA, Lu Z, Carneiro G. 2017. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Med Image Anal. 35:159–171.

Nosir Y, Vletter WB, Boersma E, Frowijn R, Ten Cate FJ, Fioretti PM, Roelandt JR. 1997. The apical long-axis rather than the two-chamber view should be used in combination with the four-chamber view for accurate assessment of left ventricular volumes and function. Eur Heart J. 18(7):1175–1185.

Organization WH, et al. 2014. Global status report on noncommunicable diseases 2014. World Health Organization.

Savarese G, Lund LH. 2017. Global public health burden of heart failure. Cardiac Fail Rev. 3(1):7–11.

Silva JF, Silva JM, Guerra A, Matos S, Costa C. 2018. Ejection fraction classification in transthoracic echocardiography using a deep learning approach. In: 2018 IEEE 31st international symposium on Computer-Based Medical Systems (CBMS), Karlstad. IEEE; p. 123–128.

Van Woudenberg N, Liao Z, Abdi AH, Girgis H, Luong C, Vaseli H, Behnami D, Zhang H, Gin K, Rohling R, et al. 2018. Quantitative echocardiography: real-time quality estimation and view classification implemented on a mobile android device. In: Stoyanov D. et al. editors. Simulation, image processing, and ultrasound systems for assisted diagnosis and navigation. Cham: Springer; p. 74–81.

Xue W, Lum A, Mercado A, Landis M, Warrington J, Li S. 2017. Full quantification of left ventricle via deep multitask learning network respecting intra- and inter-task relatedness. In: Descoteaux M., Maier-Hein L., Franz A., Jannin P., Collins D., Duchesne S., editors. Medical Image Computing and Computer Assisted Intervention (MICCAI). Lecture Notes in Computer Science, vol 10435. Cham: Springer; p. 276–284.

Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, Fan E, Aras MA, Jordan C, Fleischmann KE, et al. 2017. A web-deployed computer vision pipeline for automated determination of cardiac structure and function and detection of disease by two-dimensional echocardiography. arXiv:170607342.

Zhen X, Wang Z, Islam A, Bhaduri M, Chan I, Li S. 2016. Multi-scale deep networks and regression forests for direct bi-ventricular volume estimation. Med Image Anal. 30:120–129.