 doi:10.20944/preprints202203.0140.v1

Article

Convolutional LSTM Approach for Left Ventricle Segmentation and Estimation of Left Ventricle Ejection Fraction in Echocardiography

Jin Gyo Jeong [1*], Dong Hyun Kim [2*], Young Jae Kim [3], Kyungmin Yoo [4], Kyungeun Ha [4,5], Wook-Jin Chung [4,5†] and Kwang Gi Kim [1,3†]

[1] Department of Health Sciences and technology, GAIHST, Gachon University, Incheon 21999, Korea; wlsry4008@gamcil.com
[2] Department of Medicine, Gachon University College of Medicine, Incheon, 21565, Korea; e-mail@e-mail.com
[3] Department of Biomedical Engineering, College of Medicine, Gachon University, Incheon, 21565, Korea
[4] Gachon Cardiovascular Research Institute, Gachon University, Incheon, Korea
[5] Department of Cardiovascular Medicine, Gachon University Gil Medical Center, Incheon, Korea
† Correspondence: kimkg@gachon.ac.kr and heart@gachon.ac.kr
*These authors contributed equally to this work as first authors.
K.G.K† and W-J.C† contributed equally to this study as corresponding authors.

Abstract: Cardiovascular disease is the leading cause of death worldwide. A key factor in assessing the risk of cardiovascular disease is left ventricular functional evaluation. Left ventricular (LV) systolic function is evaluated by measuring the left ventricular ejection fraction (LVEF) using echocardiography data. Therefore, quick and accurate left ventricle segmentation is important for estimating the LVEF. However, it is difficult to accurately segment the left ventricle due to changes in the shape and area of the left ventricle during cardiac cycles. In this study, we proposed a framework that considers changes in the shape and area of the left ventricle during the cardiac cycle by applying the convolutional long short-term memory (CLSTM) approach. In addition, we evaluated the left ventricular segmentation and multidimensional quantification of the proposed system in comparison to manual and automated segmentation methods. In addition, to assess the validity of CLSTM, the values of multi-dimensional quantification metrics were compared and analyzed using graphs and Bland–Altman plots on a frame-by-frame basis. We demonstrated that the CLSTM method effectively segments the left ventricle by considering the LV activity. In conclusion, we demonstrated that LV segmentation based on our framework may be utilized to accurately estimate LVEF values.

Keywords: Left ventricular ejection fraction; Left ventricle segmentation; Convolutional long short-term memory; Echocardiography

1. Introduction

Cardiovascular disease is the highest death-causing disease in the world[1]. A key factor in assessing the risk of cardiovascular disease is left ventricular functional evaluation[2]. The primary test for evaluating the function of the left ventricle (LV) is echocardiography[3]. LV systolic function is evaluated by measuring the left ventricular ejection fraction (LVEF) using echocardiography[4]; LVEF is an important indicator for the diagnosis and prediction of the prognosis of cardiovascular disease[5]. Heart failure and myocardial disease are likely to occur in cases of abnormal LVEF values[6,7]. LVEF represents the percentage of blood released to the body when the heart contracts; it is calculated by dividing the one-time ejection of the heart by the end-diastolic volume of the LV[8].

Currently, in clinical practice, the modified Simpson's method (MSM) is recommended to evaluate LVEF in echocardiographic images[9]. MSM calculates LVEF by

 doi:10.20944/preprints202203.0140.v1

measuring the volume via segmentation of the end-systolic and end-diastolic left ventricular endocardium using A2C and A4C views[10]. In current clinical trials, the LV endocardium is manually segmented using the MSM[11].

However, manual segmentation is time-consuming, requires specialized clinical knowledge about the cardiac, and has the possibility of providing different outcomes among examiners[12]. Furthermore, limitations such as the ambiguity of myocardial distinctions caused by the presence of papillary and trabecular muscles, the variable shape of the heart structure according to different patients, and the change in the LV size due to the heartbeat are also problematic when attempting to achieve precise segmentation[13]. Therefore, it is difficult to quantitatively measure LVEF due to differences in the left ventricular volume.

Traditional image processing and machine learning approach manually extract features from data; however, inaccurate results can be obtained because they are performed using features extracted based on the specifications of the algorithm creator. In contrast, the convolutional neural network (CNN) approach is capable of automatically learning relevant features via convolution operations to achieve improved results.

With the development of deep learning techniques, attempts to perform LV segmentation in echocardiogram images have utilized automatic segmentation techniques that have been developed for medical imaging applications[14]. For example, MFP-Unet has been used to conduct LV segmentation with a multi-feature pyramid U-net[15]. The combined robustness of the existing U-net system and the flexibility of featured pyramid networks (FPNs), achieved by mixing a dilated U-net and FPN, demonstrated improved performance over traditional and machine learning approaches, with a lower runtime than found in previous studies.

ResDUnet is another system that could achieve LV segmentation in echocardiogram images[16]. Using residual blocks eliminates the need for a complex deep architecture and allows detailed features to be extracted from the segmentation process; however, the performance of this technique declines if the boundary of the ventricular wall is ambiguous.

Previous studies have also encountered limitations when utilizing frame-by-frame approaches (i.e., 2D-based segmentation of the left ventricle) without considering the inter-frame information in the echocardiogram data and the change in the LV size.

The K-Net method achieves left ventricular segmentation considering the learning of inter-frames in echocardiogram images using Bi-ResLSTM[17]. This model learns temporal variations between the previous and current frames, showing superior performance over other existing models. This method confirmed that learning is possible considering the changes in the LV size and the variations according to the temporal order in echocardiogram videos.

In this study, we propose a model that considers the LV activity and variation according to the temporal order of echocardiogram video data using a convolutional long short-term memory (CLSTM) approach. Furthermore, we evaluate the clinical usefulness of LV segmentation and LVEF estimation using this method. The detailed abbreviations and definitions used in this study are listed in Table 1.

Table 1. List of abbreviation and acronyms used in this study

Abbreviation	Definition
A2C	Apical two-chamber
A4C	Apical four-chamber
LV	Left Ventricle/Ventricular
LVAd	Left Ventricular Area at End-Diastolic
LVAs	Left Ventricular Area at End-Systolic
LVEF	Left Ventricular Ejection Fraction
LAD	Long-Axis Dimension
SADd	Short-Axis Dimension at End-Diastolic

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1

SADs	Short-Axis Dimension at End-Systolic
LVVd	Left Ventricular Volume at End-Diastolic
LVVs	Left Ventricular Volume at End-Systolic

2. Materials and Methods

2.1 Patient Demographics

In this study, echocardiogram video data were obtained from 87 patients (mean age: 41 ± 13 years; 55% male; mean weight: 63.3 ± 12 kg) from 2016 to 2018 at Gil Medical Center; the patient characteristics are listed in Table 2. Individual patient data are composed of an apical two-chamber view and an apical four-chamber view. This study was approved by the Institutional Review Board of the Gil Medical Center (IRB No. GCIRB2021-250). All images were de-identified before inclusion in this study.

Table 2. Patient characteristics

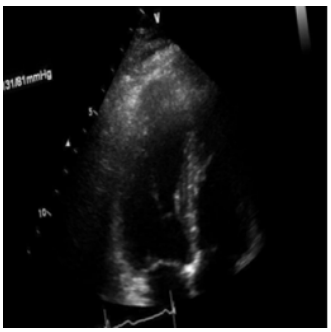
	Female	**Male**	**All**
No. of cases	39	48	87
Age (years)	40.8 ± 12.8	41.9 ± 13.8	41.8 ± 13.8
Body height (cm)	167.6 ± 8.4	167.5 ± 8.3	167.3 ± 8.4
Body weight (kg)	63.3 ± 12.8	63.3 ± 12.6	63.3 ± 12.6

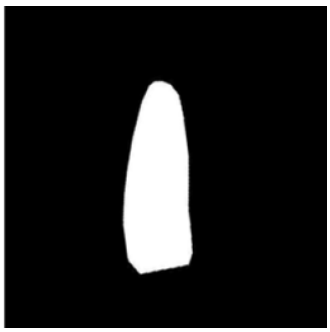
2.2 Datasets and Data Pre-Processing

All echocardiogram video data used in this study had a resolution of 636 × 436 with 40–48 frames. Data with fewer than 48 frames were equally composed of 48 frames using zero paddings. In addition, all data was resized to 256 × 256 using the ground truth (GT). Overall, 60% of the images were used for training, 20% for validation, and 20% for testing.

2.3 Data Augmentation

To avoid overfitting and improve the generalization of the model, the dataset was augmented to 243 videos with three types: (1) only random rotations (-15° to 15°); (2) random scaling (0.5 to 1.5 times); and (3) a combination of random rotations (-15° to 15°) and scaling (0.5 to 1.5 times). Figure 1 shows the results of the data augmentation process.





Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1

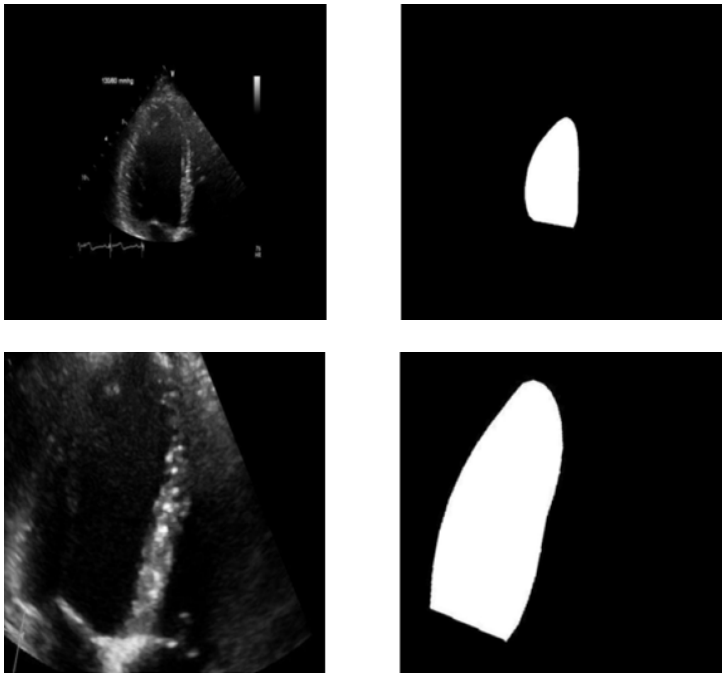


Figure 1. Augmentation images: (top) rotation, (middle) scaling, (bottom) rotation and scaling.

2.4 Bi-LSTM U-Net for LV Segmentation

The proposed framework is inspired by U-Net, which exhibits a high performance in the segmentation of medical images, and bidirectional-convolutional LSTM (Bi-CLSTM), which can learn using data with a temporal sequence[18]. U-Net consists of an encoder, decoder, and skip connection between the encoder and decoder. The encoder extracts the features of the LV via convolution calculations; the decoder restores the feature map, which becomes smaller due to convolutional calculations into the original size of the image. Figure 2 illustrates the architecture of the proposed framework.

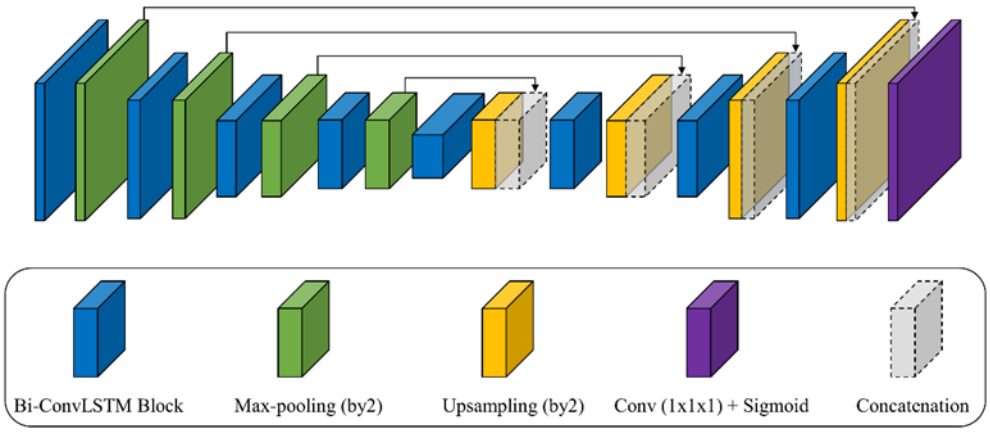


Figure 2. Proposed framework architecture.

LSTM is typically used in natural language processing, and it is primarily used for continuous data or data with a temporal sequence; CLSTM was developed for CNN models[19-21]. A previous study proposed a 3D segmentation approach that combined CLSTM for 3D fungus datasets. This approach achieved a higher performance than the generalized 3D CNN. The CLSTM approach demonstrated efficient learning abilities considering inter-slice and 3D volume data. LSTM considers a data sequence in one direction;

 doi:10.20944/preprints202203.0140.v1

however, as the heart undergoes repetitive movements wherein the area of the left ventricle expands and contracts due to contraction and relaxation, Bi-CLSTM was used in this study to consider both directions. Figure 3 shows the Bi-CLSTM block.

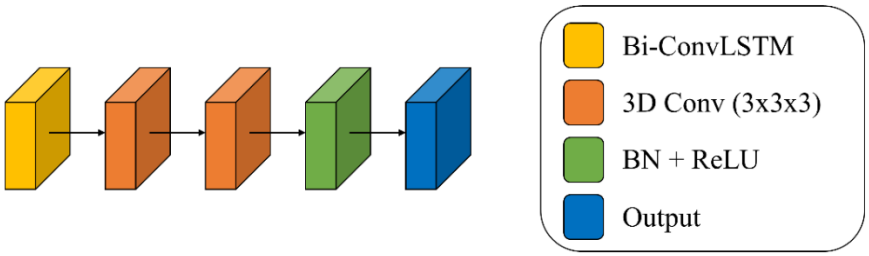


Figure 3. Bi-CLSTM block.

2.5 Implementation Details

The proposed framework was implemented in Python 3.7, TensorFlow 2.1.3, Keras 2.3.1, and was run on four NVIDIA Tesla V100 GPUs with 5120 cores and 32 GB of memory. Our framework was trained using the Adam optimizer to minimize the dice loss[22,23]. When the loss was not minimized over 10 epochs, the learning rate was reduced by multiplying the initial learning rate by 0.1, with an initial value of 0.01. The process was terminated early when the loss did not improve for 30 epochs during the training procedure. Our network terminated the training procedure early between 90 to 150 epochs.

2.6 Left ventricular ejection fraction

LVEF is a parameter with the highest clinical significance for cardiac function evaluation. LVEF is estimated using the volume displacement between the LVVd and LVVs. Therefore, accurate LVVd and LVVs information is required to estimate LVEF using echocardiography. Several methods are used to estimate the volume of the left ventricle, but the modified Simpson's method is recommended. By dividing the left ventricle into 20 disks, the left ventricular volume was calculated using Eq. 1 via modified Simpson's method. Figure 4 shows the modified Simpson's method; LVEF values were estimated using Eq. 2.

$$LV\ Volume = \frac{\pi}{4}\sum_{i=1}^{20} aibi\ \times \frac{L}{20} \tag{1}$$

$$LVEF = \frac{LVVd - LVVs}{LVVd} \times 100\ (\%) \tag{2}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1

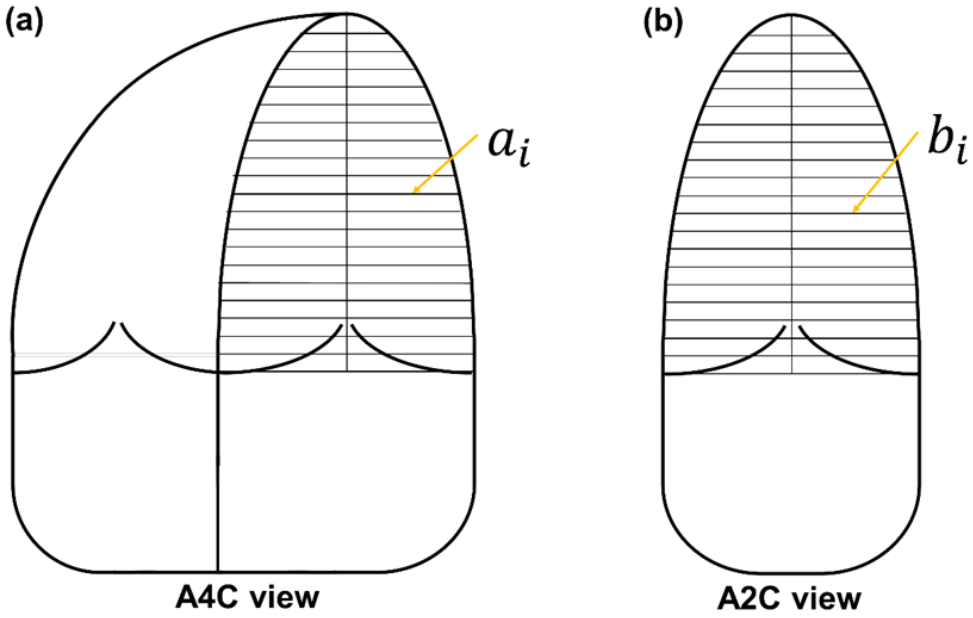


Figure 4. Illustration of the modified Simpson's method: (a) 20 disks in an apical four-chamber view and (b) 20 disks in an apical two-chamber view.

2.7 Evaluation Metrics

The performance of the proposed approach was evaluated using the following metrics: recall, precision, dice similarity coefficient (DSC), and mean absolute error (MAE). The DSC was defined as the volume of overlap between the CNN and manually labeled segmentations divided by the average segmentation volume of the two methods:

$$DSC = \frac{1}{n_t}\sum_{t}\frac{2|y^t \cap \hat{y}^t|}{|y^t| + |\hat{y}^t|} \tag{3}$$

The MAE was defined as:

$$MAE = \frac{1}{n_t}\sum_{t}|y^t - \hat{y}^t| \tag{4}$$

to evaluate the estimation precision of the 1D (i.e., LADA2C, LADA4C, SADdA2C, SADsA2C, SADdA4C, and SADsA4C), 2D (i.e., LVAdA2C, LVAsA2C, LVAdA4C, and LVAsA4C), and 3D (LVVd and LVVs) indices.

3. Results

3.1. LV Segmentation Accuracy

We analyzed the accuracy of the LV segmentation framework by comparing the 3D U-Net, 3D U-Net w/CLSTM, and our proposed framework. Figure 5 shows the visual assessment results for the continuous frame of the overlapping image in which manual segmentation by a cardiologist and the predicted LV segmentation results achieved by 3D U-Net, 3D U-net w/CLSTM, and our framework were performed. Table 3 shows the performance of our method for the A2C and A4C datasets. Our framework achieved a high segmentation accuracy of 0.912±0.109 and 0.888±0.127 for the paired apical views. It performed better than 3D U-Net in terms of the segmentation accuracy, exhibiting improvements of 1.6% and 5.2% in the A2C and A4C datasets, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1

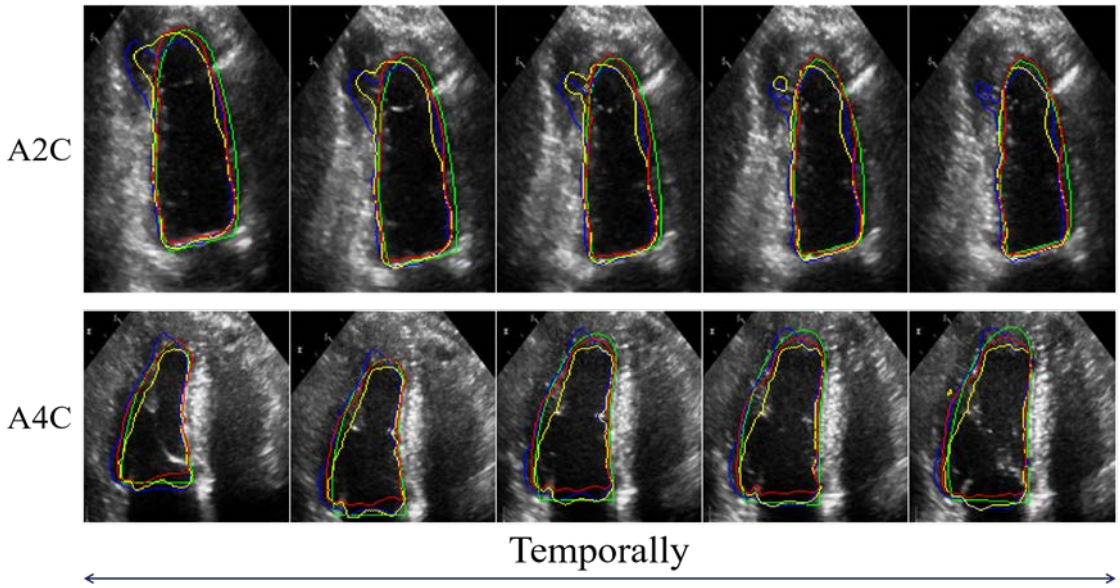


Figure 5. Consistently accurate LV segmentation along consecutive frames via the proposed framework. The ground truth and predicted segmentations of the 3D U-Net, 3D U-Net w/ CLSTM, and 3D U-Net w/ Bi-CLSTM are shown by the green, blue, yellow, and red lines, respectively.

Table 3. Segmentation performance of the proposed model on the A2C and A4C datasets considering performance evaluation metrics.

Data	Model	Recall	Precision	DSC
A2C	3D U-Net	0.940±0.064	0.878±0.147	0.896±0.129
	3D U-Net w/ CLSTM	0.927±0.072	0.897±0.142	0.901±0.125
	3D U-Net w/ Bi- CLSTM	0.951±0.056	0.893±0.133	0.912±0.109
A4C	3D U-Net	0.940±0.066	0.812±0.243	0.836±0.235
	3D U-Net w/ CLSTM	0.927±0.080	0.843±0.211	0.856±0.190
	3D U-Net w/ Bi- CLSTM	0.913±0.113	0.892±0.141	0.888±0.127

3.2 Estimated Quantification Accuracy

A comparison of the estimated LVEF and the 1D, 2D, and 3D quantification results is presented in Table 4. Considering quantification, our framework achieved a high precision for multidimensional indices with low MAEs of 0.747, 0.813, 0.303, 0299, 0.311, and 0.365 mm for 1D quantification (LADA2C, LADA4C, SADdA2C, SADsA2C, SADdA4C, and SADsA4C, respectively); 3.082, 2.662, 3.630, and 3.172 mm2 for 2D quantification (LVAdA2C, LVAsA2C, LVAdA4C, and LVAsA4C, respectively); and 13.287, 8.550, and 5.486 ml for 3D quantification (LVVd, LVVs, and LVEF, respectively). Therefore, our framework effectively segmented the changes in the LV size and provided accurate quantification results from the continuous frames.

Table 4. Quantification performance of the proposed framework under different component configurations. MAE is used to evaluate the performances of all frameworks.

	3D U-Net	**3D U-Net w/ CLSTM**	**3D U-Net w/ Bi-CLSTM**
One-dimensional metrics (mm)			
LAD_{A2C}	1.175±1.124	0.906±1.067	**0.747±0.853**
LAD_{A4C}	0.898±0.015	0.910±0.094	**0.813±0.749**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1

$SADd_{A2C}$	0.455±0.481	0.396±.0432	**0.303±0.256**
$SADs_{A2C}$	0.376±0.387	0.339±0.389	**0.299±0.265**
$SADd_{A4C}$	0.513±0.524	0.405±0.456	**0.311±0.216**
$SADs_{A4C}$	0.431±0.438	0.392±0.374	**0.365±.278**
Two-dimensional metrics (mm^2)			
$LVAd_{A2C}$	5.602±6.440	4.828±6.514	**3.082±2.933**
$LVAs_{A2C}$	3.842±3.971	3.296±4.033	**2.662±2.223**
$LVAd_{A4C}$	5.862±7.831	5.577±7.075	**3.630±2.781**
$LVAs_{A4C}$	3.869±5.043	3.685±4.399	**3.172±2.534**
Three-dimensional metrics (mL)			
LVVd	23.305±36.628	20.247±29.643	**13.287±10.653**
LVVs	10.495±14.684	9.902±12.978	**8.550±7.718**
LVEF	6.833±6.317	6.056±4.537	**5.486±4.638**

The LVEF estimated by the proposed framework and the LVEF measured by a cardiologist were compared and analyzed using a Blend–Altman plot. Figure 6 shows the Bland–Altman plot, wherein most of the LVEF values were within the 95% confidence interval, indicating equivalence.

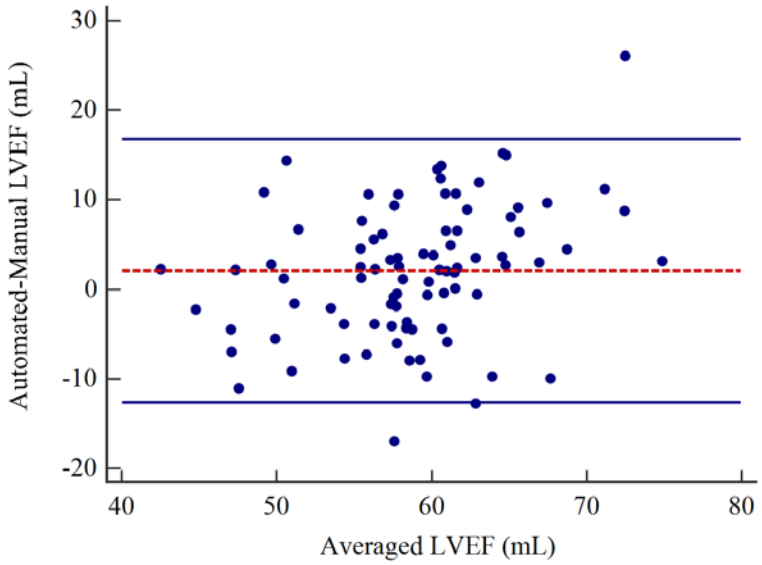
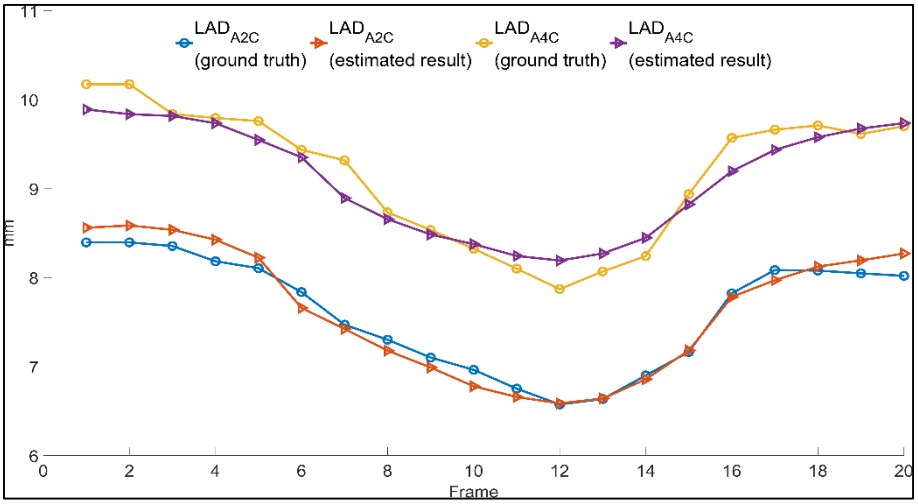


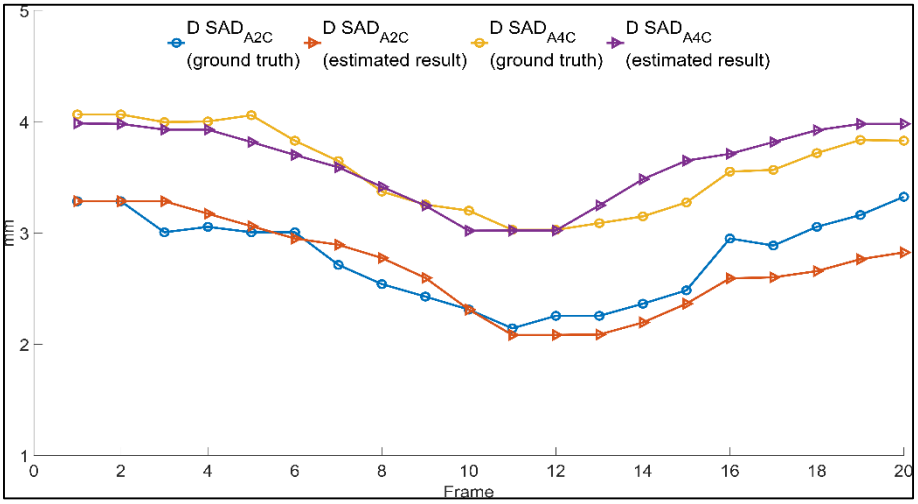
Figure 6. Agreement of LVEF values between the CNN-estimated and manually measured results.

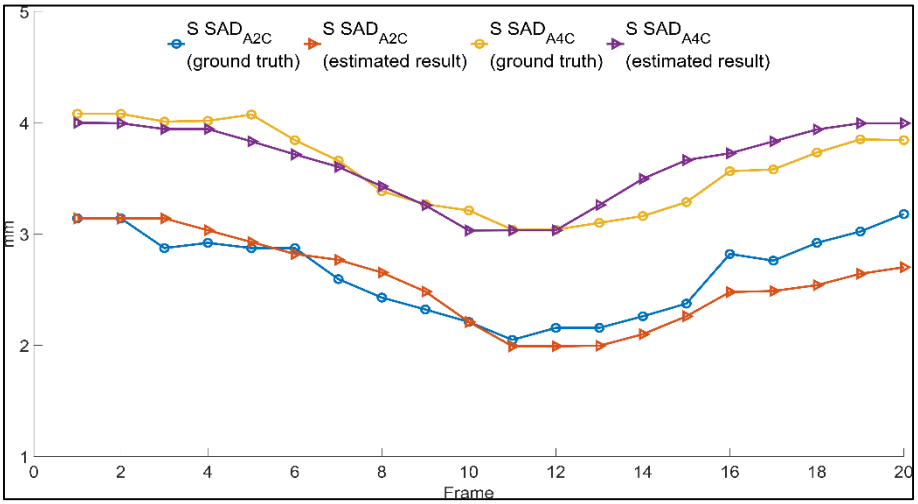
3.3 Effectiveness of Bi-CLSTM in Inter-Frames

The 1D, 2D, and 3D quantification results are presented frame-by-frame in Figure 7 to verify Bi-CLSTM is effectiveness for LV segmentation in inter-frames.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1







Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 March 2022 doi:10.20944/preprints202203.0140.v1

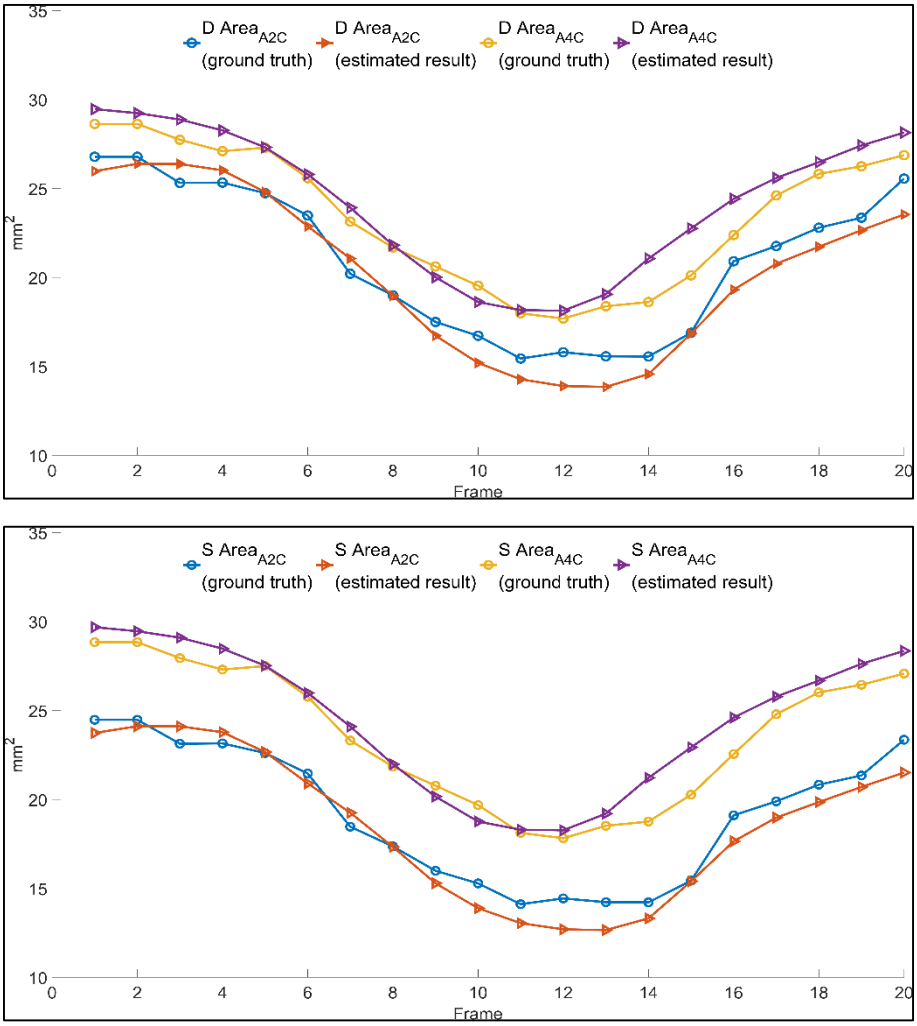


Figure 7. Estimation of high coincide indices along the cardiac cycle by the proposed framework to model LV activity. The polygonal lines present the frame-wise value of the index for the average subject.

4. Discussion

In this study, we showed that our proposed framework based on U-Net and Bi-CLSTM can be trained to perform automated and accurate LV segmentation, and LVEF estimation using echocardiography. This model achieved high performance with segmentation accuracy DSC values of 0.912 and 0.888 for A2C and A4C datasets, respectively. Compared to the reference model (3D U-Net), the performance of the proposed framework was improved by 1.6% and 5.2% for the A2C and A4C datasets, respectively. Table 3 presents the results of the CLSTM and Bi-CLSTM models compared to the basic 3D U-Net model, showing that the segmentation accuracy was improved by 2% and 3%, respectively. Figure 5 shows the training results of the proposed model. The segmentation results of the proposed model were more accurate than those of the other models.

In addition, we performed a comparative analysis of the LVEF estimated based on the LV area manually segmented by the cardiologists and the LVEF estimated by our framework. A Bland–Altman plot analysis of the LVEF results indicated the reliability of our framework considering the agreement of our results with those estimated by cardiologists. Therefore, the proposed model for LV segmentation using echocardiography may be used to accurately estimate LVEF values. As a result of the Bland–Altman plot analysis,

most of the values were within the 95% confidence interval, indicating the reliability between the compression ratio estimated by the cardiologists and the compression ratio estimated by our framework is high. Table 4 lists the MAE values for multidimensional quantification parameters. The proposed framework showed lower MAE values than the basic 3D U-Net model, indicating that the proposed model accurately models the LV activity, effective for comprehensive LV segmentation and quantification from echocardiography data. Furthermore, we presented the 1D and 2D quantification results that were estimated frame-by-frame to verify the effectiveness of Bi-CLSTM, as shown in Fig. 7. Consequently, the similarity of the estimated results to the ground truth was verified, and Bi-CLSTM was found to be effective for use in conjunction with continuous echocardiographic video data.

There are several limitations to the dataset used in this study. The echocardiographic video data used in this study were obtained from two or more models from different manufacturers; the data was difficult to normalize because of various differences such as contrast and resolution; therefore, additional work may be required to improve the applicability of this data. More extensive testing should be conducted to reduce the number of failed CNN cases, and the additional training data may be utilized to further improve the performance of the proposed model.

In conclusion, the proposed framework segments LV echocardiography data to provide high performance. The LVEF values estimated using the images segmented by the proposed framework resulted indicated the high reliability of this model. Additional training with a larger volume of data should be conducted in future studies to improve the accuracy of the segmentation results of this model over the LV area and improve the estimation of LVEF values.

Author Contributions: J.G.J. and D.H.K. drafted the manuscript. J.G.J. conducted the deep learning analysis and performed statistical analysis. K.Y., K.H., and W.J.C. collected the dataset and confirmed the comparison process. W.J.C., Y.J.K., and K.G.K. revised the manuscript. All authors have read and approved the final manuscript.

Funding: This research was supported by the Seoul R&D Program (BT190153) through the Research and Development for Regional Industry, the Korea Medical Device Development Fund grant funded by the Korean government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 9991006834, KMDF_PR_20200901_0164), (Project Number: 9991007387,KMDF_PR_20200901_0170), Gachon University research fund (GCU-202008440010) by the MSIT (Ministry of Science and ICT), Korea under the ITRC (Information Technology Research Center) support program (IITP-2021–2017-0–01630) supervised by the IITP (Institute for Information & Communications Technology Promotion), and Gil Medical Center (FRD2019-11-02(3)).

Institutional Review Board Statement: The study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Gil Medical Center, Gachon University, South Korea (IRB No. GCIRB2021-250).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The echocardiography data used to support the findings of this study are available upon request from the corresponding author.

Acknowledgments: Kwang Gi Kim† and Wook-Jin Chung† contributed equally to this study as corresponding authors. Jin Gyo Jeong* and Dong Hyun Kim* contributed equally to this work as first authors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mann, D.L.; Zipes, D.P.; Libby, P.; Bonow, R.O.; Braunwald, E. *Braunwald's heart disease : a textbook of cardiovascular medicine*, Tenth edition. ed.; Elsevier/Saunders: Philadelphia, PA, 2015; p. volumes.

2. Jafari, M.; Girgis, H.; Liao, Z.; Behnami, D.; Abdi, A.H.; Vaseli, H.; Luong, C.L.; Rohling, R.; Gin, K.; Tsang, T.S.M.; et al. A Unified Framework Integrating Recurrent Fully-Convolutional Networks and Optical Flow for Segmentation of the Left Ventricle in Echocardiography Data. In Proceedings of the DLMIA/ML-CDS@MICCAI, 2018.
3. Dong, S.; Luo, G.; Tam, C.; Wang, W.; Wang, K.; Cao, S.; Chen, B.; Zhang, H.; Li, S. Deep Atlas Network for Efficient 3D Left Ventricle Segmentation on Echocardiography. *Med Image Anal* **2020**, *61*, 101638, doi:10.1016/j.media.2020.101638.
4. Ge, R.; Yang, G.; Chen, Y.; Luo, L.; Feng, C.; Zhang, H.; Li, S. PV-LVNet: Direct left ventricle multitype indices estimation from 2D echocardiograms of paired apical views with deep neural networks. *Med Image Anal* **2019**, *58*, 101554, doi:10.1016/j.media.2019.101554.
5. Li, T.; Wei, B.; Cong, J.; Hong, Y.; Li, S. Direct estimation of left ventricular ejection fraction via a cardiac cycle feature learning architecture. *Comput Biol Med* **2020**, *118*, 103659, doi:10.1016/j.compbiomed.2020.103659.
6. Tanabe, K.; Sakamoto, T. Heart failure with recovered ejection fraction. *J Echocardiogr* **2019**, *17*, 5-9, doi:10.1007/s12574-018-0396-2.
7. Jurado-Roman, A.; Agudo-Quilez, P.; Rubio-Alonso, B.; Molina, J.; Diaz, B.; Garcia-Tejada, J.; Martin, R.; Tello, R. Superiority of wall motion score index over left ventricle ejection fraction in predicting cardiovascular events after an acute myocardial infarction. *Eur Heart J Acute Cardiovasc Care* **2019**, *8*, 78-85, doi:10.1177/2048872616674464.
8. MacIver, D.H.; Adeniran, I.; Zhang, H. Left ventricular ejection fraction is determined by both global myocardial strain and wall thickness. *Int J Cardiol Heart Vasc* **2015**, *7*, 113-118, doi:10.1016/j.ijcha.2015.03.007.
9. Lang, R.M.; Badano, L.P.; Mor-Avi, V.; Afilalo, J.; Armstrong, A.; Ernande, L.; Flachskampf, F.A.; Foster, E.; Goldstein, S.A.; Kuznetsova, T.; 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. *Eur Heart J Cardiovasc Imaging* **2015**, *16*, 233-270, doi:10.1093/ehjci/jev014.
10. Kosaraju, A.; Goyal, A.; Grigorova, Y.; Makaryus, A.N. Left Ventricular Ejection Fraction. In *StatPearls*; StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.: Treasure Island (FL), 2022.
11. Franchi, F.; Cameli, M.; Taccone, F.S.; Mazzetti, L.; Bigio, E.; Contorni, M.; Mondillo, S.; Scolletta, S. Assessment of left ventricular ejection fraction in critically ill patients at the time of speckle tracking echocardiography: intensivists in training for echocardiography versus experienced operators. *Minerva Anestesiol* **2018**, *84*, 1270-1278, doi:10.23736/S0375-9393.18.12249-8.
12. Azarmehr, N.; Ye, X.; Sacchi, S.; Howard, J.P.; Francis, D.P.; Zolgharni, M. Segmentation of Left Ventricle in 2D Echocardiography Using Deep Learning. Cham, 2020; pp. 497-504.
13. Leclerc, S.; Smistad, E.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; Espeland, T.; Rye Berg, E.A.; Belhamissi, M.; Israilov, S.; Grenier, T.; et al. LU-Net: A Multistage Attention Network to Improve the Robustness of Segmentation of Left Ventricular Structures in 2-D Echocardiography. *IEEE Trans Ultrason Ferroelectr Freq Control* **2020**, *67*, 2519-2530, doi:10.1109/TUFFC.2020.3003403.
14. Leclerc, S.; Jodoin, P.-M.; Løvstakken, L.; Bernard, O.; Smistad, E.; Grenier, T.; Lartizien, C.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; et al. RU-Net: A refining segmentation network for 2D echocardiography. **2019**, 1160-1163.
15. Moradi, S.; Oghli, M.G.; Alizadehasl, A.; Shiri, I.; Oveisi, N.; Oveisi, M.; Maleki, M.; Dhooge, J. MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography. *Phys Med* **2019**, *67*, 58-69, doi:10.1016/j.ejmp.2019.10.001.
16. Amer, A.; Ye, X.; Zolgharni, M.; Janan, F. ResDUnet: Residual Dilated UNet for Left Ventricle Segmentation from Echocardiographic Images. *Annu Int Conf IEEE Eng Med Biol Soc* **2020**, *2020*, 2019-2022, doi:10.1109/EMBC44109.2020.9175436.

17. Ge, R.; Yang, G.; Chen, Y.; Luo, L.; Feng, C.; Ma, H.; Ren, J.; Li, S. K-Net: Integrate Left Ventricle Segmentation and Direct Quantification of Paired Echo Sequence. *IEEE Trans Med Imaging* **2020**, *39*, 1690-1702, doi:10.1109/TMI.2019.2955436.
18. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Cham, 2015; pp. 234-241.
19. Patraucean, V.; Handa, A.; Cipolla, R.J.A. Spatio-temporal video autoencoder with differentiable memory. **2015**, *abs/1511.06309*.
20. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.-Y.; Wong, W.-k.; Woo, W.-c. Convolutional LSTM Network: a machine learning approach for precipitation nowcasting. In Proceedings of the Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1, Montreal, Canada, 2015; pp. 802–810.
21. Stollenga, M.F.; Byeon, W.; Liwicki, M.; Schmidhuber, J. Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. In Proceedings of the Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2, Montreal, Canada, 2015; pp. 2998–3006.
22. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* **2014**.
23. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Jorge Cardoso, M. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2017)* **2017**, *2017*, 240-248, doi:10.1007/978-3-319-67558-9_28.