2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)

Identification of the left ventricle endocardial border on two-dimensional ultrasound images using the convolutional neural network Unet

Vasily Zyuzin[1], Porshnev Sergey[1], Andrey Mukhtarov[1], Tatyana Chumarnaya[1,2], Olga Solovyova[1,2]
Anastasia Bobkova[3], Vladislav Myasnikov[4]
[1]*Ural Federal University*,
[2]*Institute of Immunology and Physiology Ural branch of Russian Academy of Sciences (RAS)*
[3]*OOO "IT-Med"*
Ekaterinburg, Russia
[4]*Samara National Research University*
Samara, Russia
v.v.zyuzin@urfu.ru

***Abstract*—Nowadays ultrasound studies of the heart, also called echocardiography (EchoCG), are widespread in modern cardiology. One of the most important steps in estimating the health of the heart is the tracking and segmentation of the left ventricular (LV) endocardial border from EchoCG, which is used for measuring the ejection fraction and assessing the regional wall motion [1]. The disadvantage of these methods is the necessity to apply image processing manually or in a semi-automatic mode, which requires special knowledge and skills. As a result, the issue of an automatic tracking and segmentation of the LV on EchoCG-images is an actual and practical problem.**

The capabilities of the fully trained model of the convolutional neural network Unet for automatic identification of the LV region are explored in this paper. The obtained accuracy of LV segmentation is up to 92.3%, which suggests the expediency of using Unet for automatic identification of the LV endocardial border on ultrasound images.

***Index Terms*—segmentation, ultrasound images, left ventricle, deep learning, Unet**

I. Introduction

Ultrasound studies for the non-invasive methods of visualization of internal organs are widespread in various fields of medicine [1], [2]. In cardiology, this method is known as echocardiography (EchoCG). One of the most important steps in estimating the health of the heart is the tracking and segmentation of the left ventricular (LV) endocardial border from EchoCG, which is used to measure the ejection fraction and to assess the regional wall motion [3]. Typically, the ultrasound images of the LV are analyzed by an expert (e.g., a cardiologist), who segments the endocardial border of the LV manually or in a semi-automatic mode. The expert defines region of interest by indicating 3 basis points (edges of mitral annulus and apex) (Fig. 1). The manual and semi-automatic segmentation of the LV exposes the following issues: 1) it is a tedious and time demanding task that can be only performed by a specialized clinician; and 2) it is prone to poor repeatability. These issues can be solved with the use of an automatic LV segmentation system, which has the potential to improve the workflow in a clinical site and to decrease the variability between user segmentations. At present, many studies on the development of automatic algorithms for delineation of the LV on ultrasound images obtained by the method of two-dimensional EchoCG are carried out [4]–[7]. However, the issues of generalizing the experience of experts in the set of formal rules for LV segmentation have not been solved yet, despite the existing guidelines of the American Society of Echocardiography [8]. As a result, the issue of an automatic tracking and segmentation of the LV on EchoCG-images is an actual and practical problem.

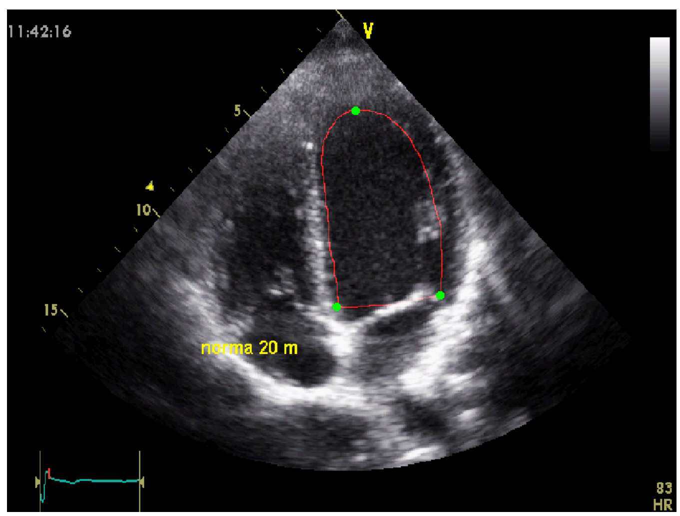

Fig. 1. Ultrasonic image of the heart in the apical four chamber view. Red curve is the endocardial border of the LV, marked by an expert. Green points - 3 basis points (edges of mitral annulus and apex)

II. Data

Ultrasound LV images in classical 2D apical four-chamber view for 94 patients were recorded during the entire cardiac cycle simultaneously with EchoCG recording by ultrasound system Philips IE33. From the data set, frames at the time of the minimum volume of the LV (end systole) and at the time of maximum LV volume (end diastole) were selected for each patient. In total, 188 frames have been used, the size of frames varied from 320×240 to 832×644 pixels. In each of the images used there was an endocardial border of the LV, marked by an expert. It should be noted that all images have been resized to 640×480 by using bilinear interpolation.

III. Unet architecture

In our studies, we used the deep convolutional neural network (CNN) Unet described by Olaf Ronneberger et al. [9]. The choice of the CNN has been made because it is suitable for a segmentation of medical images, even if the dataset for training is small. The Unet architecture is presented in tab. I. The architecture of the CNN Unet consists of two parts (tab. I):

TABLE I
UNET ARCHITECTURE

Layer Name	Block Type	Output Resolution	Output Width
Input		640×480	1
Down 1_1	conv 3×3	640×480	32
Down 1_2	conv 3×3	640×480	32
Pooling 1	maxpooling	320×240	32
Down 2_1	conv 3×3	320×240	64
Down 2_2	conv 3×3	320×240	64
Pooling 2	maxpooling	160×120	64
Down 3_1	conv 3×3	160×120	128
Down 3_2	conv 3×3	160×120	128
Pooling 3	maxpooling	80×60	128
Down 4_1	conv 3×3	80×60	256
Down 4_2	conv 3×3	80×60	256
Pooling 4	maxpooling	40×30	512
Accross 1	conv 3×3	40×30	512
Accross 2	conv 3×3	40×30	512
Accross Up 1	upsampling	80×60	256
Merge 1 (with Down 4_2)	concatenate	80×60	512
Up 1_1	conv 3×3	80×60	256
Up 1_2	conv 3×3	80×60	256
Up 1	upsampling	160×120	128
Merge 2 (with Down 3_2)	concatenate	160×120	256
Up 2_1	conv 3×3	160×120	128
Up 2_2	conv 3×3	160×120	128
Up 2	upsampling	320×240	64
Merge 3 (with Down 2_2)	concatenate	320×240	128
Up 3_1	conv 3×3	320×240	64
Up 3_2	conv 3×3	320×240	64
Up 3	upsampling	640×480	32
Merge 4 (with Down 1_2)	concatenate	640×480	64
Up 4_1	conv 3×3	640×480	32
Up 4_2	conv 3×3	640×480	32
Output	conv 1×1	640×480	1

- The contracting path converts the image into a feature set. It consists of repeated application of two convolutional layers (conv 3x3) and 2x2 max pooling operation (maxpooling).
- The expansive path converts the feature set into an image. Every step of it consists of repeated application of two convolutional layers (conv 3x3), the upsampling layer that halves the number of feature channels and concatenation with the correspondingly feature level of the contracting path.

It should be noted that behind each convolution layer, except the last layer, there is an activation layer with ELU function (Exponential linear units):

$$f(x) = \begin{cases} x & \text{if } x \geq 0 \\ e^x - 1 & \text{otherwise} \end{cases}, \tag{1}$$

where x is the pixel value of the image after the convolution layer.

The last convolution layer converts the value of pixels into the interval [0, 1] (the pixel does not belong to the LV area or belongs to the LV area), so here the activation function of sigmoid is used:

$$f(x) = \frac{1}{1 - e^{-x}} \tag{2}$$

The activation layers extract the significant attributes of the image when the model is trained.

IV. Metric for model evaluation

The Dice coefficient as a quantitative measure of LV segmentation quality is calculated by the formula:

$$Dice = \frac{2 \sum_i \sum_j y_{i,j}^{true} y_{i,j}^{pred}}{\sum_i \sum_j y_{i,j}^{true} + \sum_i \sum_j y_{i,j}^{pred}} \tag{3}$$

where $y_{i,j}^{true}$ is the pixel value of the image with the LV expert area, taking the values 0 or 1, $y_{i,j}^{pred}$ is the pixel value of the image with the LV region obtained by the model and taking values into the interval [0, 1].

For training the model a loss function has been used. Its value is equal to the value of the Dice coefficient taken with the opposite sign.

V. Experiments

Assessment of the chosen architecture has been carried out by using cross validation. For this the initial data was divided into five blocks. At each step, the model has been trained in a training subsample consisting of four blocks, and one block has been used as a control subsample to evaluate the quality of the model. Thus, five models of CNN with trained weights have been constructed. It should be noted that the model optimizer used Adam (Adaptive Moment Estimation) [9], the batch size is 4. The CNN architecture trained on the initial data achieved segmentation with an average accuracy of 87.6% (tab. II). An example of segmented endocardial border is shown in fig. 2

In order to improve the quality of segmentation, new models have been built on artificially augmented data. Augmented data have been obtained by using:

- random vertical shifting of the image up or down up to 20% of the image height;

- random horizontal shifting of the image left or right up to 20% of the image width;
- random increase or decrease of the image in the range from 0.5 to 2 image sizes.

For the estimation of these new models quality the cross validation has been carried out as described before. Note that the augmentation has been applied only to the training subsamples, which allowed to increase its volume up to 4000 images. The validation has been carried out on the original unchanged data. The use of data augmentation has increased the accuracy of LV segmentation up to 92.3% (tab. II). An example of segmented endocardial border is shown in fig. 3.

TABLE II
ACCURACY OF TRAINED MODELS: BEFORE AND AFTER AUGMENTATION

Model	Dice (initial data)	Dice (augmented data)
1	0.864	0.927
2	0.890	0.918
3	0.877	0.924
4	0.875	0.920
5	0.875	0.927
Average	0.876	0.923

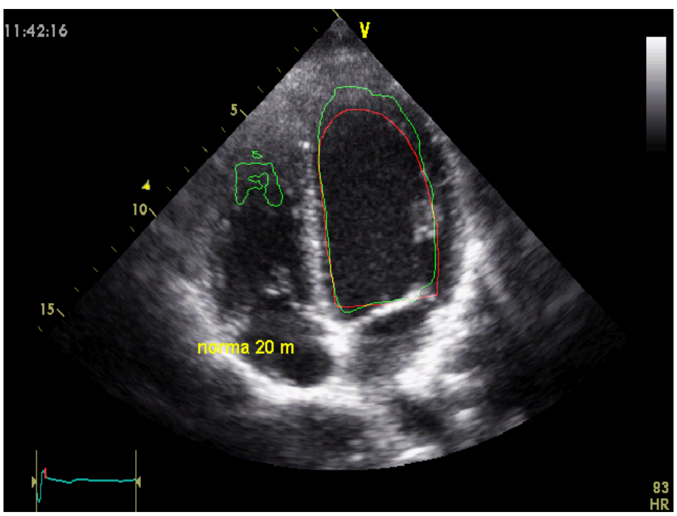

Fig. 2. Ultrasound image with segmented LV border. Red curve is the endocardial border of the LV marked by an expert. Green curve is the endocardial border marked by model trained on initial data

VI. CONCLUSION

The presented results show the possibility of using the CNN Unet for automatic segmentation of the LV area on ultrasound images. Thus, Unet architecture can be used for segmentation LV on other heart projections.

Further studies are connected with comparing of different famous CNNs like FCN, SegNet and etc. The detailed analysis is needed for varios group of patients because of specific LV form that appears with heart diseases.

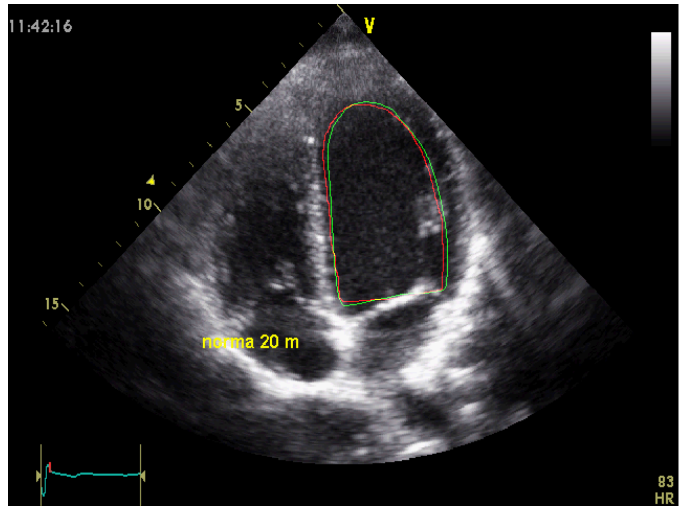

Fig. 3. Ultrasound image with segmented LV border. Red curve is the endocardial border of the LV marked by an expert. Green curve is the endocardial border marked by model trained on augmented data

ACKNOWLEDGMENT

The study was carried out with the financial support of the RFFI in the framework of the scientific project No 17-31-50073 ”mol_nr”, the Program of the Presidium RAS #27 (project -18-118020590030-1) and supported by RF Government Act #211 of March 16, 2013 (agreement 02.A03.21.0006).

REFERENCES

[1] Flower M. A. (ed.), Webb’s physics of medical imaging. CRC Press, 2012.
[2] Feigenbaum H., Echocardiographia. M. : Vidar, 1999.
[3] J. A. Noble and D. Boukerroui, “Ultrasound image segmentation: A survey,” IEEE Trans. Med. Imag., vol. 25, no. 8, 2006, pp. 987–1010.
[4] AlcevskaL E. “Segmentation of the Left Ventricle of the Heart in 2D Ultrasound Images using Convolutional Neural Networks,” //Master, Department of Signals and Systems, CHALMERS UNIVERSITY OF TECHNOLOGY, CHALMERS UNIVERSITY OF TECHNOLOGY, EX093/2016. 2016.
[5] Carneiro G., Nascimento J. C., Freitas A. “The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods,” IEEE Transactions on Image Processing, . 21, . 3, 2012, pp. 968–982.
[6] Carneiro G., Nascimento J. C. “Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data,” IEEE transactions on pattern analysis and machine intelligence, . 99, No. 1, 2013.
[7] Raynaud C. et al. “Handcrafted features vs ConvNets in 2D echocardiographic images,” Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on IEEE, 2017, pp. 1116–1119.
[8] R. M. Lang, L. P. Badano, V. Mor-Avi, J. Afilalo, A. Armstrong, L. Ernande, F. A. Flachskampf, E. Foster, S. A. Goldstein, T. Kuznetsova, 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,” Journal of the American Society of Echocardiography, 28(1):139, 2015.
[9] Ronneberger O., Fischer P., Brox T. “U-net: Convolutional networks for biomedical image segmentation,” International Conference on Medical image computing and computer-assisted intervention, Springer, Cham, 2015, pp. 234–241.
[10] Kingma D. P., Ba J. “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.