E3S Web of Conferences **460**, 04026 (2023) https://doi.org/10.1051/e3sconf/202346004026
BFT-2023

Solving the problem of calculating strain indices in echocardiography using deep learning neural networks

Anna Alekhina[1,2*], *Mikhail* Dorrer[1], *Mikhail* Sadovskiy[3,4,5], *Vitaly* Sakovich[5], and *Igor* Demichev[5]

[1] Siberian State University of Science and Technology named after M.F. Reshetnev, 660037 Krasnoyarsk, Russia
[2] Siberian Federal University, Artificial Intelligence Laboratory, 660041 Krasnoyarsk, Russia
[3] Institute of Computational Modeling SB RAS, 660036 Krasnoyarsk, Russia
[4] Federal Siberian Research and Clinical Centre FMBA, 660000 Krasnoyarsk, Russia
[5] Krasnoyarsk State Medical University (KrasSMU), 660022 Krasnoyarsk, Russia

Abstract. The paper sets the task of calculating the parameters of deformation of the heart muscle according to echocardiogram data under interference conditions, for example, in the study of children. Indicators of deformation (strain value) of the heart muscle were used by us to determine the presence and severity of dysfunction of the chambers of the heart in atrial septal defect – a congenital heart defect characterized by the presence of communication between the right and left atria. The problem was solved by analyzing the video stream obtained from the installation of echocardiography using a set of deep learning neural network architectures designed for image segmentation. The study was conducted for the U-net architecture. As a result of processing the video stream, it was possible to solve the problem of segmentation of the walls of the heart muscle and binding of key points in the condition of interference in the removal of an echocardiogram on child patients unable to remain motionless during the study. The obtained indicators provide the cardiologist with important information for determining the dysfunction of the chambers of the heart (especially the right atrium, the most compromised chamber of the heart in the studied cases) with a defect of the atrial septum.

1 Introduction

Echocardiography is an important ultrasound imaging technique widely used to assess the structure and function of the heart. An echo signal can provide valuable diagnostic information, but image analysis requires a lot of time and resources, and also depends significantly on differences in echocardiography equipment. The breakthrough in the field of image processing with the help of artificial neural networks that took place in the 2010s made it possible to automate image processing and reduce analysis time, as well as reduce the dependence of processing procedures on the specifics of equipment. However, final solutions

* Corresponding author: dorrer_mg@sibsau.ru

in this area have not yet been received and work on processing an echocardiogram using artificial neural networks is currently actively underway.

Atrial septal defect is one of the most widespread congenital heart defects in the modern population, occurring with a frequency of 3.89 per 1000 children and 0.88 per 1000 adults. In the case of preservation of fetal atrial communication after birth, an atrial septal defect is formed. It includes both real septal membrane anomalies and additional anomalies that provide communication between the two atria. In most cases, small fenestra of the atrial septum close spontaneously during development. Large persistent tissue defects may require percutaneous or surgical intervention to avoid consequences such as the development of heart failure, stroke, arrhythmia and pulmonary hypertension.

In this paper, we solved a particular, but quite important in practice, problem of measuring strain, the value of which is planned to be used in the future as one of the criteria for constructing the tactics of dynamic management of patients with atrial septal defect (DMPP). The specificity of this study is that in a significant part of cases it is necessary to produce it in children, including newborns. A significant difficulty in interpreting the echocardiography of such patients is that they are unable to remain motionless during the echocardiogram removal session, which is why the image is constantly shifting in all three coordinates, and this makes it difficult to measure the dynamics of the size of the heart muscle by existing methods.

2 Related works

Due to its prevalence and severity of possible consequences for the patient's health, an atrial septal defect is an object of active interest for both practitioners and medical scientists. A study of the publications of the last three years shows a large number of works devoted to the study of this pathology.

Thus, the work [1] is devoted to determining the etiology of atrial septal defect, describing the clinical picture of a patient with atrial septal defect, available treatment options and management of atrial septal defect, as well as explaining interprofessional team strategies to improve treatment. coordination and communication for patients with atrial septal defect.

In the article [2], the authors conducted a comparative study of two methods of atrial septal defect treatment - a minimally invasive method of atrial septal defect closure, including total peripheral cannulation and an anterior mini-thoracotomy incision 5 cm or less in length, with access to median sternotomy.

Work [3] describes the examination of children with atrial septal venous sinus defect (SV-DMPP) for the presence of concomitant vascular anomalies.

In [4], acute changes in biventricular longitudinal deformation after transcatheter closure of the atrial septal defect and its relationship with the size of the device were evaluated. The magnitude of the global longitudinal deformation of the right ventricle (2-dimensional strain) was used as an important diagnostic feature.

The strain index was used in the article [5] to assess the effect of percutaneous closure of the atrial septal defect (DMPP) on the function of the left ventricle (LV) in adult patients.

Approaches to processing echocardiographic images by means of artificial neural networks are also being actively studied.

Thus, in [6], the authors used a logically transparent deep neural network to classify echocardiography images, obtaining an average classification accuracy of 98.2%.

The authors of [7] proposed a set of convolutional neural networks MENN designed to analyze images of the left ventricle both in the brightness mode along the long axis (B) and in the movement mode along the short axis (M). The efficiency of their proposed architecture was additionally confirmed on two preclinical models and achieved excellent correlation indicators (value Pearson correlation coefficient r from 0.85 to 0.99) between the results of automatic and manual analysis of echocardiography images.

The paper [8] describes the use of a deep neural network with an attention mechanism and a residual feature aggregation module for segmentation of the left ventricle in transesophageal echocardiography images during cardiopulmonary resuscitation. The aim of the work was to determine the best position for effective chest compression using the systolic function of the left ventricle.

Work [9] describes the solution of the problem of detecting myocardial infarction on echocardiography recordings. The model proposed by the authors is a pipeline consisting of two-dimensional convolutional neural networks that performs preliminary data processing by segmenting the left ventricular chamber from an apical four-chamber representation, followed by a three-dimensional convolutional neural network that performs binary classification of the fact of a heart attack on the analyzed record. The model achieved high accuracy rates - 90.9% accuracy, 100% precision, 95% recall and 97.2% F1_score.

The authors of [10] solved the problem of tracking the aortic valve in the echocardiography video stream. Two convolutional neural network architectures were used to solve the detection problem - a single-shot Multibox detector (solid-state drive) faster and a region-based convolutional neural network (RNN).

Thus, it can be stated that the topic chosen in the work is quite relevant in modern conditions at the junction of such sciences as medicine and computer science in terms of the tasks of artificial intelligence and computer vision.

3 Methods and materials

3.1 Applied methods

To solve this problem, it was decided to use the Unet segmentation model [11] and image preprocessing by Gauss filtering methods [12], the Canny boundary search operator [13] and standard OpenCV library operations [14].

To calculate the strain parameter, it is necessary to find the extreme points of the analyzed atrium. Before that, due to the problem of strong image noise and fragmentary visibility of the area of interest, it is necessary to find fragments of the wall contour.

Search for the upper and lower contour by searching for the maximum and minimum location of the centroid of a closed contour and then finding the leftmost point of the upper fragment of the contour and the lowest point of the lower fragment of the contour.

After obtaining the values of the length of the distance vector between the extreme points, an averaging filter was used to smooth out the outliers caused by errors in the selection of the contours of the heart muscle.

Formula (1) for calculating the strain parameter:

$$\varepsilon(t) = \frac{L(t) - L(t_0)}{L(t_0)} \tag{1}$$

where L(t) is the length of the distance vector between the extreme points on the current frame, L(t0) is the length of the distance vector between the extreme points on the previous frame.

3.2 Experimental materials

To develop an algorithm for calculating the atrial strain parameter, it is necessary to pre-process a video fragment obtained from an echocardiogram. For segmentation of the atrial walls and calculation of strain indicators, video clips of EchoCG sessions of 40 patients diagnosed with atrial septal defect were collected, in which indications for surgical correction of the defect were determined. The study was conducted in the Federal Agricultural Research Center, the Russian Federation, in the city of Krasnoyarsk. The age of patients ranges from 0 to 18 years.

The original resolution of the video fragments has 1344x1000 pixels. 25 frames per second, image in RGB color space. The video is dominated by strong noise, characteristic when using EchoCG installations. After removing the extra fragments of the system, the image became 800x600 pixels in size in the RGB color space. An example of the images is shown in Figure 1.

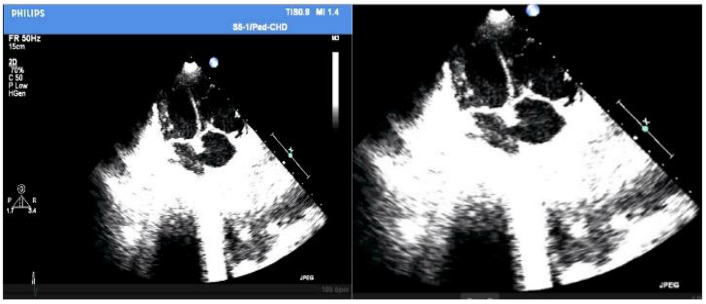

Fig. 1. Frame fragmentation for neural network training preparation

Fragmented images were transferred to the experts to mark up the data. They were tasked with manually segmenting the area of the right and left atria under the “atrium” class as an area of interest for further training of neural networks. To train the model, the sample was divided into training, validation and test samples in the ratio of 70%:20%:10%. In the end, the following sample was obtained: 1500 images of the heart in motion (1 compression cycle) of patients in the dataset. 3000 marked atria (left and right). Figure 2 shows an example of manual segmentation of atrial regions and obtaining a common mask for further training of the neural network. The layout of the dataset was performed in the Labelme markup program [15].

Fig. 2. The marking of the atria in the image, where 1) the mask of the left atrium, 2) the mask of the right atrium, 3) the general mask of both atria, 4) the addition of the image of the heart and the general mask of the atria.

Neural network training was performed on a personal computer with the following characteristics: NVIDIA GeForce 2060, 16 Gb RAM.

4 Results

4.1 Results of segmentation model training

The created model based on the architecture of the convolutional neural network Unet [11] has a total number of 6,502,786 neurons, of which 4,658,882 are training ones.

The input of the neural network is 512x512 pixels in the RGB color space. After segmentation by the model, the image size increases to the size of the analyzed video fragment, in this case to the size of 800 by 600 in the RGB color space. The full structure of the created model is shown in the figure. The training took place in 18 epochs, 147 iterations in 1 epoch, the tensor dimension of 32 images. Learning outcomes loss0.013 val_loss0.017. SparseCategoricalCrossentropy loss function.

```
Model: "model_9"
__________________________________________________________________________________________________
 Layer (type)                   Output Shape         Param #     Connected to
==================================================================================================
 input_13 (InputLayer)          [(None, 512, 512, 3  0           []
                                )]

 model_8 (Functional)           [(None, 256, 256, 9  1841984     ['input_13[0][0]']
                                6),
                                 (None, 128, 128, 1
                                44),
                                 (None, 64, 64, 192
                                ),
                                 (None, 32, 32, 576
                                ),
                                 (None, 16, 16, 320
                                )]

 sequential_25 (Sequential)     (None, 32, 32, 512)  1476608     ['model_8[0][4]']

 concatenate_11 (Concatenate)   (None, 32, 32, 1088  0           ['sequential_25[0][0]',
                                )                                 'model_8[0][3]']

 sequential_26 (Sequential)     (None, 64, 64, 256)  2507776     ['concatenate_11[0][0]']

 concatenate_12 (Concatenate)   (None, 64, 64, 448)  0           ['sequential_26[0][0]',
                                                                  'model_8[0][2]']

 sequential_27 (Sequential)     (None, 128, 128, 12  516608      ['concatenate_12[0][0]']
                                8)

 concatenate_13 (Concatenate)   (None, 128, 128, 27  0           ['sequential_27[0][0]',
                                2)                                'model_8[0][1]']

 sequential_28 (Sequential)     (None, 256, 256, 64  156928      ['concatenate_13[0][0]']
                                )

 concatenate_14 (Concatenate)   (None, 256, 256, 16  0           ['sequential_28[0][0]',
                                0)                                'model_8[0][0]']

 conv2d_transpose_31 (Conv2DTra  (None, 512, 512, 1)  1441       ['concatenate_14[0][0]']
 nspose)
```

Fig. 3. The architecture of the trained model based on the Unet architecture

4.2 Preliminary preparation of the image for the calculation of the strain

A video from an ultrasound machine with a heartbeat rate of 5-10 iterations is fed into the system for automatic calculation of the strain parameter. The video image size is 1344x1000 pixels in the RGB color channel. A fragment is cut from a video with a window of 600x800 pixels in the RGB color channel. The resulting image is reduced to the size of the input layer of the trained U-net neural network (512x512 pixels) and at the output we get segmented masks of the right and left atria of the heart. Then the original image and mask are multiplied and an image of segmented areas of 800x600 pixels in the RGB color channel is obtained. For further contour search, the following operations must be performed: grayscale conversion, Gaussian filtering with an 11x11 pixel window to remove coarse noise and contour search using the Canny operator with parameter 30, 150. Since the image is very noisy, most of the contours are unclosed. To solve this problem, the pixel value was supplemented with the dilate function [16]. The result of the operation are closed fragments of the visible contour of the atrial wall of the heart prepared for strain calculation. Figure 4 shows a step-by-step description of the image preprocessing.

Fig. 4. Algorithm for obtaining the contour of the area of interest of the heart.

4.3 Calculation of the strain value of the right atrium

After obtaining the contour of the atria, it is necessary to find and fix the extreme upper point and the extreme lower point of the analyzed area. Since there is a strong noise and partial visibility of the organ in the ultrasound images, the visible part of the atrium wall is outlined indistinctly. Figure 5 shows examples of obtaining contours with varying degrees of noise in the zone of interest.

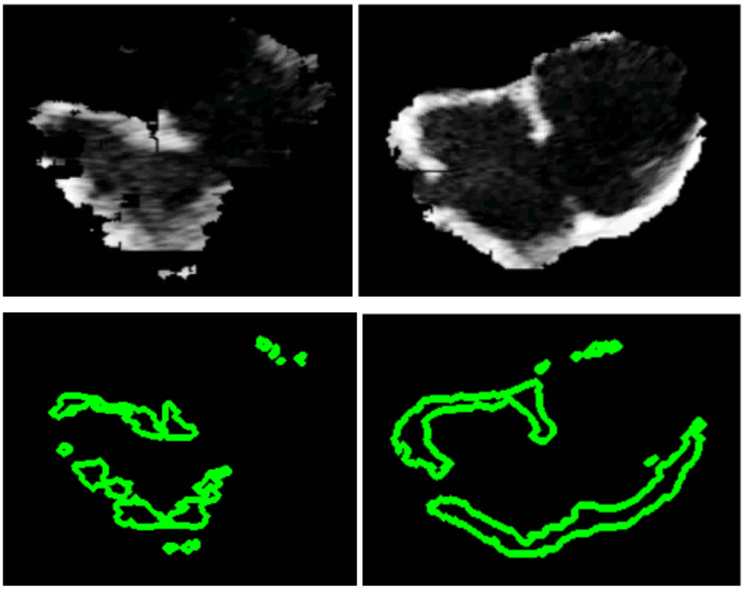

Fig. 5. Examples of contour extraction from EchoCG video frames, on the left – a highly noisy atrial region, on the right – an atrial region clean of noise

After finding the extreme points of the necessary contour fragments, the length of the resulting distance vector is calculated. Figure 6 shows the results of finding the points and the vector found. The red dot is the left extreme point of the upper wall fragment, the green dot is the lower point of the lower wall fragment, the green frame is the bounding box of the found closed contour, the pink line is the distance vector between the extreme points.

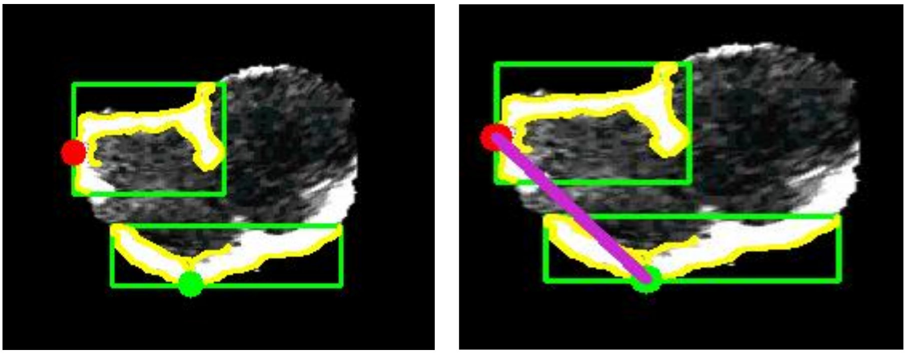

Fig. 6. The result of the search for the extreme points of the visible sections of the walls of the left atrium.

4.4 Calculation of the instantaneous and average values of the strain of the left atrium

After forming a list of the type [[number_frame], [lenght]], the data was sorted in ascending order and copied into a one-dimensional ordered array. Then, a median filter was used to remove noise in the data obtained after calculating the length of the distance vector between the extreme points. After the received noise-free signal, the instant strain was calculated according to formula 1. For a better understanding, the data obtained has been converted into percentages. In the table you can see the results of calculating the strain in one stroke of compression of the right atrium, the data arrays obtained and the stages of their transformation

Table 1. Protocol for step-by-step calculation of instant strain

Frame number	Original length	After the median filter	Instant Strain	Instant Strain, %
0	75.610	75.610	0	+0
1	100.004	75.610	0.322	+32
2	52.325	100.004	-0.395	-40
3	121.490	60.440	0.514	+51
4	60.440	91.547	-0.339	-34
5	91.547	60.440	0.032	+3
6	58.051	62.393	0	+0
7	62.393	62.393	0.613	+61
8	112.445	100.687	0	+0
9	100.687	100.687	0	+0
10	80.529	100.687	0.010	+1
11	108.756	101.710	0	+0
12	101.710	101.710	0.010	+1
13	38.948	85.023	0	+0
14	85.023	55.901	-0.164	-16
15	55.901	85.023	-0.342	-34
16	102.004	102.004	0.520	+52
17	128.082	102.004	0.199	+20

18	58.034	123.709	0	+0
19	123.709	111.072	0.212	+21
20	111.072	123.709	-0.102	-10
21	150.416	111.072	0.113	+11
22	49.040	91.005	-0.102	-10
23	91.005	91.005	-0.180	-18
24	130.667	91.005	0	+0
25	10	100.498	0	+0
26	100.498	93.005	0.104	+10
27	93.005	93.005	-0.074	-7
28	67.601	93.005	0	+0

Figure 7 shows graphs of the received data and the stages of their change.

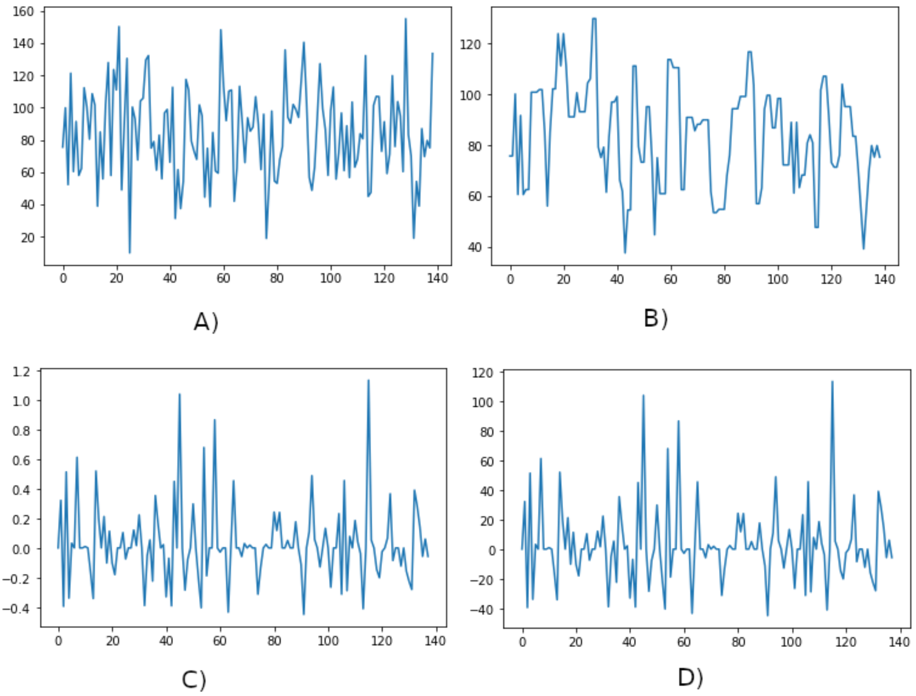

Fig. 7. Strain calculation graphs, where A) is a graph of the source data, C) is a graph of data using a median filter, C) is a graph of the instant strain, D) is a percentage display of the instant strain

5 Conclusion

The authors expect that the resulting solution will improve the quality of diagnostics and help functional diagnostics doctors and cardiologists more accurately determine the necessary timing of planned surgical treatment of DMPP and create a new examination protocol using new decision support methods.

The prospect of work from the point of view of applied information technologies is to solve the problem of strong left atrium noise in the EchoCG image. The authors plan to use the method of restoring the invisible wall of the right atrium using generative-adversarial neural networks (GAN) to solve this problem, which will allow calculating the independent parameters of the instantaneous and average strain for both atria and thereby diagnose in more detail the dysfunction of the heart caused by a defect of the atrial septum.

References

1. O. Elsaka et al, Asian J. Med. Princ. Clin. Pract. **4(4),** 207–220 (2021)
2. A. Bhattarai et al, J. Nepal Health Res. Counc. Nepal, **19(4),** 725–729 (2022)
3. A. Akbar et al, Cureus (2022)
4. A. Alkhateeb et al, Cardiol. Young **32(5),** 746–754 (2022)
5. T. Miki et al, Int. J. Cardiovasc. Imaging **37(3),** 903–912 (2021)
6. M. Huang et al, Eur. Hear. J. - Cardiovasc. Imaging **23(1)** (2022)
7. C. Duan et al, Am. J. Physiol. Circ. Physiol. **323(4),** H628–H639 (2022)
8. S. Kang et al, PLoS One / ed. Liu X. **18(1),** 0280485 (2023)
9. O. Hamila et al, Multimed. Tools Appl. **81(26),** 37417–37439 (2022)
10. M.H.A. Nizar et al, Curr. Med. Imaging Former. Curr. Med. Imaging Rev. **16(5),** 584–591 (2020)
11. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation (2015)
12. J. Andrews, Low-Pass Risetime Filters for Time Domain Applications (2023)
13. J.A. Canny, IEEE Trans. Pattern Anal. Mach. Intell. **6,** 679–698 (1986)
14. OpenCV [Electronic resource]. URL: https://opencv.org/
15. LabelMe [Electronic resource]. URL: http://labelme.csail.mit.edu/Release3.0/
16. Eroding and Dilating [Electronic resource] // Open Source Computer Vision. URL: https://docs.opencv.org/3.4/db/df6/tutorial_erosion_dilatation.html