

Myocardium Segmentation using Two-Step Deep Learning with Smoothed Masks by Gaussian Blur

Vitalii Slobodzian^a, Pavlo Radiuk^a, Anastasiia Zingailo^b, Olexander Barmak^a, and Iurii Krak^{c,d}

^a Khmelnytskyi National University, 11, Instytuts'ka str., Khmelnytskyi, 29016, Ukraine

^b Clinic of Oxford Medical Khmelnytskyi, 10, Podilska str, Khmelnytskyi, 29013, Ukraine

^c Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska str., Kyiv, 01601, Ukraine

^d Glushkov Cybernetics Institute, 40, Glushkov ave., Kyiv, 03187, Ukraine

Abstract

Nowadays, cardiac magnetic resonance images face challenges in distinguishing between inflamed and non-inflamed tissues due to subtle color variations rather than clear density distinctions. Pixel values in these images vary based on individual subjects and the MRI equipment, making them inconsistent across different training datasets. Thus, detecting inflamed tissues in MRIs largely depends on the expertise of interpreting physicians, making it time-consuming and complicating the training of accurate classifiers. To address this issue, in this study, we propose a novel approach for myocardium segmentation on MRI images utilizing a two-stage neural network process coupled with mask refinement. The initial network outlines the myocardium, which is then fine-tuned by the second network for precise myocardium segmentation. A key enhancement involves mask post-processing via Gaussian blur, where the blur coefficient is automatically adjusted. Experimental outcomes demonstrated an increase in the Dice coefficient from 0.889 to 0.894 upon removing non-essential labels. Moreover, using a dual-model approach for myocardium localization and contour definition elevated the coefficient to 0.938. Employing the Gaussian blur during mask resizing culminated in an impressive average Dice coefficient of 0.955.

Keywords ¹

Cardiac MRI, myocardium segmentation, medical image analysis, improved mask, deep learning, human-in-the-loop

1. Introduction

Segmenting ventricles in cardiac magnetic resonance imagery (MRI) using artificial intelligence (AI) systems and means has become pivotal in diagnosing a plethora of cardiac ailments, encompassing conditions like pulmonary hypertension, dysplasia, coronary heart disease, and cardiomyopathies. The MRIs that display myocardium, the right ventricle (RV), and the left ventricle (LV) pose significant challenges due to the heart's intricate anatomy and motion. The myocardium's segmentation is particularly taxing and time-consuming due to the following factors [1]: (i) unclear ventricular boundaries resulting from blood movement and the partial volume effect, (ii) wall irregularities within the ventricle that match the grey scale of adjacent tissues, and (iii) the right ventricle's intricate crescent shape that fluctuates based on the MRI slice [2]. Presently, cardiologists manually undertake myocardium segmentation in medical facilities. This labor-intensive procedure demands approximately 10 to 15 minutes of a specialist's time [2] and is susceptible to variations both between and within

¹IDDM'2023: 6th International Conference on Informatics & Data-Driven Medicine, November 17–19, 2023, Bratislava, Slovakia
EMAIL: vitalii.slobodzian@gmail.com (V. Slobodzian); radiukpavlo@gmail.com (P. Radiuk); stjascha89@gmail.com (A. Zingailo); alexander.barmak@gmail.com (O. Barmak); yuri.krak@gmail.com (I. Krak)
ORCID: 0000-0001-8897-0869 (V. Slobodzian); 0000-0003-3609-112X (P. Radiuk); 0009-0004-8026-9639 (A. Zingailo); 0000-0003-0739-9678 (O. Barmak); 0000-0002-8043-0785 (I. Krak).



© 2023 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

experts. As a result, the functional evaluation of the myocardium has often been deemed secondary to the LV and RV, leaving ample room for advancements in myocardium segmentation.

While many solutions to the myocardium segmentation issue have employed conventional methodologies [3-6], like traditional machine learning (ML), their performance, despite being commendable, has not attained the accuracy benchmark set by human experts (0.90 Dice score) [2]. Many of these strategies are tailored for specific datasets, necessitating substantial modifications for alternative data. Conversely, deep learning (DL) techniques [7-10] have equaled the efficacy of these traditional methods without requiring extensive alterations. Though they have not yet exceeded human precision, and in some instances might fall short of traditional methods, they exhibit superior generalizability across diverse datasets.

In this study, we propose a novel approach based on the use of the traditional myocardial segmentation method for preliminary localization of the image size to the original using Gaussian blurring and automatic selection of the blurring coefficient, which allows for improving the accuracy of segmentation. Our contribution aims to obtain more accurate results for processing cardiac MRI.

The paper's structure is as follows: Section II offers a concise review of the myocardium segmentation literature, encompassing both traditional ML and DL techniques. In Section III, we elaborate on the localization and segmentation methods, dataset, and data processing techniques. Section IV showcases our comparative findings across different segmentation strategies. Finally, our conclusions and plans for future research are presented in Section V.

2. Related works

The body of work on myocardium is not as extensive as that on LV and RV segmentation [11]. Traditional methods used for myocardium segmentation encompass techniques like graph cuts, deformable models, level sets, and atlas-based strategies. Although these methods are primarily tailored to heart geometry, they provide valuable insights for effective myocardium segmentation. Notably, grid generation [4] and shape model-based techniques [5] are favored for RV segmentation. A recent method introduced in [12] employs a hybrid graph-based approach, aggregating per-slice segmentation results to achieve comprehensive cardiac segmentation. Nevertheless, our research advocates for a DL approach to myocardium segmentation, targeting enhanced adaptability and practicality.

The U-Net model [13] is among the most favored DL architectures for a wide range of biomedical segmentation tasks. The study in [14] delves into the applicability of U-Net for myocardium segmentation, utilizing a standard U-Net framework. Additionally, both [7] and [14] investigated the dilated variant of U-Net. Notably, the dilated U-Net reportedly surpasses the original U-Net's performance [16], offering a broader field of view without an uptick in parameter count. The modified U-Net described in [8], which employs a single convolution layer in place of dual convolution layers at each level, aligns with the results presented in [14]. An approach rooted in the fully convolutional network (FCN) was introduced by [15], but it has a significant limitation, i.e., only high-level functions are considered in the decoder, while low-level functions are ignored. Ensembling DL models has been notably effective for resolving ambiguous cases. Majority and average voting emerge as the preferred ensembling methods. We also explore two-stage methodologies in the subsequent section.

In [17], a strategy was introduced that employs a cascaded convolutional neural network (CNN) [9] network for isolating the region-of-interest (ROI), followed by another CNN for fully automatic segmentation of the right ventricle in cardiac MRI. Similarly, in [18], the authors put forward a method combining CNN and stacked autoencoders to identify the ROI and subsequently outline the myocardium.

A growing concern within the scientific community revolves around the "black box" nature of AI algorithms, particularly in the realm of healthcare diagnostics. For instance, in our previous work [19], we emphasized the importance of trust and transparency in AI systems, noting that while AI can produce high results, it may not always meet the normative service expectations of professionals like doctors. The research also presents a human-in-the-loop approach, which, despite its promising results, has limitations related to the need for updated databases for MRI. Another work [20] delves into the integration of a quality control tool in the MRI segmentation pipeline. This tool aims to facilitate a human-in-the-loop framework for DL-based analysis, especially in cardiac segmentation. The research

underscores the importance of time and effort efficiency for human experts in a collaborative framework with AI.

A survey of the literature reveals a multitude of architectural choices within the DL realm tailored for myocardium segmentation. Several enhancements [21], ranging from network modifications, and post-processing techniques like the fully connected conditional random field, and cyclic learning rate scheduler training methods, to the amalgamation of diverse models, have the potential to elevate the performance of DL models in the myocardium segmentation.

It is worth discussing in more detail the works that use the same dataset that it used as based in this research. The approach [22] involved testing the U-Net and FCN architectures with various hyperparameters. They also experimented with the use of 2D and 3D convolutional layers and training with Dice loss function versus cross-entropy loss. The authors' best-performing architecture was U-Net with 2D convolutional layers trained using cross-entropy loss. The Dice coefficient for myocardium segmentation was 0.901. Approach [23] implemented an ensemble of 2D and 3D U-Net architectures (with residual connections at the upsampling layers). In the 3D network, due to significant inter-slice gaps in input images, pooling, and upsampling operations are performed only in the short-axis plane. Additionally, due to memory constraints, the 3D network has fewer feature maps. Both networks were trained with the Dice loss function. The Dice coefficient for myocardium segmentation was 0.919. Approach [24] implemented the "M-Net" architecture, where the main difference from U-Net is that the feature maps of the decoding layers are concatenated with the same layers. The corresponding network was trained with a weighted cross-entropy loss function. The Dice coefficient for myocardium segmentation was 0.895.

Overall, in cardiac MRI, which leverages the principles of nuclear magnetic resonance, discerning between inflamed and non-inflamed soft tissues presents inherent challenges. This complexity stems from the subtle variations in image appearance: inflamed tissue typically appears darker than its non-inflamed counterpart. However, these differences are always not vivid. Rather than offering precise indications of tissue density, MRI images primarily manifest varying shades of darkness or lightness. Compounding this issue, the exact pixel values within these images are contingent upon both the individual subjects and the specific MRI machines used, rendering them inconsistent across diverse image datasets. Such datasets are crucial for training neural networks or machine learning methods. As a result, the expertise of medical professionals becomes paramount in interpreting these images, posing a significant obstacle to the effective training of classifiers for identifying inflamed soft tissues.

3. Methodology of research

3.1. Method of myocardium segmentation

To overcome the issue formalized in the previous section, specifically to enhance the quality of myocardial inflammation recognition using DL, it is crucial to accurately segment the ROI (in our case, the myocardium segment). The approach proposed by the authors is dedicated to achieving the most precise segmentation of the myocardial region in MRI images.

The approach consists of the sequential use of two DL models: for localization and segmentation. The first model is responsible for the localization of the area where the myocardium is located. The result of the method is a mask with one segment that localizes the area where the myocardium and the left ventricle are located. After a segment of the myocardial region is found, the image is cropped to include only the myocardium with a small frame. The second DL model is responsible for a more accurate search for a segment of the myocardium without a left ventricle. Accordingly, the result of her work is a mask with a marked segment of the myocardium. Since all images are scaled to the same size before training and testing the models, the final step is to resize the mask to match the original size of the input image.

The main steps of the approach are described below.

1. **Step 1: Localization.**

The first step is responsible for the localization of the image area where the myocardium is located. A DL model for localization uses randomly generated 80/20 training and testing datasets respectively. The fastai library [25] is used for training using a pre-trained network based on the ResNet architecture [26]. It is a popular residual neural network used for object classification and detection.

In this work, we used the ResNet-34 version, which had 34 layers (convolutional, pooling, and fully connected layers).

Before starting training, all images are reduced to a single size with the same number of channels. Experimentally, it is determined that the best results for this step are obtained for an image size of 250×250 pixels.

The training process is described below. First, the optimal learning rate is determined, and then the model is trained at high speed for 10 epochs, after which all model weights are unfrozen. After defrosting, the learning rate is determined again, and another 10 learning epochs are performed with changed values. This approach to training helps to increase the accuracy of the model. Using the predicted mask, images are cropped with a change in the aspect ratio to square and centered according to the found mask. Additionally, it is added a small frame, 15% of the size of the mask.

2. Step 2: Segmentation.

The second step is responsible for predicting the segment of the myocardium. The DL segmentation model uses three datasets – to train the segmentation model, to train a linear regression model (to predict the value of the standard deviation for the Gaussian blur method [27] based on the width of the image), and a set to validate the segmentation model. These sets are randomly generated in a 60/20/20 ratio.

Same as with training a localization model, the images are scaled to the same size. The size of images for training is determined experimentally. The model trained with the image size 60×60 showed the best statistical results.

The process of training a segmentation model is similar to training a localization model and includes determining the optimal learning rate, high-speed learning, unfreezing the model weights, re-determining the learning rate, and training with changed values.

Since the model is trained and uses resizing images, in step 2 it is necessary to return the image to its original size. However, resizing usually results in loss of detail and the appearance of unwanted artifacts, which in turn affects the final evaluation of the result.

To solve the described problem, the use of blurring techniques is proposed, aimed at smoothing transitions between pixels and creating a more natural and smooth effect when resizing images. Commonly used blurring methods include Gaussian blur, Box Blur, Median Filter, Bilateral Filter, Mean Shift Filter, and Laplacian of Gaussian Filter.

Box Blur is a simple method that uses the average pixel values in a local area. While this method is computationally efficient, it results in less natural blurring and can lead to loss of details in the image. Median Filter replaces each pixel value with the median value in its neighborhood, helping remove impulse noise but not always providing a smooth transition between different areas. Bilateral Filter considers pixel intensity and spatial proximity, preserving edges, but it can be computationally expensive. Mean Shift Filter shifts each pixel towards the mode of pixel distribution in its local area, ensuring smooth transitions between different parts of the image. Laplacian of Gaussian Filter applies Gaussian blur and then detects edges in the image.

Compared to other methods, Gaussian blur stands out for its optimal balance between result quality and computational complexity. It provides effective image blurring while preserving natural transitions and details. Although other methods may have their advantages in specific scenarios, Gaussian blur remains one of the most versatile and efficient approaches in image processing tasks. Therefore, this method is recommended for use. The following formalism is proposed:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (1)$$

where $G(x, y)$ – the value of the Gaussian filter at a point (x, y) , σ – standard deviation, which determines the degree of blurring, (x, y) – coordinates of a pixel in the image.

It is worth noting that σ is a Gaussian blur parameter that affects the degree of blurring. The choice of σ is important to achieve the best resizing results. Linear regression was used to automatically select an acceptable value. For the prepared dataset, the model was trained as follows. The myocardial mask is defined, and the size of the mask is increased to the original one with σ values from 0 to 1 in steps of 0.1. The size of the image and the best σ value for it are used to train the model. The width of the image is used as an input parameter for the regression model, and the value of σ is used as the output parameter. The trained model can predict the optimal value of σ for any new image size. This approach allows to

automate the selection of the σ parameter for Gaussian blur when resizing the image, which helps preserve the quality and details of the mask and helps improve the final results.

The final step is to map the mask to the input, non-localized image.

The general scheme of the approach is shown in Fig. 1.

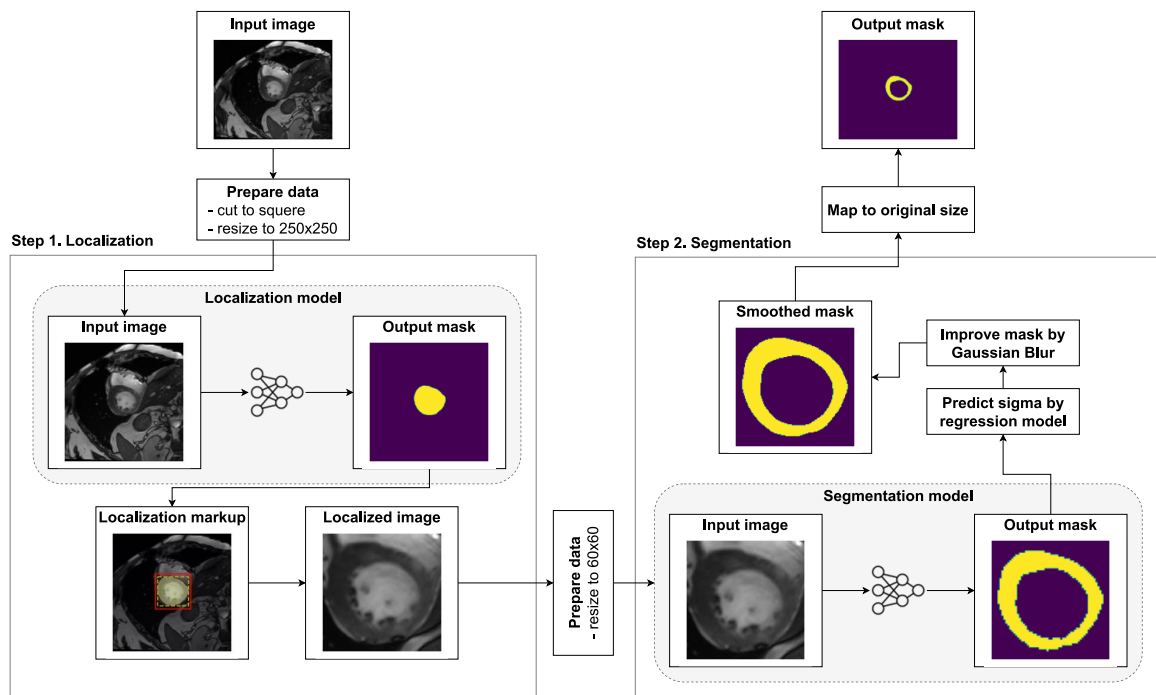


Figure 1: The general scheme of the approach.

3.2. Evaluation of the quality of the obtained results

To validate the proposed approach, namely, to evaluate the similarity between two segmentation masks, it is proposed to use the Dice coefficient [28]. It is a statistical measure that checks the similarity of two sets. It is often used to compare segmentation results with predicted and true masks. It is defined as follows:

$$Dice = \frac{2 \times |A \cap B|}{|A| + |B|}, \quad (2)$$

where A – the set of pixels that belong to the predicted segmentation, B – the set of pixels that belongs to the true segmentation, $|A|$ – the number of elements in the set A , $|B|$ – the number of elements in the set B , $|A \cap B|$ – the number of elements that belong to the set A , and the set B (this is the area of overlap between the predicted and true masks).

The Dice coefficient can take values from 0 to 1, so the results were interpreted as follows:

- *Dice coefficient* = 0.0: indicates no shared pixels between segments, which means no similarity between predicted and true segmentation.
- *Dice coefficient* = 1.0: indicates full identity between segments, meaning perfect agreement between predicted and true segmentation.

Using the Dice coefficient allows objective assessment of the segmentation accuracy, so, to determine how accurately the segmentation results correspond to the true masks.

3.3. Dataset

A modified ACDC dataset [29] was used to train and test the proposed DL models. The set consists of 150 patients and contains five groups: 30 healthy patients; 30 patients with a previous myocardial infarction; 30 patients with dilated cardiomyopathy; 30 patients with hypertrophic cardiomyopathy; 30 patients with an anomaly of the right ventricle.

Data about each patient includes physical parameters, a set of images, and true masks (ground truth label) at different stages of the cardiac cycle. True masks contain marked segments for the myocardium, left and right ventricles. To solve the given problem, two new datasets with certain modifications were created based on the original dataset.

The first dataset is used to train and test the localization model. Only the masks are modified in this dataset. The right ventricular mask was removed and the myocardial and left ventricular masks were combined. These modifications are illustrated in Fig. 2.

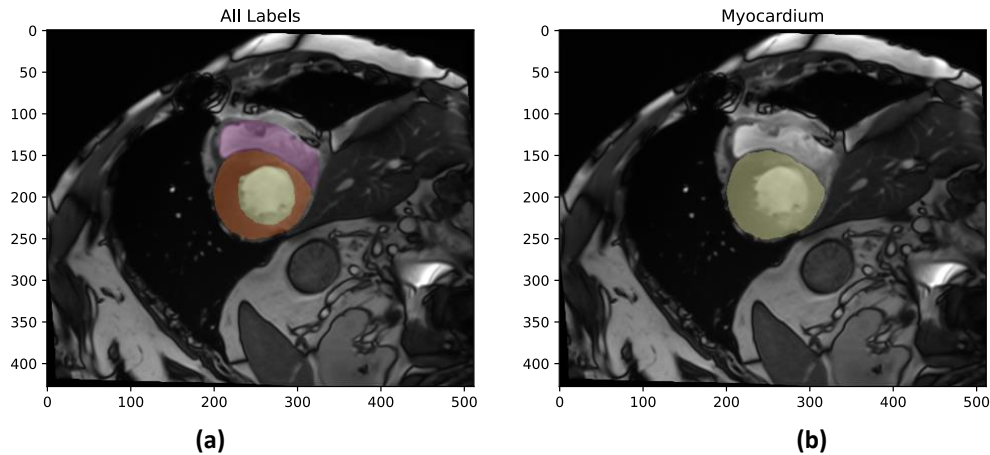


Figure 2: Modification of the mask for the first dataset: (a) with all labels and (b) myocardium area.

The second dataset is used to train the segmentation model. Three modifications were applied to this dataset:

1. Removed masks of segments of the right and left ventricles.
2. Improved the contours of the masks denoting the segment of the myocardium.
3. The dataset is filtered by several criteria.

Below, we will consider and illustrate the applied modifications.

3.3.1. Removing masks of segments of the right and left ventricles

Left and right ventricular segment masks are not used during training, so this information is redundant. This modification is aimed at reducing information noise, which has a positive effect on the quality of education. Also important is that reducing the classes for training improves the learning speed and performance of the model. Fig. 3 shows the appearance of the original mask and the modified mask with only the myocardial segment.

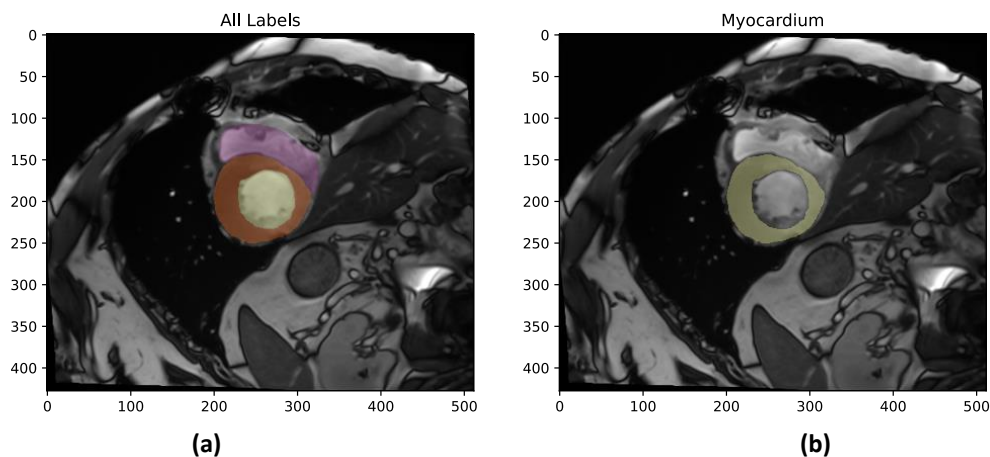


Figure 3: Removing masks of segments of the (a) right and (b) left ventricles.

3.3.2. Improving the contours of the masks denoting the segment of the myocardium

During the training and testing of the models using the original dataset, it was found that some masks have minor inaccuracies. Even a small difference during training and testing can significantly affect the evaluation of the result. These inaccuracies can be either an accidentally marked part of the image or a missed part of the myocardium.

For example, in Fig. 4, it can be seen that an extra area was detected in the upper part of the left ventricle, which was marked as part of the myocardium in the original dataset (A). However, in the edited set, this area was removed. Also, in the right part of the myocardium, an area was found, on which part of the muscle tissue was not indicated in the original dataset, but in the updated set this area was marked (B).

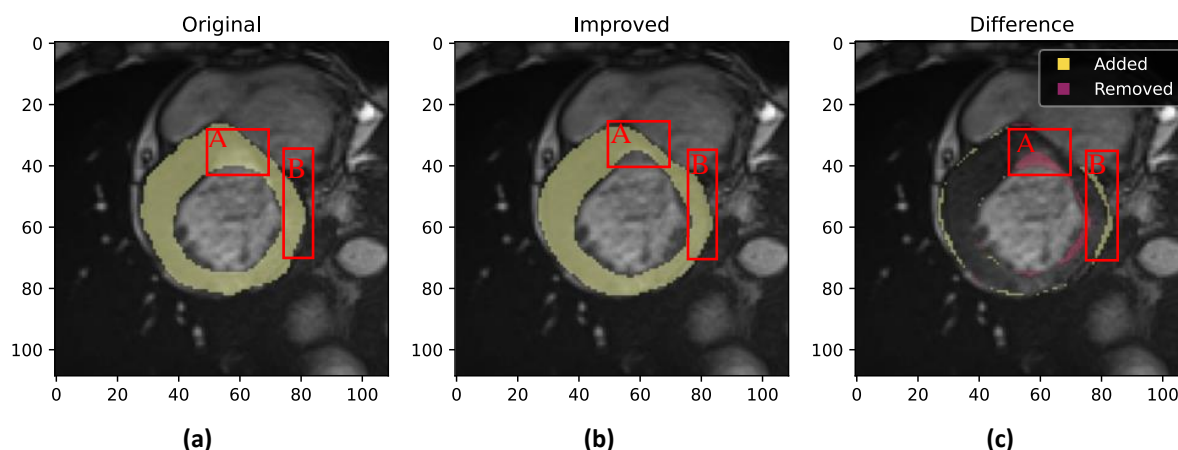


Figure 4: Presentation of (a) original mask and (b) modified mask: A – removed part of the mask, B – added part of the mask. Part (c) shows the difference between the original and predicted masks.

To validate the updated set of images, we developed a tool with an interface shown in Fig. 5.

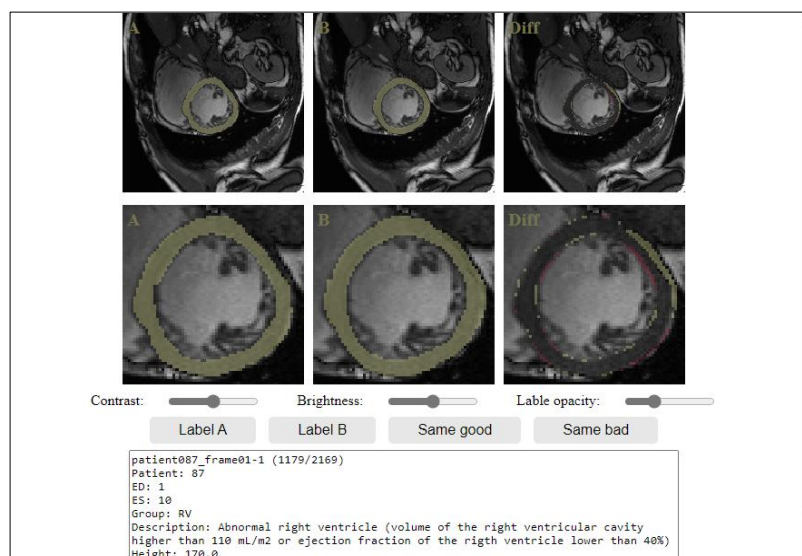


Figure 5: A tool for validating the updated masks.

The developed application lets users analyze images and masks in their original size or localized area, adjust the contrast and brightness to facilitate the analysis of darkened or lightened areas of the image, adjust the transparency of the masks to accurately determine the overlap, and view information about patients and their diagnoses. Users can choose one of four image evaluation options: better mask

A, better mask B, same good, or same bad. Using the application, the practicing cardiologist formed the final dataset.

3.3.3. Filtrating by several criteria

Some images have poor resolution or poor quality in general, which negatively affects the performance of the model. In addition, the original dataset is not focused only on the myocardium, so some images that, for example, made sense with right ventricular segments, may not have any value in the study of the myocardium. The dataset was filtered according to the following criteria:

- The myocardium is partially or completely not visualized.

In some images, the segment of the myocardium is not visualized or is partially visualized. This is mostly because the cut was performed incorrectly, for example, at the moment of opening the mitral or tricuspid valve. If the myocardium in the image is not visualized clearly enough, this can lead to insufficient or even wrong information for training the model. For example, in Fig. 6, you can see that the lower part of the myocardium is poorly visualized, but it is marked on the mask.

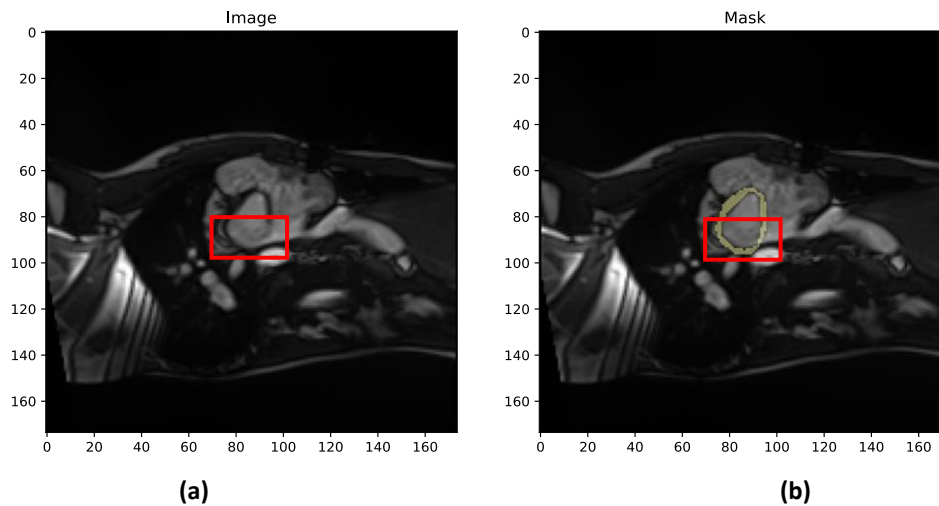


Figure 6: The myocardium is partially visualized in (a) the original image and (b) the image with the segmented myocardium.

Such masks can cause teaching the network to supplement the mask for such cases. This can lead to false segmentations, so it is better to remove such images from the training dataset.

- The myocardial segment is too small.

The original dataset contains images with the myocardial segment starting from 9×9 pixels. This size is not enough to provide quality training. Usually, the reason for such cases is that the picture was taken at a certain period of the cardio cycle when the myocardium is maximally shortened. Such images stand out from the general dataset, and this leads to confusion of the model. In such cases, there is the loss of details and the difficulty of determining features and characteristics for model training. For the final dataset, it was used images with the size of the myocardial segment exceeding 30×30 pixels.

- The left ventricle is partially or completely not visualized.

In some images, the segment of the left ventricle is not visualized or visualized partially. This is mostly because the cut was made closer to the top of the heart, so the walls look thicker, and the cavity of the left ventricle is not visualized. It can also be affected by certain pathologies, for example, hypertrophic cardiomyopathy.

The presence of such images in the training dataset leads to incorrect training of the model. This has been noticed that it causes mistakes in the recognition of myocardium in cases with cardiomyopathy, trabecularity, or spongy myocardium. At the same time, the number of such images in the original dataset is too small for training correct segmentation. The same criterion includes images that do not show the contours of the left ventricle clearly enough.

- Image with insufficient or excessive brightness level.

Some images or their areas are darkened or too bright. Images that are too dark or too bright can affect the quality of model training. Areas with an insufficient or excessive level of brightness lead to the loss of important details and, as a result, negatively affect the quality of learning.

As a result of the conducted filtration, 734 images were removed. The number of patients and their groups corresponds to the original dataset, as only specific images for each patient were removed according to the criteria described above. Therefore, the generated dataset contains 2169 images of 150 patients, which can be divided into 5 groups of 30.

The modifications made to the original dataset resulted in a significant improvement in the quality and efficiency of the myocardial segmentation model. Removing the masks of the right and left ventricles, improving the contours of the myocardial masks, and filtering images based on a set of criteria contributed to reducing informational noise and increasing segmentation accuracy.

3.4. Image size adjustment and data augmentation

Correct presentation of input images is important for successful training of neural network models. The more accurately the training sample approximates the general set of images (that will be input to the system), the higher the maximum achievable quality of the result will be. For the neural network architecture used in the research and the problem under consideration, it is important to investigate the effect of data augmentation and the effect of the size of the input image on the result.

For a neural network with the ResNet architecture, it is important that all images are the same size and have the same number of channels. Therefore, before training (and subsequently before identification), the images are reduced to a single size. The choice of size is an important step, as it directly affects the learning results. Too small a size may cause excessive loss of detail when resizing large images, which is obviously unacceptable in medical image analysis. However, this will significantly speed up the training time of the model and the overall performance of the model. On the contrary, if the images are too large, it will negatively affect the speed of the model, and when stretching small images, it may cause artifacts to appear.

The research conducted an experiment to determine the optimal value for image size. The average Dice coefficient value, calculated based on 10 runs of model training and testing using different image sizes, was used to determine the optimal value.

For the localization stage, the experiment was conducted using images ranging from 100 to 325 pixels in size. The experiment results are shown in Fig. 7. As evident from the graph, smaller-sized images exhibited poorer results, and as the image size increased, the model's efficiency improved. This is because significant compression of image size leads to the loss of image details. However, after increasing the size beyond 250 pixels, a decrease in model efficiency was observed. This happened because most images were significantly stretched to fit this size, causing excessive deformation.

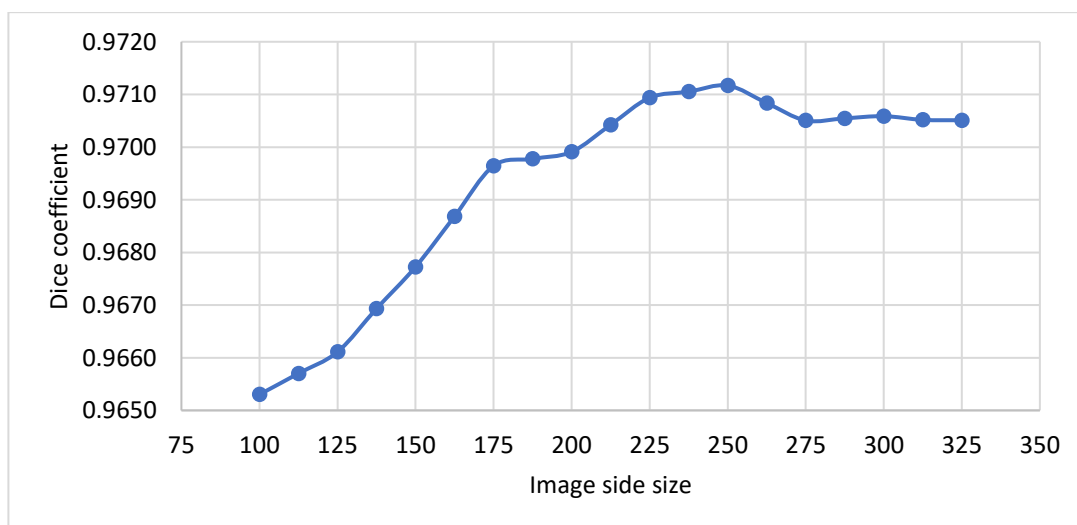


Figure 7: Dependence of results on input image size for the localization stage.

For the segmentation stage, a similar experiment was conducted where images ranging from 30 to 100 pixels were analyzed. The experiment results are depicted in Fig. 8. Similar to the previous plot, a gradual increase in the Dice coefficient can be observed with the increase in image size, followed by a deterioration in results after a certain threshold. From the graph, it can be seen that the highest Dice coefficient value corresponds to the 60x60 size.

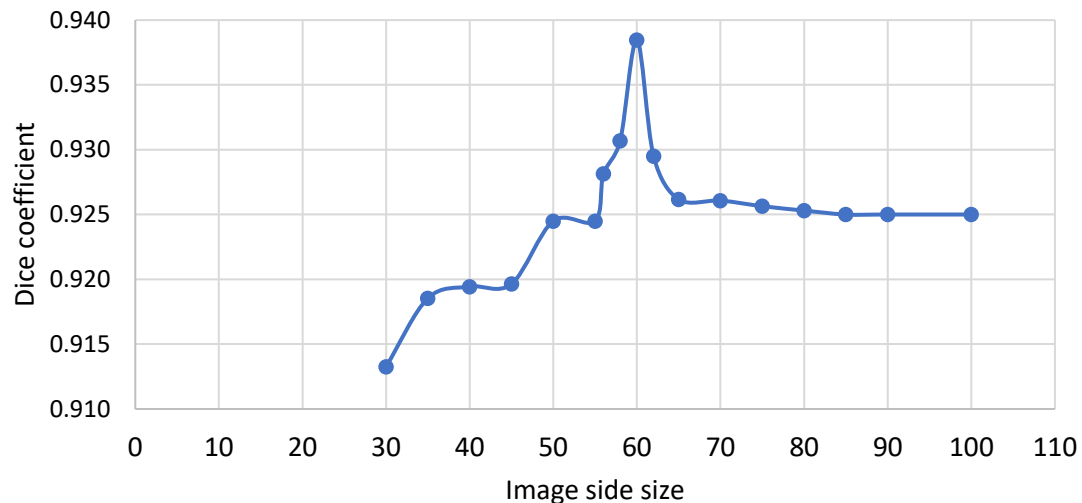


Figure 8: Dependence of results on input image size for the segmentation stage.

As can be seen from the graph, images of smaller size showed a worse result, and along with the increase in image size, the efficiency of the model increased as well. At the same time, after increasing the size by more than 60 pixels it appears a degradation of the model's performance, which is caused because most of the images were significantly stretched to fit this size, so they were too deformed.

In addition, it was conducted an experiment with the introduction of data augmentation to improve the model's effectiveness. Namely, it was added a change of brightness and contrast with a random value within a certain range. The results showed that by increasing the intensity of adding brightness and contrast, the quality of learning became worse (Fig. 9).

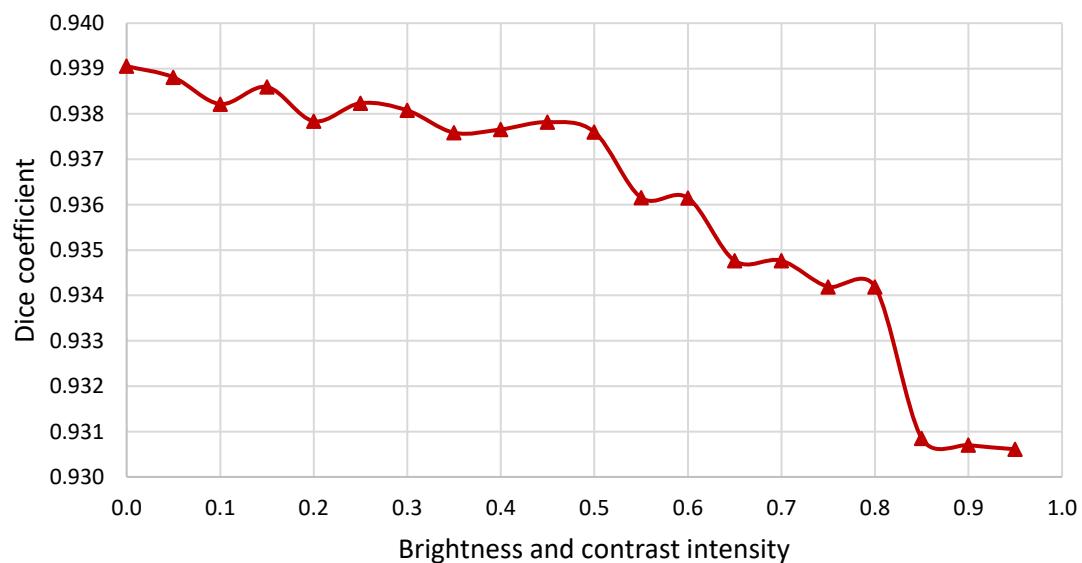


Figure 9: Dependence of results on brightness and contrast change intensity.

Due to the ineffectiveness of this approach, it was not included in the final model.

4. Results and discussion

To evaluate the capability of the approach, it was performed a series of tests using the arithmetic average value of the Dice coefficient calculated from 10 runs of training and testing the model. To ensure the objectivity of the results, the test and training datasets were formed randomly. This made it possible to avoid distortions due to random selection of data and ensure reproducibility of results. Several approaches were tested and compared, and the results are presented in Table 1.

Table 1
Comparison of results of different approaches.

| Approach | Avr Dice | Min Dice | Max Dice |
|---|----------|----------|----------|
| Without localization, using all mask labels | 0.889 | 0.139 | 0.970 |
| Without localization, using only myocardium mask labels | 0.894 | 0.148 | 0.973 |
| With localization, using a standardized size | 0.938 | 0.824 | 0.982 |
| With localization, using the original size | 0.948 | 0.825 | 0.984 |
| With localization, using the original size with Gaussian blur | 0.955 | 0.828 | 0.986 |

The first experiment includes the use of the model without localization and using the original masks. The average Dice coefficient in this case was 0.889.

The second experiment tested the model without localization and using modified masks containing only the myocardial label. The average Dice coefficient was 0.894.

The next experiments were conducted on different images, specifically, with localization, using standardized size (Fig. 10(a)), with localization, using the original size (Fig. 10(b)), and with localization, using the original size with Gaussian blur (Fig. 10(c)).

In the third experiment (Fig. 10(a)), the model was tested using standardized-size images with an average Dice coefficient of 0.938. The model scored an accuracy of 0.948 on an instance image, shown in Fig. 10(a).

In the fourth experiment (Fig. 10(b)), the model was tested with a resizing of the image to the original, which led to a decrease in the average result to 0.948. The model scored an accuracy of 0.940 on an example image, shown in Fig. 10(b).

The fifth (last) experiment (Fig. 10(c)) involved resizing the image to its original size followed by Gaussian blurring. This led to an improvement in the results to the level of 0.955. The model scored an accuracy of 0.953 on an example image, illustrated in Fig. 10(c).

It should be noted that the proposed approach to segmentation of the myocardium on MRI images, which sequentially uses two neural networks, may lose its effectiveness when using images of poor quality.

Therefore, it also has its limitations and potential problems, especially when working with low-quality images:

- **Partial visualization of the myocardium:** If the myocardium is partially or completely not visualized in the image, then the model may form an incorrect mask because it is based on a certain difference between the myocardium and the surrounding tissues. This can lead to incorrect segmentation or to the absence of a myocardial segment in the image.
- **Small size of the area where the myocardium is located:** The area where the myocardium is located has a size of up to 30×30 pixels. This can make it difficult to accurately segment the myocardium, as noise or other structures may be highlighted.
- **Insufficient visualization of the left ventricle:** If the left ventricle is partially or completely not visualized in the image, then the model may mistakenly mark areas that do not belong to the myocardium.
- **Lack of brightness or overbrightness:** Images with a lack of brightness or overbrightness may cause the method to have issues correctly segmenting the myocardium.
- **In cases where the patient has certain pathologies (for example, cardiomyopathy, trabecularity, or spongy myocardium),** the quality of contour definition may drop slightly because the dataset does not contain enough of such cases for quality training.

As follows, the use of the proposed approach has its limitations and requires attention to the selection of input data, especially when working with images of low quality or obvious pathological conditions of the myocardium.

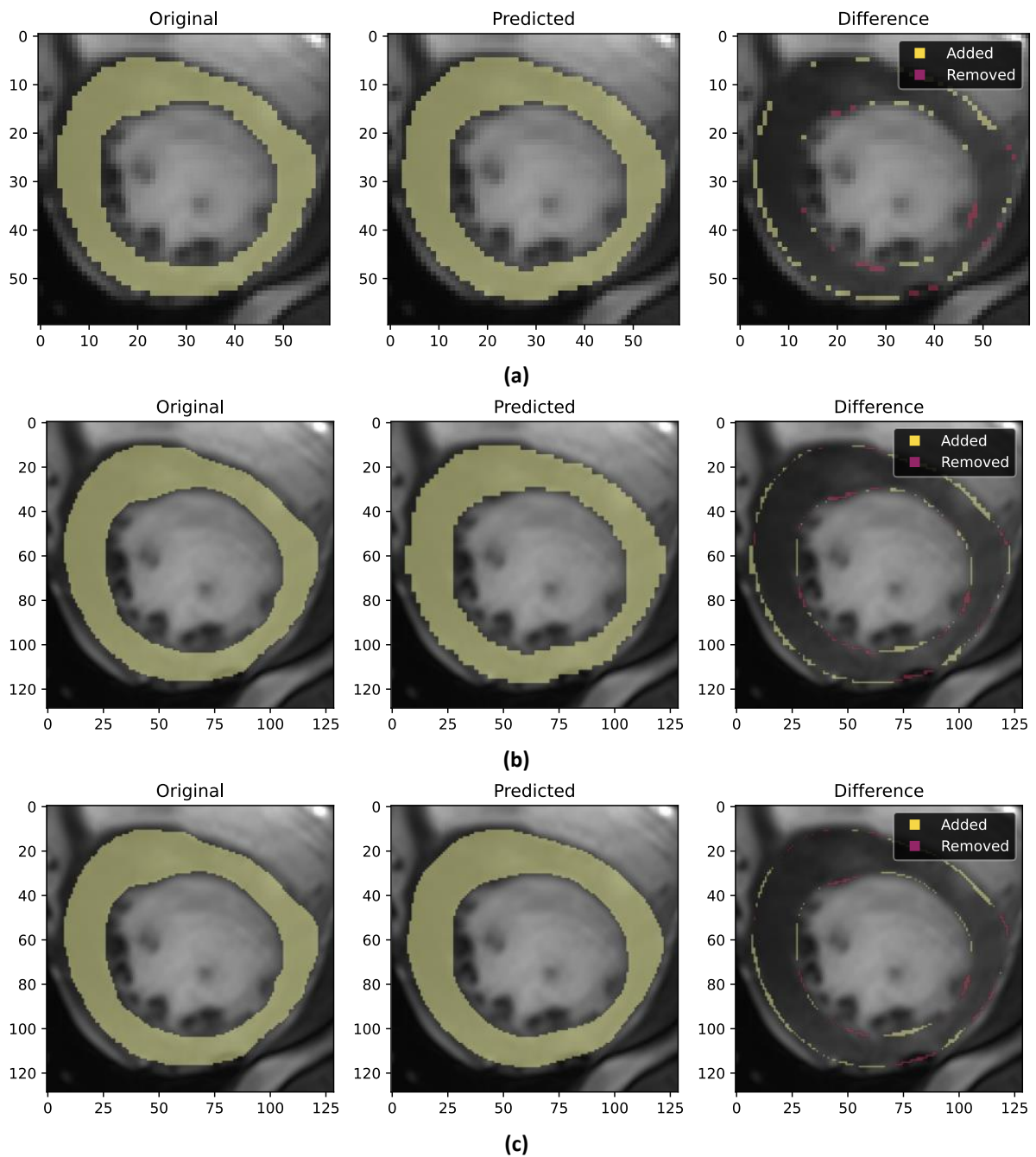


Figure 10: The segmented myocardium with the original and predicted masks obtained through different experiments: (a) Image with localization, using standardized size, (b) image with localization, using original size, and (c) image with localization, using original size with Gaussian blur.

Comparing the proposed approach with other approaches using the original dataset, it can be seen that the results are at a similar level. The comparison of the results can be seen in Table 2.

Table 2

Comparison of results of existing approaches using original database.

| Approach | Dice |
|---|-------|
| With localization, using a standardized size | 0.905 |
| With localization, using the original size with Gaussian blur | 0.915 |
| Baumgartner et al. [22] | 0.901 |
| Isensee et al. [23] | 0.919 |
| Jang et al. [24] | 0.895 |

From the provided table, it can be seen that the proposed two-stage segmentation approach shows a slightly lower result (0.905) compared to the winner of the ACDC segmentation challenge Isensee et al. (0.919). Although the additional improvement in mask accuracy using the Gaussian method enhances the result by one percent (0.915), it is still lower than the best-described approaches. Therefore, while the approach to mask improvement using Gaussian blurring demonstrates increased myocardium segmentation efficiency, the two-stage segmentation method requires further refinement to enhance the final outcome. Meanwhile, the proposed approaches for dataset filtering and myocardium contour refinement with the assistance of a medical professional have shown significant improvement, leading to the model training accuracy of 0.955.

5. Conclusions and Future work

This study explores a strategy for myocardium segmentation in MRI images, utilizing a two-step neural network process followed by mask enhancement. Initially, the first network produces a mask delineating the myocardium and left ventricle areas. Subsequently, the second network refines the segmentation specific to the myocardium. A pivotal element of this approach is the post-processing of the mask using a Gaussian blur, wherein the blur coefficient is automatically selected.

To evaluate the segmentation quality, the Dice coefficient was employed, offering an objective measure of alignment between the predicted and actual masks. Throughout the study, various segmentation strategies were tested and compared. Our experiments revealed that excluding superfluous labels from the masks (on the right and left ventricles) marginally improved the average Dice coefficient from 0.889 to 0.894. However, deploying one model to pinpoint the myocardial region on images and a secondary model to define the precise myocardial contours escalated the coefficient to 0.938. Additionally, the method's efficacy was underscored by the significant enhancement in performance, reaching an average Dice coefficient of 0.955, when using the Gaussian blur method with auto-adjusted blur coefficient during mask resizing.

The limitations of the proposed approach include poor-quality input images, particularly those with inadequate myocardium visualization, limited myocardium areas, suboptimal left ventricle visualization, inappropriate brightness levels, or specific pathologies like cardiomyopathy, trabeculation, or a spongy myocardium.

In conclusion, our novel approach, which integrates preliminary myocardial localization with Gaussian blurring for image size restoration, demonstrates improved segmentation accuracy. This precision augments the MRI image analysis results, particularly in myocardium image classification and identification.

Further research will be directed at improving the accuracy of models in the initial stages – localization and segmentation, and using the obtained results to improve existing solutions that detect areas with inflammatory tissue in cardiac MRI.

6. References

- [1] X. Zhuang et al., Cardiac segmentation on late gadolinium enhancement MRI: A benchmark study from multi-sequence cardiac MR segmentation challenge, *Medical Image Analysis*, vol. 81, p. 102528, 2022, doi:10.1016/j.media.2022.102528.
- [2] T. F. Ismail et al., Cardiac MR: From theory to practice, *Front. Cardiovasc. Med.*, vol. 9, p. 826283, 2022, doi:10.3389/fcvm.2022.826283.
- [3] F. Zabihollahy, S. Rajan, and E. Ukwatta, Machine learning-based segmentation of left ventricular myocardial fibrosis from magnetic resonance imaging, *Curr Cardiol Rep*, vol. 22, no. 8, p. 65, 2020, doi:10.1007/s11886-020-01321-1.
- [4] B. Villard, V. Grau, and E. Zacur, Surface mesh reconstruction from cardiac MRI contours, *Journal of Imaging*, vol. 4, no. 1, Art. no. 1, 2018, doi:10.3390/jimaging4010016.
- [5] S. Wang et al., Deep generative model-based quality control for cardiac MRI segmentation, in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. Part IV*, A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. Racoceanu, and L. Joskowicz, Eds., in *Lecture Notes in Computer Science*, vol. 12264. Lima, Peru, October 4–8, 2020: Springer International Publishing, 2020, pp. 88–97. doi:10.1007/978-3-030-59719-1_9.
- [6] V. Russo, L. Lovato, and G. Ligabue, Cardiac MRI: Technical basis, *Radiol med*, vol. 125, no. 11, pp. 1040–1055, Nov. 2020, doi:10.1007/s11547-020-01282-z.
- [7] G. Simantiris and G. Tziritas, Cardiac MRI segmentation with a dilated CNN incorporating domain-specific constraints, *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 6, pp. 1235–1243, 2020, doi:10.1109/JSTSP.2020.3013351.
- [8] H. Cui, C. Yuwen, L. Jiang, Y. Xia, and Y. Zhang, Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images, *Computer Methods and Programs in Biomedicine*, vol. 206, p. 106142, 2021, doi:10.1016/j.cmpb.2021.106142.
- [9] P. Radiuk, O. Barmak, and I. Krak, An approach to early diagnosis of pneumonia on individual radiographs based on the CNN information technology, *The Open Bioinformatics Journal*, vol. 14, no. 1, pp. 92–105, Jun. 2021, doi:10.2174/1875036202114010093.
- [10] I. Krak, V. Kuznetsov, S. Kondratiuk, L. Azarova, O. Barmak, and P. Radiuk, Analysis of deep learning methods in adaptation to the small data problem solving, in *Lecture Notes in Data Engineering, Computational Intelligence, and Decision Making*, S. Babichev and V. Lytvynenko, Eds., in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 149. Cham: Springer International Publishing, 2023, pp. 333–352. doi:10.1007/978-3-031-16203-9_20.
- [11] R. M. Wehbe et al., Deep learning for cardiovascular imaging: A review, *JAMA Cardiol.*, Sep. 2023, doi:10.1001/jamacardio.2023.3142.
- [12] S. Ciyamala Kushbu and T. M. Inbamalar, Interactive one way contour initialization for cardiac left ventricle and right ventricle segmentation using hybrid method, *Journal of Medical Imaging and Health Informatics*, vol. 11, no. 4, pp. 1037–1054, 2021, doi:10.1166/jmihi.2021.3562.
- [13] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, U-Net and its variants for medical image segmentation: A review of theory and applications, *IEEE Access*, vol. 9, pp. 82031–82057, 2021, doi:10.1109/ACCESS.2021.3086020.
- [14] Z. Chen, J. Bai, and Y. Lu, Dilated convolution network with edge fusion block and directional feature maps for cardiac MRI segmentation, *Front. Physiol.*, vol. 14, p. 1027076, 2023, doi:10.3389/fphys.2023.1027076.
- [15] Y. Dang, D. Anand, and A. Sethi, Pixel-wise segmentation of right ventricle of heart, in *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON-2019)*, Kochi, India, 17-20 October 2019: IEEE Inc., Oct. 2019, pp. 1797–1802. doi:10.1109/TENCON.2019.8929229.
- [16] I. Krak, O. Barmak, and P. Radiuk, Detection of early pneumonia on individual CT scans with dilated convolutions, in *Proceedings of the 2nd International Workshop on Intelligent Information Technologies & Systems of Information Security (IntelITSIS-2021)*, T. Hovorushchenko, O. Savenko, P. Popov, and S. Lysenko, Eds., Khmelnytskyi, Ukraine, March 24–26, 2021: CEUR-WS.org, 2021, pp. 214–227. Accessed: May 09, 2021. [Online]. Available: <http://ceur-ws.org/Vol-2853/>

- [17] Y. Luo, L. Xu, and L. Qi, A cascaded FC-DenseNet and level set method (FCDL) for fully automatic segmentation of the right ventricle in cardiac MRI, *Med Biol Eng Comput*, vol. 59, no. 3, pp. 561–574, Mar. 2021, doi:10.1007/s11517-020-02305-7.
- [18] M. Lin, M. Jiang, M. Zhao, E. Ukwatta, J. A. White, and B. Chiu, Cascaded triplanar autoencoder M-Net for fully automatic segmentation of left ventricle myocardial scar from three-dimensional late gadolinium-enhanced MR images, *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2582–2593, 2022, doi:10.1109/JBHI.2022.3146013.
- [19] P. Radiuk, O. Kovalchuk, V. Slobodzian, E. Manziuk, and I. Krak, Human-in-the-loop approach based on MRI and ECG for healthcare diagnosis, in *Proceedings of the 5th International Conference on Informatics & Data-Driven Medicine*, Lyon, France, 18-20 November: CEUR-WS.org, 2022, pp. 9–20. [Online]. Available: <https://ceur-ws.org/Vol-3302/paper1.pdf>
- [20] D. M. Yalcinkaya et al., Temporal uncertainty localization to enable human-in-the-loop analysis of dynamic contrast-enhanced cardiac MRI datasets. *arXiv*, 2023. doi:10.48550/arXiv.2308.13488.
- [21] L. Li et al., MyoPS: A benchmark of myocardial pathology segmentation combining three-sequence cardiac magnetic resonance images, *Medical Image Analysis*, vol. 87, p. 102808, 2023, doi:10.1016/j.media.2023.102808.
- [22] C. F. Baumgartner, L. M. Koch, M. Pollefeys, and E. Konukoglu, An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation, in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, M. Pop, M. Sermesant, P.-M. Jodoin, A. Lalande, X. Zhuang, G. Yang, A. Young, and O. Bernard, Eds., in *Lecture Notes in Computer Science*, vol. 10663. Cham: Springer International Publishing, 2018, pp. 111–119. doi:10.1007/978-3-319-75541-0_12.
- [23] F. Isensee, P. F. Jaeger, P. M. Full, I. Wolf, S. Engelhardt, and K. H. Maier-Hein, Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features, in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, M. Pop, M. Sermesant, P.-M. Jodoin, A. Lalande, X. Zhuang, G. Yang, A. Young, and O. Bernard, Eds., in *Lecture Notes in Computer Science*, vol. 10663. Cham: Springer International Publishing, 2018, pp. 120–129. doi:10.1007/978-3-319-75541-0_13.
- [24] Y. Jang, Y. Hong, S. Ha, S. Kim, and H.-J. Chang, Automatic segmentation of LV and RV in cardiac MRI, in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, M. Pop, M. Sermesant, P.-M. Jodoin, A. Lalande, X. Zhuang, G. Yang, A. Young, and O. Bernard, Eds., in *Lecture Notes in Computer Science*, vol. 10663. Cham: Springer International Publishing, 2018, pp. 161–169. doi:10.1007/978-3-319-75541-0_17.
- [25] J. Howard and S. Gugger, Fastai: A layered API for deep learning, *Information*, vol. 11, no. 2, Art. no. 2, 2020, doi:10.3390/info11020108.
- [26] R. Wightman, H. Touvron, and H. Jégou, ResNet strikes back: An improved training procedure in timm. *arXiv*, 2021. doi:10.48550/arXiv.2110.00476.
- [27] E. S. Gedraite and M. Hadad, Investigation on the effect of a Gaussian blur in image filtering and segmentation, in *Proceedings ELMAR-2011, Zadar, Croatia, 14-16 September 2011: IEEE Inc.*, 2011, pp. 393–396. Accessed: Sep. 29, 2023. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/6044249>
- [28] A. C. Ogier, A. Bustin, H. Cochet, J. Schwitter, and R. B. van Heeswijk, The road toward reproducibility of parametric mapping of the heart: A technical review, *Front. Cardiovasc. Med.*, vol. 9, p. 876475, 2022, doi:10.3389/fcvm.2022.876475.
- [29] O. Bernard et al., Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the Problem Solved?, *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2514–2525, Nov. 2018, doi:10.1109/TMI.2018.2837502.