CYBERNETICS AND PHYSICS, VOL. 14, NO. 1, 2025, 62–66

DEEP LEARNING METHODS FOR RIGHT VENTRICLE SEGMENTATION IN RADIONUCLIDE IMAGING

Petr Larochkin
Saint-Petersburg State University
Russia
plarochkin@gmail.com

Elena Kotina
Saint-Petersburg State University
Russia
e.kotina@spbu.ru

Dmitry Girdyuk
Saint-Petersburg State University
Russia
d.girdyuk@spbu.ru

Evgeny Ostroumov
N.V. Sklifosovsky Research Institute
for Emergency Medicine
Moscow, Russia
ostroumoven@sklif.mos.ru

Article history:
Received 22.05.2025, Accepted 21.06.2025

Abstract

This study presents a deep learning-based approach for the automatic segmentation of the right ventricle (RV) in gated myocardial perfusion SPECT images (gSPECT). Unlike the left ventricle (LV), the RV poses significant segmentation challenges due to its complex anatomy, thinner walls, and lower perfusion. We manually annotated 384 SPECT volumes and propose the use of the ResUNetSE3D neural network, incorporating both anatomical and phase imaging data to enhance segmentation accuracy. The model achieved a Dice coefficient of 0.8272 and a Jaccard index of 0.7086. These results demonstrate the feasibility of fully automated RV segmentation, laying the groundwork for future clinical applications in quantitative cardiac assessment.

Key words

Gated single photon emission computed tomography (gSPECT), mathematical modeling, myocardial perfusion imaging, deep learning, image segmentation

1 Introduction

Gated single-photon emission computed tomography (gSPECT) plays a crucial role in the diagnosis of cardiovascular disease by providing detailed information on myocardial perfusion and function [Ostroumov et al., 2023]. This imaging modality is widely used for detecting ischemic heart disease, assessing myocardial viability, and stratifying patient risk. However, the complexity of SPECT image processing necessitates the development of specialized methods. Recently, deep learning techniques have demonstrated high effectiveness in medical imaging tasks, enabling automated data analysis and improving diagnostic accuracy [Miller et al., 2025; Schmidt et al., 2023].

Segmentation of the right ventricle (RV) in radionuclide images remains a challenging task in cardiology. Unlike the left ventricle (LV), the RV is not always visulized on nuclear imaging or is not visulized clearly and also has a more complex structure, making it difficult to process [Ostroumov et al., 2023]. Most existing algorithms are tailored to LV analysis [Ploskikh and Kotina, 2021; Kotina, 2022; Germano et al., 2016; Germano and Slomka, 2019], while RV segmentation remains significantly less explored. In addition, RV analysis is typically not included in standard image processing software packages.

The extraction of information about the RV from nuclear tomoscintigraphy has long attracted the attention of researchers [Kotina et al., 2012; Farag et al., 2019; Ostroumov et al., 2015]. A semi-automatic approach to the quantitative analysis of RV myocardium was used in software "Karfi" was developed on the base of SPbU [Kotina et al., 2014].

Some articles discuss the construction of semi-automatic methods for non-synchronized myocardial perfusion tomoscintigraphy studies [Entezarmahdi et al., 2023], in [Zhao et al., 2023], automatic construction of RV contours using studies with 8 synchronization intervals is considered. The present work proposes a deep learning-based approach for fully automatic RV segmentation in gated SPECT images.

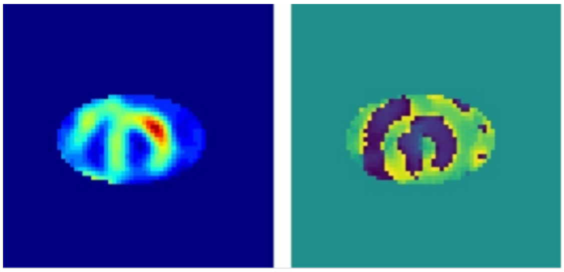


Figure 1. Representative slices from an original gated SPECT image (left) and the corresponding phase image (right)

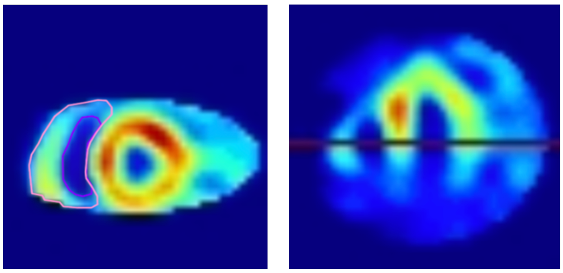


Figure 2. Contours annotated in CVAT. Left: epicardial and endocardial boundaries of the right ventricle on a coronal slice. Right: a line on the transverse slice indicates the correspondence between the coronal and transverse planes.

2 Problem Statement

Automatic segmentation of the right ventricular (RV) myocardium in gated SPECT images remains a challenging task [Ostroumov et al., 2023] due to the typically thinner RV wall, lower perfusion compared to the left ventricle (LV), and more complex anatomical structure. The advancement of artificial intelligence (AI) methods offers new opportunities for the automatic segmentation of medical images, including RV myocardium delineation.

This study focuses on gated myocardial perfusion SPECT acquisitions with 16 synchronization intervals. The primary objective is to develop a neural network model for the automatic segmentation of the RV myocardium. The task involves both data annotation and model training. We also explore the inclusion of a phase parametric channel in the input data to enhance segmentation performance. A series of experiments with varying parameters is conducted, and the results are evaluated using selected performance metrics.

3 Data

DICOM (Digital Imaging and Communications in Medicine) is a widely used standard that facilitates the exchange of medical images (such as SPECT, PET, CT, and MRI) and associated data across different systems. These files, typically with a `.dcm` extension, contain both the image and a standardized header with tags that store key information such as patient demographics and scan settings.

The dataset comprises studies, each consisting of N three-dimensional volumes corresponding to N temporal frames of a representative cardiac cycle acquired using gated myocardial perfusion SPECT imaging [Kotina, 2022]:

$$P_1(i,j,k),\ldots,P_N(i,j,k), i=\overline{1,n},\ j=\overline{1,m},\ k=\overline{1,s}.$$

In this study, we consider $N = 16$ time intervals.

In addition to the arrays mentioned above, a phase array $\Phi(i,j,k)$ (or phase image) is also considered, which is constructed as follows:

$$\Phi(i,j,k) = \frac{180}{\pi}\arctan\left(\frac{b_1(i,j,k)}{a_1(i,j,k)}\right), \quad (1)$$

where

$$a_1(i,j,k) = \frac{2}{N}\sum_{s=1}^{N}\left(P_s(i,j,k)\cos\left(\frac{2\pi(s-1)}{N}\right)\right), \quad (2)$$

$$b_1(i,j,k) = \frac{2}{N}\sum_{s=1}^{N}\left(P_s(i,j,k)\sin\left(\frac{2\pi(s-1)}{N}\right)\right). \quad (3)$$

The phase three-dimensional images, calculated from the source data, provide additional information about the RV myocardium.

Figure 1 presents coronal slice from an original gated SPECT image and its corresponding phase image.

Image annotation was performed manually using the CVAT platform [Corporation, 2022]. For each slice, two masks were created: an outer mask corresponding to the epicardial contour of the right ventricular (RV) myocardium and an inner mask outlining the RV cavity (see Figure 2).

During the annotation process, an inter-slice interpolation and tracking system was utilized, reducing the need for manual labeling on individual slices by propagating annotations across frames automatically.

The final dataset consists of 384 three-dimensional SPECT myocardial perfusion images. Each slice has a resolution of 64×64 pixels, and the number of slices per volume ranges from 20 to 60. For simplicity, we assume that each volume is represented as a $64 \times 64 \times 64$ array. If the original scan contains fewer slices, the volume is zero-padded accordingly to match the target shape.

The dataset was partitioned into three subsets for training, validation, and testing of the model with the following proportions: 75% for training, 12.5% for validation, and 12.5% for testing.

4 Model

4.1 ResUNetSE3D Architecture

Having relatively small volume size and dense data representation, we used the fully-convolutional ResUNetSE3D neural network architecture [Toubal et al.,

ID	Experiment Configuration	Dice	MeanIoU
a	All slices: ✗, In: 2, Out: 1	0.8272	0.7086
b	All slices: ✗, In: 1, Out: 1	0.8040	0.6786
c	All slices: ✓, In: 2, Out: 1	0.7179	0.5646
d	All slices: ✓, In: 1, Out: 1	0.7210	0.5701
e	All slices: ✗, In: 2, Out: 2	0.7893	0.6594
f	All slices: ✗, In: 1, Out: 2	0.7932	0.6668
g	All slices: ✓, In: 2, Out: 2	0.6263	0.4669
h	All slices: ✓, In: 1, Out: 2	0.6525	0.4967

Table 1. Evaluation metrics for different experimental configurations

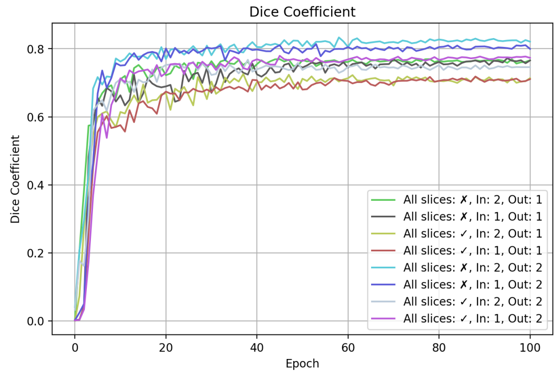


Figure 3. Values of the Dice coefficient on the validation set for different configurations during training process

2020; Wolny, 2020]. Its structure is based on the original UNet [Ronneberger et al., 2015] with encoder-decoder architecture and skip connections. Both encoder and decoder consist of four (32, 64, 128 and 256 channels size) double 3D-convolutional layers ($3 \times 3 \times 3$ kernel, 3D batch normalization and LeackyReLU activation) with residual connections followed by ”Squeeze and Excitation” blocks [Hu et al., 2017]. The decoder block uses transposed convolution for upsampling.

4.2 Training Configurations

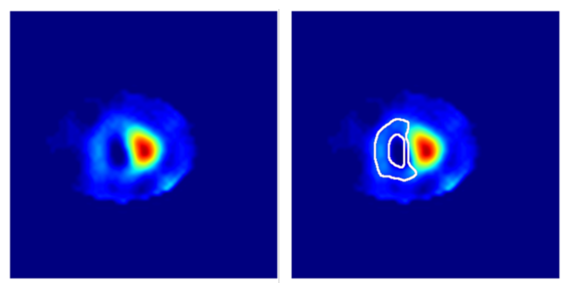
The experiments were performed using the ReSUNetSE3D model to process three-dimensional images. The input to the model consisted of either one or two channels; in the latter case, a phase image was included as an additional input. Similarly, the model produced either one or two output channels. When two outputs were used, the model generated two masks: an outer mask encompassing the entire right ventricle and an inner mask delineating the right ventricular cavity. Depending on the configuration, either all tomoscintigraphy slices or only those containing the right ventricle were used for training. The loss function employed was the BCE Dice Boundary loss [Kervadec et al., 2021], with weight coefficients of 0.5, 0.25, and 0.25, respectively. Optimization was carried out using the AdamW optimizer with a learning rate of 1×10^{-3} and a weight decay of 1×10^{-5}. The learning rate was reduced by a factor of 0.5 every 20 epochs. Training was conducted for 100 epochs with a batch size of 16.

Figure 3 shows the segmentation quality (Dice coefficient) on the validation set during model training process. These metrics were calculated for different model configurations at each training epoch. The following notations are used in the figures: ”All slices: ✓or ✗” corresponds to the ”layer type” parameter in the configuration; ”In: 1, 2” indicates the use or absence of the phase channel; ”Out: 1, 2” represents the number of output channels (whether the mask was predicted as a whole or as a combination of the outer and inner masks).

5 Results

The results presented in Table 5 demonstrate the impact of different model configurations on segmentation quality. The configuration “All slices: ✗, In: 2, Out: 1” showed the highest metric values: Dice is 0.8272 and Jaccard is 0.7086. This may indicate the potential benefit of using two input channels to improve segmentation accuracy. The second result was achieved with the configuration “All slices: ✗, In: 1, Out: 1”, where the Dice score reached 0.8040 and the Jaccard index was 0.6786.

Several evaluated configurations relied on the assumption that segmentation would be performed only on slices containing the right ventricle, rather than on the entire volume. Although this approach introduces a manual preprocessing step, the identification of relevant slices is relatively simple and has been found to improve segmentation quality.



a b

Figure 4. Example of a coronal slice from the original image (a) with the corresponding manually delineated right ventricular contour (b)

Figure 4 presents an example of a coronal slice along with the corresponding contours obtained through manual annotation.

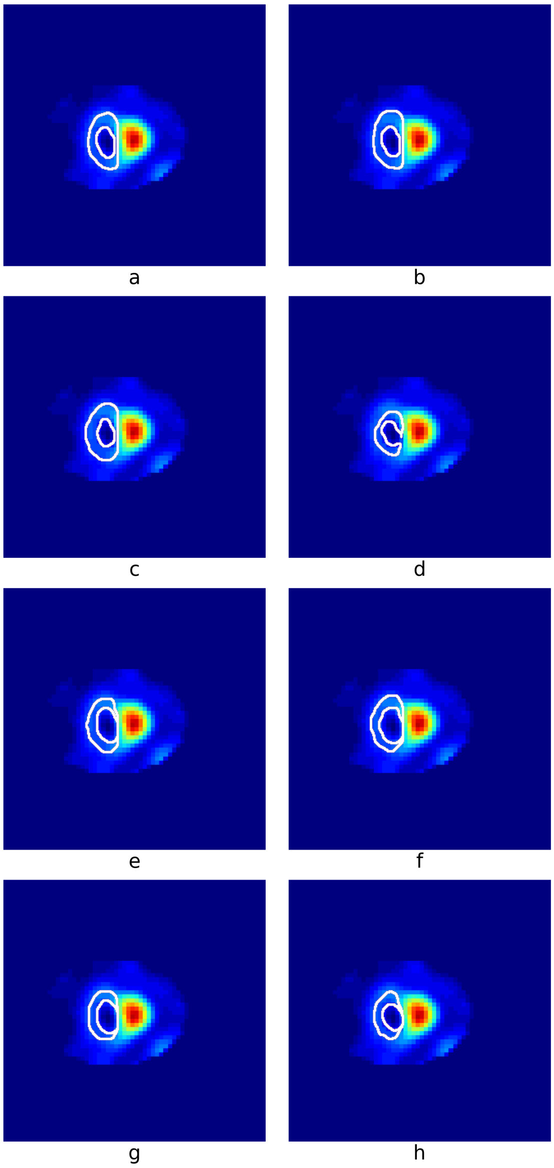


Figure 5. Masks and smoothed contours obtained in different experimental configurations, corresponding to the data presented in Table 1

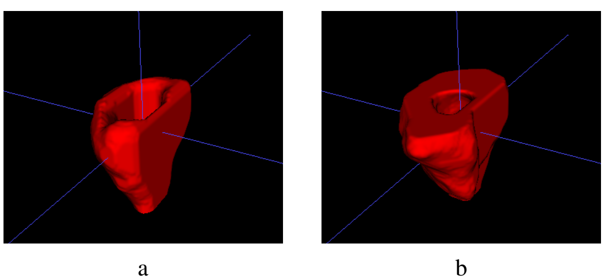


Figure 6. 3D visualization of the segmented region produced by the best-performing model configuration (a) and the manually annotated region (b)

Figure 5 shows the masks and smoothed contours obtained in different experimental configurations, corresponding to the data presented in Table 5.

Figure 6 shows a 3D visualization of the segmented region produced by the best-performing model configuration alongside the corresponding manually annotated region.

6 Conclusion

In this work, we proposed a method for automated segmentation of the right ventricle (RV) of the heart on gated SPECT images with 16 synchronization intervals.

The ResUNetSE3D model achieved a segmentation accuracy of Dice is 0.8272 and Jaccard = 0.7086. This result was obtained using a dual-channel input configuration (original volume and phase volume) and slices containing the right ventricle.

The obtained results lay the foundation for fully automated quantitative assessment of clinically significant parameters of the right ventricle.

While the current approach focuses on segmenting a single three-dimensional volume, an important extension for future research would be to develop a model that incorporates the entire gated study, consisting of all 16 3D volumes. This would allow the model to utilize temporal information across all synchronization intervals, potentially enhancing segmentation performance and enabling more detailed functional characterization of the right ventricle.

References

Corporation, I. (2022). Cvat: Computer vision annotation tool. https://github.com/opencv/cvat. Accessed: 2025-04-30.

Entezarmahdi, S., Faghihi, R., Yazdi, M., et al. (2023). Qcard-nm: Developing a semiautomatic segmentation method for quantitative analysis of the right ventricle in non-gated myocardial perfusion spect imaging. *EJNMMI Physics*, 10:Article 21.

Farag, A., Heo, J., Tauxe, L., et al. (2019). Detection and quantitation of right ventricular reversible perfusion defects by stress spect myocardial perfusion imaging: A proof-of-principle study. *Journal of Nuclear Cardiology*, 26(1):266–271.

Germano, G., Kavanagh, P., Fish, M., et al. (2016). Same-patient processing for multiple cardiac spect studies: Improving lv segmentation accuracy. *Journal of Nuclear Cardiology*, 23(6):1435–1441.

Germano, G. and Slomka, P. (2019). Assessing lv remodeling in nuclear cardiology. *Journal of Nuclear Cardiology*, 26(1):233–235.

Hu, J., Shen, L., Albanie, S., Sun, G., and Wu, E. (2017). Squeeze-and-excitation networks.

Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J., and Ben Ayed, I. (2021). Boundary loss for highly unbalanced segmentation. *Medical Image Analysis*, 67:101851.

Kotina, E. (2022). *Mathematical and Computer Modeling in Nuclear Medicine*. VVM Publishing, St. Petersburg.

Kotina, E., Ostroumov, E., and Ploskikh, V. (2012). Left and right ventricular phase analysis of gated spect myocardial perfusion. *European Journal of Nuclear Medicine and Molecular Imaging*, 39(2):213.

Kotina, E., Ovsyannikov, D., Ostroumov, E., et al. (2014). Indis-karfi: Program for processing functional cardiac studies. Software Certificate No. 2014662434. Registered software certificate.

Miller, R., Kavanagh, P., Lemley, M., et al. (2025). Artificial intelligence-enhanced perfusion scoring improves the diagnostic accuracy of myocardial perfusion imaging. *Journal of Nuclear Medicine*, 66(3).

Ostroumov, E., Kotina, E., Ploskikh, V., et al. (2015). Visualization of acute myocardial infarction of right ventricle free wall after cardiac surgery by gated spect with phase images. *European Journal of Nuclear Medicine and Molecular Imaging*, 42(Suppl. 1):S509.

Ostroumov, E., Migunova, E., Kotina, E., et al. (2023). Visualization of the right ventricle in myocardial perfusion spect before and after revascularization in patients with postinfarction cardiosclerosis. *Transplantology*, 15(2):200–215.

Ploskikh, V. and Kotina, E. (2021). Challenges of gated myocardial perfusion spect processing. *Cybernetics and Physics*, 10(3):171–177.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv e-prints*, page arXiv:1505.04597.

Schmidt, I., Kotina, E., and Buev, P. (2023). Deep learning muscle segmentation model for CT images in DICOM format. *Cybernetics and Physics*, 12(3):201–206.

Toubal, I. E., Duan, Y., and Yang, D. (2020). Deep learning semantic segmentation for high-resolution medical volumes. In *2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–9.

Wolny, A. (2020). 3d u-net model for volumetric semantic segmentation written in pytorch. https://github.com/wolny/pytorch-3dunet/. Accessed: 2025-04-30.

Zhao, C., Shi, S., He, Z., et al. (2023). Spatial-temporal v-net for automatic segmentation and quantification of right ventricle on gated myocardial perfusion spect images. *Medical Physics*, 50(12):7415–7426.