EJNMMI Research

ORIGINAL RESEARCH Open Access

Fully automatic left ventricle segmentation in ^{82}Rb PET/CT Using a semi-supervised nnU-net

Mohammadreza Amirian[1,2†], Arthur Chevalley[1,2†], María Martín Asiain[1,2], Ran Klein[3], Robert DeKemp[4], Eric Moulton[6], John O. Prior[1], Christel H. Kamani[1,5], Mario Jreige[1*†] and Adrien Depeursinge[1,2†]

Abstract

Background Quantification of myocardial blood flow (MBF) with ^{82}Rb PET/CT requires accurate delineation of the left ventricle (LV). Manual or semi-automated contouring remains time-consuming and error-prone, particularly in hypoperfused myocardium. We developed and validated a fully automatic LV segmentation pipeline using nnU-Net applied to ^{82}Rb PET/CT. A manual, multimodal segmentation protocol integrating dynamic PET and CT was established in a single-center cohort of 40 non-gated PET/CT series (20 patients, rest & stress), including challenging cases with extensive necrosis (median 21%). The resulting ground truth masks were used for five-fold cross-validation, and semi-supervised learning incorporated 805 additional unlabeled dynamic PET series (504 patients). Model performance was compared with an optimized semi-automatic thresholding baseline (35% $\mathrm{SUV}_{\max}$).

Results The nnU-Net significantly outperformed the baseline, achieving a mean Dice of 87.8[85.6, 89.2]% vs 75.1[72.9, 76.9]%, recall 89.1[86.1, 91.4]% vs 82.6[79.1, 85.4]%, and precision 88.1[84.2, 90.4]% vs 70.2[67.2, 73.0]%. The improvement was most pronounced in hypoperfused regions, where recall increased by 20–30% compared to thresholding. Semi-supervised learning modestly enhanced model robustness across both rest and stress acquisitions.

Conclusions A deep-learning-based approach enables fully automatic LV segmentation in ^{82}Rb PET/CT with near-expert accuracy. This framework eliminates manual interaction, supports large-scale MBF quantification, and paves the way for reproducible, high-throughput cardiac PET analysis in clinical and research workflows.

Keywords Left ventricle segmentation, ^{82}Rb PET/CT, Deep convolutional neural networks

†Mohammadreza Amirian, Arthur Chevalley, Mario Jreige and Adrien Depeursinge contributed equally to this work.

*Correspondence:
Mario Jreige
mario.jreige@chuv.ch
[1]Department of Nuclear Medicine and Molecular Imaging, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland
[2]Institute of Informatics, HES-SO Valais-Wallis, Sierre, Switzerland
[3]Division of Nuclear Medicine and Molecular Imaging, University of Ottawa, Ottawa, ON, Canada
[4]Division of Cardiology, University of Ottawa Heart Institute, Ottawa, ON, Canada
[5]Department of Cardiology, CHUV, Lausanne, Switzerland
[6]Jubilant DraxImage Inc., Kirkland, QC, Canada

Introduction

^{82}Rb PET/CT is a well-established modality for quantitative myocardial perfusion imaging (MPI), enabling robust assessment of myocardial blood flow (MBF) and coronary flow reserve in patients with suspected or known coronary artery disease (CAD) or coronary microvascular dysfunction (CMD) [1–4]. MBF quantification provides incremental diagnostic and prognostic value over relative perfusion imaging, particularly for detecting multivessel disease, microvascular disease, and for guiding decisions on revascularization [5, 6].

Accurate delineation of the left ventricle (LV) is essential for extracting reliable time–activity curves (TACs) and computing MBF, as segmentation errors propagate through kinetic modeling and may directly affect clinical interpretation and patient management [7–11]. This challenge is accentuated in dynamic acquisitions and in diseased myocardium, where hypoperfusion, thinning, or scar complicate boundary definition.

In routine practice, LV segmentation is performed manually or with semi-automated tools (FlowQuant, Corridor 4DM, QPET). These approaches remain limited by operator dependency, fixed geometric assumptions such as uniform wall thickness, and reduced performance in heterogeneous tracer distributions. Such limitations hinder reproducibility, slow clinical workflow, and restrict the feasibility of large-scale MBF studies [12–14].

Deep learning has transformed cardiac segmentation in MRI and CT [15–18], and more recently in SPECT and PET using tracers such as ^{13}N-ammonia, ^{18}F-FDG, or ^{18}F-FPTP [19–22]. However, no published work has addressed fully automatic LV segmentation in ^{82}Rb PET, despite its growing adoption and the recognized need for standardized, reproducible MBF quantification [23, 24]. Moreover, existing PET convolutional neural network (CNN)-based methods rarely evaluate performance in hypoperfused or necrotic segments, although these regions critically affect downstream MBF estimation and clinical decision-making.

Recent advances in cardiac image segmentation have explored a wide spectrum of methodologies, including atlas-based registration [17, 25], statistical shape modeling [16, 26], model-based frameworks [27], and deformable-model approaches applied across CT, MRI, and nuclear cardiology [28, 29]. More recently, deep learning architectures such as U-Net and its derivatives [30, 31], attention-based models [20, 32], and spatiotemporal networks designed for gated SPECT and PET [31, 33] have achieved state-of-the-art performance in several modalities. PET-specific segmentation methods have also been reported for tracers including 18F-FPTP [17], 18F-FDG [34], and 13N-ammonia [18], yet these have predominantly targeted gated or static acquisitions with relatively preserved perfusion patterns. Dynamic ^{82}Rb PET presents additional challenges–short half-life, rapid kinetics, low-count early frames, and a higher prevalence of hypoperfused segments–making existing methods insufficiently validated for this context (see Fig. 1).

To address this gap, we developed a comprehensive framework that combines a multimodal manual segmentation protocol, integrating early and late dynamic PET information with CT anatomical guidance, to generate high-quality ground-truth LV masks, together with a fully automatic deep learning pipeline based on nnU-Net. The latter incorporates semi-supervised learning to exploit a large cohort of unlabeled dynamic ^{82}Rb PET studies, thereby improving model robustness and generalizability across a wide range of perfusion patterns.

This study presents the first comprehensive evaluation of deep learning-based LV segmentation in ^{82}Rb PET/CT, including stratified analysis in hypoperfused myocardium–an essential step toward automated, standardized MBF quantification in clinical practice.

Methods

Study population

Between January 2019 and December 2022, all consecutive patients who underwent clinically indicated rest–stress ^{82}Rb PET/CT at our institution were retrospectively screened. The study was approved by the Cantonal Ethics Commission (CER-VD, PB_2018-01513) and conducted according to the Declaration of Helsinki. For development and evaluation of the segmentation

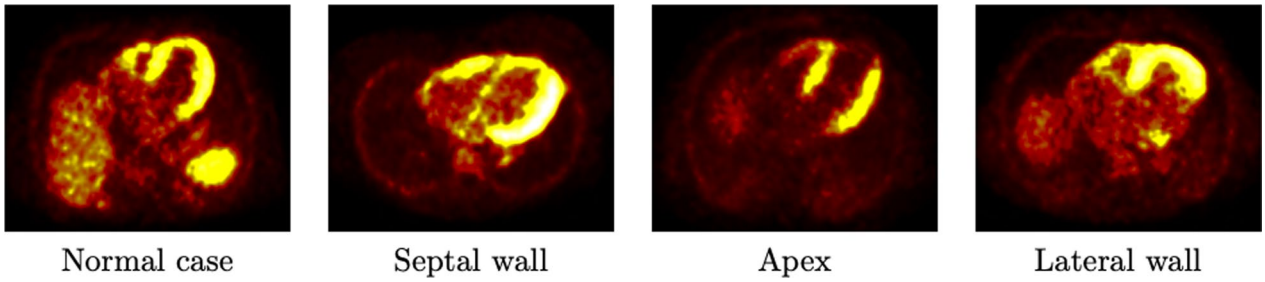

Fig. 1 Examples illustrating challenges in LV segmentation on ^{82}Rb PET/CT. Hypoperfusion, myocardial thinning, and heterogeneous tracer uptake complicate delineation of LV myocardium, particularly when using threshold-based methods

Table 1 Baseline characteristics and relevant imaging findings (All patients, n=20)

Baseline Parameters	All (n=20)	Hemodynamic	All (n=20)
Age (years)	71 ± 7	Sinus rhythm	15 (75%)
Gender (M:F)	17:3	Atrial fibrillation/Flutter	5 (25%)
Height (cm)	172 ± 8	Bundle branch block	1 (5%)
Weight (kg)	79 ± 16	Pacemaker rhythm	3 (15%)
BMI (kg/m^2)	26 ± 4	**Rest Scans**	
Cardiovascular Risk Factors		Heart Rate (bpm)	69±15
Hypertension	14 (70%)	Systolic BP (mmHg)	113±19
Diabetes mellitus	9 (45%)	Diastolic BP (mmHg)	61±12
Smoking history	12 (60%)	**Stress Scans**	
Family history of CAD	2 (10%)	Heart Rate (bpm)	81±18
Dyslipidemia	18 (90%)	Systolic BP (mmHg)	107±19
Cardiovascular History		Diastolic BP (mmHg)	56±14
Previous PCI	16 (80%)	Total burden percentages (median & range)	17% (8–43%)
Previous CABG	5 (25%)	Necrosis percentages (median & range)	21% (5–40%)

framework, 20 patients (40 dynamic PET series including rest and stress acquisitions) were selected. All patients were processed with Corridor 4DM to compute sum rest score (SRS), sum stress score (SSS) and sum difference score (SDS=SSS-SRS) used in clinical routine to characterise necrosis percentage ($\frac{SRS}{68} \cdot 100\%$) and ischemia percentage ($\frac{SDS}{68} \cdot 100\%$). This cohort was intentionally enriched with challenging clinical presentations, including extensive hypoperfusion (high Sum Rest/Stress Scores), myocardial thinning to ensure rigorous testing under conditions where segmentation is most difficult. Baseline characteristics are summarized in Table 1. To support semi-supervised learning, an additional dataset of 805 unlabeled dynamic PET series from 504 patients was included, as long as the complete dynamic sequence and a diagnostic-quality CT were available.

PET/CT acquisition and preprocessing

All PET/CT examinations were acquired on a Biograph Vision PET/CT system (Siemens Healthineers, Erlangen, Germany) using standard dynamic ^{82}Rb imaging protocols. List-mode PET data were reconstructed into 20 dynamic frames (12 × 8 s, 5 × 12 s, 1 × 30 s, 1 × 60 s, and 1 × 120 s) using the vendor's ordered-subset expectation maximization (OSEM) algorithm with time-of-flight and point-spread-function modeling. Reconstructions included corrections for attenuation, scatter, randoms, dead time, and radioactive decay. Low-dose CT scans were acquired for attenuation correction and to provide anatomical guidance for segmentation (Care kV, Care4D; parameters: 100 kVp, 20 mAs, slice thickness 3 mm, pitch 1). PET images were resampled to an isotropic voxel grid prior to training to ensure spatial consistency across subjects and cross-validation folds.

Manual ground-truth segmentation

Reference LV segmentations were generated using a multimodal protocol designed to optimize contouring in hypoperfused myocardium (Fig. 2). Late dynamic PET frames, corresponding to maximal myocardial uptake, served as the primary source for LV delineation. Early frames were inspected to verify the location of the blood pool and avoid its inclusion in the myocardial mask. The registered CT was used to validate myocardial boundaries, particularly in regions with severely reduced uptake or absent perfusion (Fig. 3). All segmentations were produced by a trained imaging specialist and reviewed by a board-certified nuclear medicine physician with expertise in cardiac PET, ensuring anatomical accuracy and consistency across the dataset.

Automatic segmentation using nnU-Net

A fully automatic segmentation pipeline was developed using the nnU-Net framework, which automatically configures preprocessing, network architecture, and training

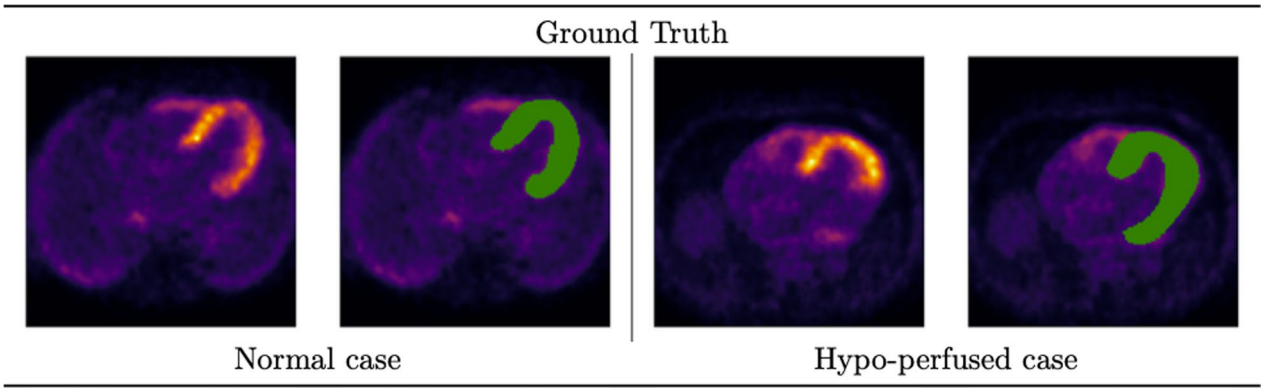

Fig. 2 Representative manual ground-truth LV masks on ^{82}Rb PET/CT. Late PET frames with maximum myocardial uptake were used for contour delineation, while early frames helped discriminate blood pool from myocardium

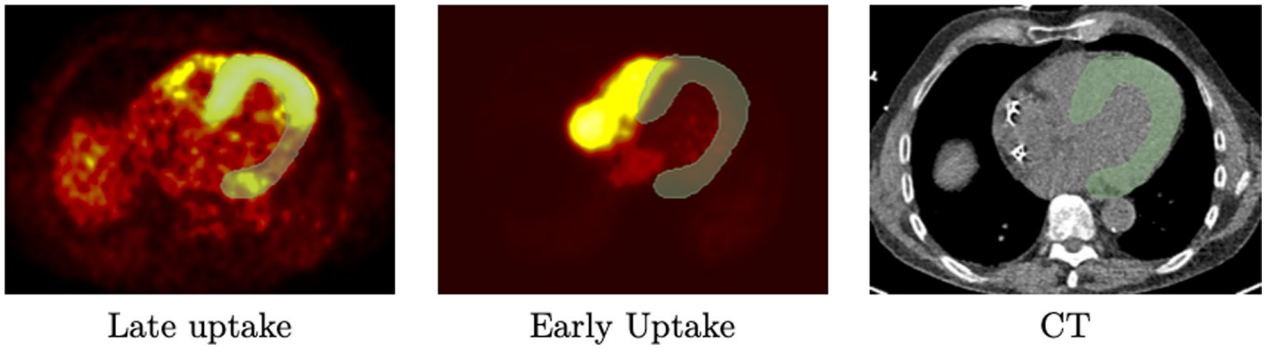

Fig. 3 Multimodal manual segmentation workflow integrating late PET, early PET, and CT anatomical guidance. The combination of modalities improves boundary definition in regions with reduced or absent tracer uptake

hyperparameters. The network input consisted of the last five dynamic PET frames of each study as distinct input channels, corresponding to the clinically relevant static perfusion phase while minimizing blood-pool contamination. Training and evaluation followed a patient-level five-fold cross-validation strategy in which rest and stress acquisitions from the same patient were always assigned to the same fold. In each fold, twelve patients formed the training set, four the validation set, and four the test set. The model was trained using a combined Dice–cross-entropy loss function with nnU-Net's default spatial and intensity augmentations. All experiments were performed on a dedicated workstation equipped with a modern GPU.

Semi-supervised learning

To further improve robustness and generalizability, a semi-supervised learning approach was implemented. For each cross-validation fold, the model trained on the manually annotated dataset was used to generate pseudo-labels for all 805 unlabeled PET series. These pseudo-labeled volumes were then combined with the manual ground truth (GT) to retrain a second-stage model. Pseudo-labels were included as generated, without any confidence thresholding or quality control. Rest and stress scans were treated identically and mixed during training. Manual and pseudo-labeled data were used without additional weighting. Final evaluation was performed exclusively on the manually annotated test subsets for each fold, thereby maintaining strict separation between training and testing data and preventing information leakage.

Baseline semi-automatic thresholding

A threshold-based semi-automatic method was implemented as a clinically interpretable comparator. For each PET series, the maximum standardized uptake value ($SUV_{\max}$) within the LV region was identified, and a threshold of 35% of this value – determined empirically to optimize performance – was applied. To restrict the segmentation to the cardiac region, an approximate bounding region around the LV was defined, and only the largest connected component was retained. This method reflects commonly used clinical approaches and served as a meaningful reference for assessing the added benefit of deep learning–based segmentation.

Evaluation metrics and statistical analysis

Segmentation performance was assessed using the Dice coefficient, recall, and precision. To specifically evaluate performance in compromised myocardium, recall was additionally computed within progressively lower uptake percentiles of the ground-truth mask, allowing a focused assessment of segmentation behavior in hypoperfused regions. All metrics are reported by aggregating the five-fold individual performances and conducting a bootstrap to estimate the mean performance and the associated 95% confidence interval (CI). Necrosis percentage ($\frac{SRS}{68} \cdot 100\%$) is used to stratify rest exams evaluation. Similarly, total burden percentage ($\frac{SSS}{68} \cdot 100\%$) is used for stress exams as the latter are also impacted by ischemia. Due to the selected cohort, patients who underwent viability study, the stratification threshold was set to the median value of each percentage, i.e. median necrosis of 21% and total burden of 25%. To further investigate hypoperfusion impact, we compared rest and stress exams according to the ischemia percentage ($\frac{SDS}{68} \cdot 100\%$) with a median value of 5%. Comparisons between methods were performed using paired statistical tests (paired *t*-test or Wilcoxon signed-rank test depending on normality). Statistical significance was set at $p < 0.05$. All analyses were performed using python (version 3.14.0) and the SciPy library (version 1.15.3). This study was conducted and reported in accordance with the Checklist for Artificial Intelligence in Medical Imaging (CLAIM) guidelines (Supplementary Table S1).

Table 2 Overall segmentation performance comparing the baseline thresholding method and the nnU-Net model across all ^{82}Rb PET/CT studies

	↓Metric/Method→	Thresholding (Baseline)	nnU-Net (Classic)	nnU-Net (Semi-Supervised)
Rest & Stress (40 Scans)	Dice %	75.1[72.9, 76.9]	87.6[85.3, 89.0]	**87.8[85.6, 89.2]**
	Recall %	82.6[79.1, 85.4]	89.0[86.1, 91.3]	**89.1[86.1, 91.4]**
	Precision %	70.2[67.2, 73.0]	87.7[83.8, 90.1]	**88.1[84.2, 90.4]**
Rest (20 Scans)	Dice %	75.1[72.0, 77.8]	87.6[83.4, 89.6]	**87.9[83.8, 90.0]**
	Recall %	82.7[77.5, 86.9]	89.1[84.9, 92.3]	**89.3[84.6, 92.6]**
	Precision %	70.3[65.8, 74.5]	88.0[80.0, 91.2]	**88.4[80.6, 91.5]**
Stress (20 Scans)	Dice %	75.0[72.0, 77.3]	87.5[85.1, 89.3]	**87.6[85.4, 89.5]**
	Recall %	82.3[77.6, 86.3]	88.7[84.6, 91.9]	**88.8[84.8, 92.0]**
	Precision %	70.1[66.1, 74.0]	87.5[82.1, 90.3]	**87.8[82.4, 90.6]**

All patients are aggregated and bootstrapped to estimate the mean and 95%CI. Separate stress/rest evaluations are provided as the stress exam is subject to both necrosis and ischemia. Bold indicates best performance

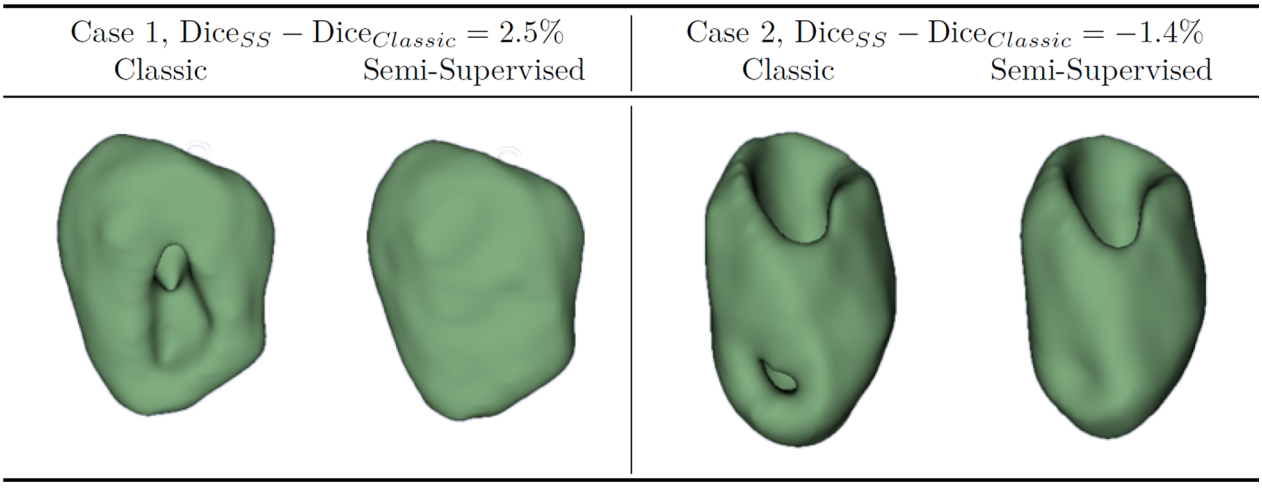

Fig. 4 Qualitative comparison of semi-supervised and classic nnU-Net models on two cases with similar Dice but qualitative differences

Results

A total of 40 dynamic ^{82}Rb PET/CT series from 20 patients (rest and stress) with manual GT LV delineations were included in the evaluation. The cohort exhibited substantial variability in perfusion status, with a median necrosis of 21% (range, 5–40%), enabling detailed assessment of segmentation accuracy in both normal and structurally altered myocardium. For the subgroup analysis, series were divided into low- and high-necrosis/total burden groups using the cohort median value (21%/25%) to ensure balanced subgroup sizes.

Overall segmentation performance

Across all studies, nnU-Nets outperformed the semi-automatic thresholding baseline for every metric (Table 2). The mean Dice coefficient improved from 75.1[72.9, 76.9] % using thresholding to 87.8[85.6, 89.2]% using nnU-Net. Recall increased from 82.6[79.1, 85.4]% to 89.1[86.1, 91.4]%, and precision from 70.2[67.2, 73.0]% to 88.1[84.2, 90.4]%. When comparing the *Classic* nnU-Net, i.e. trained a single time without the semi-supervised approach, with the semi-supervised (SS) nnU-Net, no statistical difference is achieved. However, due to our limited cohort size and patient-based differences, i.e. quantitative differences and qualitative improvements as shown in Fig. 4, it was decided only to focus on the semi-supervised models for the evaluation. Further work with a larger cohort will alleviate this problem by increasing inter-subject variability. Unless otherwise noted, nnU-Net refers specifically to the semi-supervised training framework throughout this work. Visual inspection confirmed that nnU-Net generated smooth and anatomically coherent contours, whereas thresholding frequently included the right ventricle and produced discontinuous masks in hypoperfused regions (Fig. 5).

Rest–stress comparison

Rest and stress acquisitions showed similar performance patterns (Table 2). nnU-Net reached Dice scores of 87.9[83.8, 90.0]% at rest and 87.6[85.4, 89.5]% during stress, while thresholding achieved 75.1[72.0, 77.8]% and 75.0[72.0, 77.3]%, respectively. Last two rows of

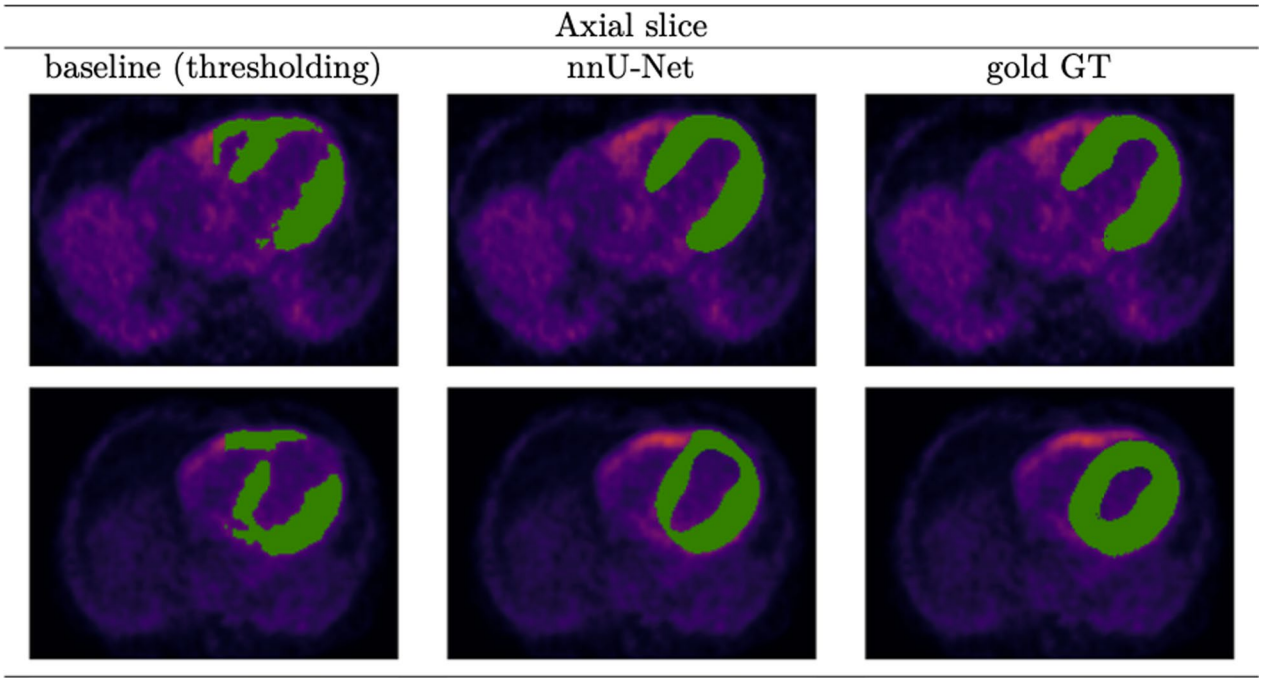

Fig. 5 Comparison of segmentation outputs for two example images (rows) for the baseline thresholding method and the proposed nnU-Net model. nnU-Net generates smooth and anatomically consistent LV contours, while thresholding frequently misclassifies the right ventricle and produces discontinuous masks in hypoperfused regions. Manual GT is shown for reference

Table 3 Overall segmentation performance comparing the baseline thresholding method and the nnU-Net model

	Stratification	**↓Metric/Method→**	**Thresholding**	**nnU-Net**
			(Baseline)	**(Semi-Supervised)**
Rest	Necrosis≤21% (10 Scans)	Dice %	75.8[72.9, 79.8]	**88.0[79.3, 91.3]**
		Recall %	88.2[83.9, 91.6]	**91.3[85.9, 94.5]**
		Precision %	67.5[62.1, 74.8]	**86.7[73.4, 91.6]**
	Necrosis > 21% (10 Scans)	Dice %	74.6[68.5, 78.2]	**87.9[84.9, 90.2]**
		Recall %	77.3[69.6, 83.7]	**87.3[79.2, 92.5]**
		Precision %	73.0[66.8, 77.4]	**90.0[85.1, 94.0]**
Stress	Total burden≤25% (10 Scans)	Dice %	72.9[67.7, 76.2]	**88.0[84.1, 90.0]**
		Recall %	84.5[76.6, 89.9]	**88.2[83.9, 92.6]**
		Precision %	64.7[59.4, 68.1]	**89.2[79.3, 93.0]**
	Total burden > 25% (10 Scans)	Dice %	77.2[74.7, 79.9]	**87.3[84.1, 90.0]**
		Recall %	80.3[73.9, 85.1]	**89.4[82.5, 93.8]**
		Precision %	75.5[70.8, 79.9]	**86.5[79.7, 89.7]**
Mean difference between stress and rest				
Stress-Rest	Ischemia≤5% (11 Scans)	Dice difference %	0.9[−0.9, 3.4]	**0.7[−2.1, 3.7]**
		Recall difference %	0.6[−4.6, 5.3]	**0.1[−2.8, 3.6]**
Stress-Rest	Ischemia > 5% (9 Scans)	Dice difference %	**−1.4[−3.6, 0.9]**	**−1.4[−6.0, 0.2]**
		Recall difference %	−1.7[−4.7, 1.9]	**−1.2[−7.3, 1.1]**

All predictions are aggregated and separated on rest and stress. Rest is stratified on necrosis ($\frac{SRS}{68} \cdot 100\%$) extent with a median of 21% (SRS of 14). Stress is stratified on total burden ($\frac{SSS}{68} \cdot 100\%$) extent with a median of 25% (SSS of 17). Finally, ischemia percentage ($\frac{SDS}{68} \cdot 100\%$) was used to compare rest-stress scans with a median of 5% (SDS of 3). CIs computed using *scipy* bootstrap. Bold indicates best performance

Table 3 shows minor declines in stress performance were observed in patients with extensive ischemia, consistent with increased regional hypoperfusion, but these differences remained small and did not reach statistical significance. Patient-level comparisons revealed no meaningful performance degradation between rest and stress in individuals with low ischemia, consistent with clinical insights.

Stratified Performance

Segmentation performance varied according to the extent of necrosis and total burden (Table 3). In rest

Table 4 Recall of thresholding and nnU-Net within successive^{82}Rb PET uptake percentiles of the LV myocardium (0–20%, 20–35%, 35–50%, etc.), assessing segmentation performance in hypoperfused and necrotic regions

		Recall				
		20%	**35%**	**50%**	**75%**	**100%**
Rest	Thresholding	30.1[19.4, 43.4]	51.8[38.6, 63.1]	65.4[55.0, 73.8]	76.9[70.0, 82.6]	82.7[77.5, 86.9]
	nnU-Net (SS)	64.3[52.4, 74.1]	73.6[62.6, 81.3]	79.7[70.8, 85.7]	85.9[79.8, 90.2]	89.3[84.6, 92.6
Stress	Thresholding	29.0[17.6, 41.5]	50.0[37.1, 61.1]	64.7[55.3, 72.6]	76.4[70.2, 81.7]	82.3[77.6, 86.3]
	nnU-Net (SS)	63.3[51.9, 72.8]	71.9[62.1, 79.9]	78.3[70.6, 84.6]	85.1[79.8, 89.4]	88.8[84.8, 92.0]

CIs computed using *scipy* bootstrap

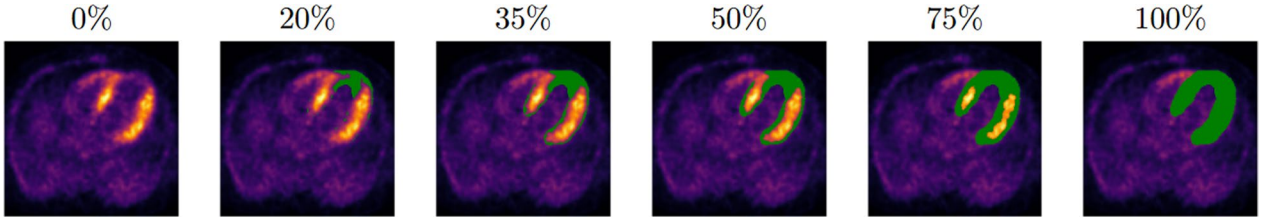

Fig. 6 Illustration of the successive ^{82}Rb PET uptake percentiles of the LV myocardium. Each panel shows the same axial slice, with voxels belonging to the LV myocardium within the corresponding uptake percentile range highlighted in green (0–20%, 20–35%, 35–50%, 50–75%, and 75–100%). Corresponding segmentation performance is detailed in Table 4

patients with necrosis ≤21%, nnU-Net achieved a Dice score of 88.0[79.3, 91.3]%, compared with 75.8[72.9, 79.8]% for thresholding. In patients with necrosis >21%, performance decreased slightly for both methods, but nnU-Net maintained significantly higher Dice (87.9[84.9, 90.2]% vs 74.6[68.5, 78.2]%), confirming its robustness across a spectrum of perfusion defects. Similarly, in stress patients with total burden ≤ 25% nnU-Net Dice is significantly higher than thresholding (88.0[84.1, 90.0]% vs 72.9[67.7, 76.2]%). Compared to stress patients with higher total burden, nnU-Net maintains its superiority and stability compared to thresholding (87.3[84.1, 90.0]% vs 77.2[74.7, 79.9]%). Representative examples of both approaches, including typical failure modes of thresholding in hypoperfused regions, are shown in Fig. 5.

Segmentation performance in hypoperfused regions

To evaluate segmentation accuracy specifically in regions of reduced tracer uptake, recall was computed within successive PET intensity percentiles of the GT (Table 4, Fig. 6). In low-uptake regions (20–35th percentile), nnU-Net markedly outperformed thresholding. For rest scans, recall improved from 30.1[19.4, 43.4]% to 64.3[52.4, 74.1]% at the 20th percentile (+34.2%) and from 51.8[38.6, 63.1]% to 73.6[62.6, 81.3]% at the 35th percentile (+21.8%). For stress scans, improvements were even larger, with gains of +34.3% and +21.9%, respectively. These findings demonstrate that deep learning substantially reduces segmentation failures in precisely the regions where threshold-based approaches struggle most.

Qualitative analysis

Fig. 5 provides a qualitative comparison illustrated the ability of nnU-Net to maintain continuous myocardial contours even in the presence of severe perfusion defects, lateral wall thinning, or apical involvement. In contrast, thresholding consistently produced perforated masks in necrotic segments and misclassified the right ventricle its uptake exceeded the threshold. In extreme cases of transmural necrosis or severe apical thinning, nnU-Net occasionally underestimated the myocardial border (Fig. 7).

Discussion

This study demonstrates that a deep learning–based pipeline using nnU-Net can generate fully automatic and clinically reliable LV contours on ^{82}Rb PET/CT. Accurate LV segmentation is essential for quantitative MBF analysis, which has been shown to provide substantial diagnostic and prognostic value for CAD and CMD [1–4]. MBF and coronary flow reserve predict long-term cardiovascular outcomes, improve detection of multivessel disease, and guide decisions regarding invasive revascularization [5, 6]. Because segmentation inaccuracies propagate through TAC extraction and kinetic modelling, reliable automated contouring is crucial to ensure consistent MBF quantification and clinical interpretation [10, 11]. Accordingly, this work was designed as a segmentation-focused methodological study, and MBF values were not estimated using either nnU-Net or the thresholding baseline.

Across the full dataset, nnU-Net substantially outperformed a clinically inspired thresholding baseline for all segmentation metrics. The improvement was most marked in structurally altered myocardium, where

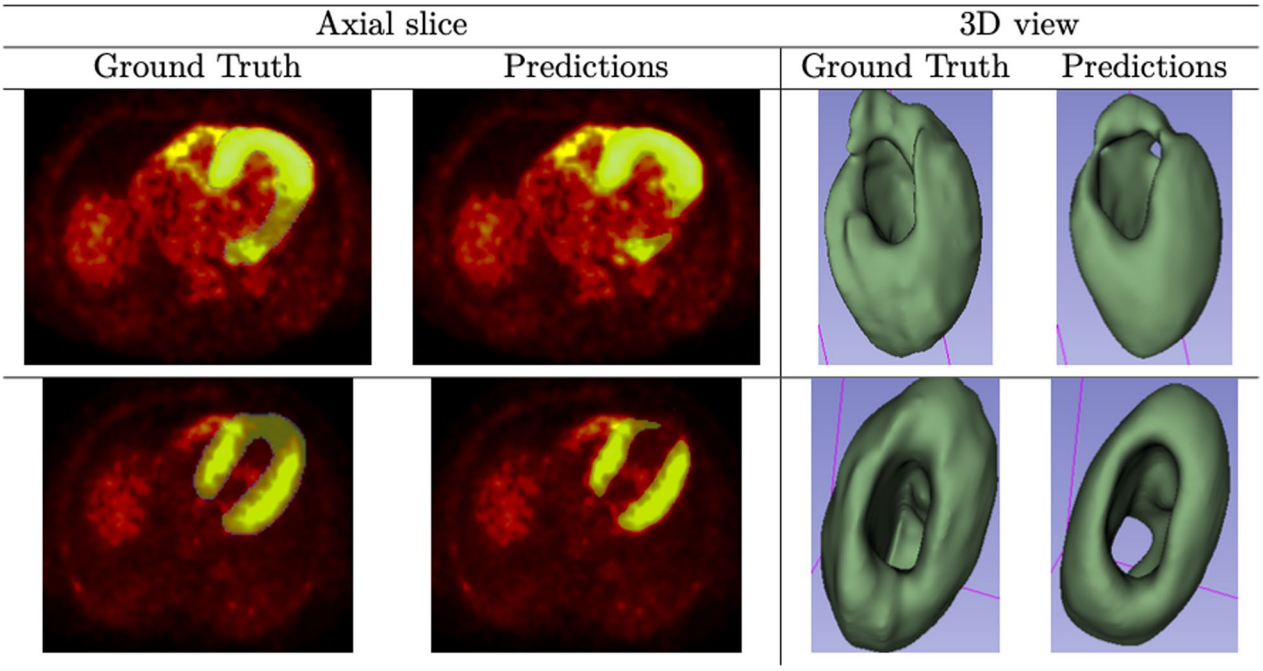

Fig. 7 Challenging segmentation cases for the nnU-Net model. Top row: transmural lateral scar. Bottom row: severe apex thinning

thresholding frequently failed due to reduced tracer uptake. In patients with higher hypoperfusion, SRS > 14 or SSS >17, nnU-Net maintained high Dice performance whereas thresholding degraded significantly, highlighting the model's robustness in tissues affected by scarring, thinning, or severe hypoperfusion. This advantage was further confirmed by percentile-based recall analysis, in which nnU-Net provided 25–40% higher recall in the lowest-perfusion regions–precisely where intensity-based segmentation collapses. These results agree with prior work showing that deep learning methods outperform threshold-based and atlas-based approaches in PET segments with heterogeneous uptake [17, 18, 34], and align with studies demonstrating that individualized LV segmentation improves diagnostic accuracy for hemodynamically significant CAD [35].

Recent CT-based foundation segmentation models and CT-guided pipelines have enabled automated delineation of cardiac structures for PET/SPECT analysis [36–38]. However, CT-derived contours rely on accurate PET–CT co-registration and may be affected by misalignment or motion-related mismatch, which are expected to be significant with stress exams and particularly at the end of the PET acquisition. By segmenting the LV directly in PET space (late frames), our method avoids CT-to-PET contour alignment and is therefore less prone to co-registration–related failure modes. This approach maintains robust segmentation performance in low-contrast and hypoperfused myocardium, which is particularly relevant for dynamic ^{82}Rb PET myocardial perfusion analysis.

To mitigate the limited inter-patient variability impact due to the limited dataset size, a semi-supervised learning scheme was used, which allowed the model to leverage more than 800 unlabelled dynamic PET series. This substantially increased the diversity of myocardial morphologies and uptake patterns encountered during training. Even though there is no significant increase in performance, patient-based differences are in favour of the semi-supervised training scheme for our limited cohort. Benefits of pseudo-labelling and semi-supervision have recently been reported in PET and hybrid imaging segmentation tasks [34, 39, 40]. Given the scarcity of manually segmented PET datasets and the high cost of expert delineation–especially in dynamic imaging–semi-supervised strategies represent a practical and scalable solution for clinical deployment of AI-based segmentation.

The clinical implications of fully automatic LV segmentation in ^{82}Rb PET/CT are substantial. Quantitative PET is increasingly used for risk stratification, detection of balanced ischemia, CMD assessment, and management of complex CAD. Automated segmentation removes a major source of variability in MBF workflows, reduces processing time, and supports reproducibility across operators and institutions. Because MBF values influence decisions regarding revascularization and predict major adverse cardiac events [5, 6], an accurate and automated pipeline may contribute to more consistent patient triage and more reliable longitudinal follow-up. Furthermore, automated LV segmentation is essential for high-throughput retrospective studies and multicenter

MBF harmonization, both of which are critical steps toward standardized clinical adoption [3, 41].

Despite strong overall performance, several limitations must be acknowledged. First, model performance decreased slightly in rare cases with extreme perfusion deficits, such as severe apical thinning or near-transmural scar, where PET activity approaches background levels. These failure modes reflect the inherent dependence of voxel-based models on local image intensity. Future work may incorporate shape-preserving loss functions [42, 43], 3D mesh-based post-processing [44], or hybrid PET-CT or PET-MRI inputs to further enhance robustness in anatomically challenging regions. Second, although cross-validation and semi-supervision mitigated the risk of overfitting, this is a single-center study; multicenter external validation will be required to evaluate generalizability across different scanners, reconstruction kernels, and patient populations. Third, the manual LV segmentations used as reference standard were produced by a single specialist and reviewed by a single expert; inter-observer variability was not assessed and may limit the robustness of the GT labels. In addition, the 35% $SUV_{\max}$ threshold baseline was intentionally chosen as a simple and reproducible segmentation method and should not be interpreted as representative of state-of-the-art clinical software (e.g. FlowQuant, Corridor4DM, QPET). Accordingly, our results should not be interpreted as a direct comparison against specific commercial software tools. Finally, although segmentation quality is known to influence MBF values, this study did not directly assess the impact of automatic delineation on MBF or coronary flow reserve. Future work should evaluate the clinical interchangeability of automated and manual workflows, including within-subject comparisons of MBF and coronary flow reserve across segmentation approaches, and quantify propagation of segmentation errors into MBF.

Looking ahead, several developments could further advance automated LV segmentation for quantitative ^{82}Rb PET/CT. Incorporating temporal information across the full dynamic sequence, integrating multimodal inputs such as gated PET, CT, or MR, and leveraging emerging diffusion-based or foundation AI models may substantially enhance performance in low-count and heterogeneous uptake regions. Large-scale multicenter validation, spanning different scanners, reconstruction methods, and patient populations, will be essential to establish model generalizability and clinical interchangeability. Ultimately, embedding robust automated segmentation within end-to-end MBF pipelines and evaluating its impact on diagnostic accuracy, prognostic value, and clinical decision-making will define its readiness for routine clinical practice.

Conclusions

This study presents the first fully automatic framework for left-ventricle segmentation in ^{82}Rb PET/CT using a deep learning model enhanced with semi-supervised learning. The proposed nnU-Net–based approach achieved near-expert segmentation accuracy, substantially outperforming threshold-based methods, particularly in hypoperfused and necrotic myocardium. Its robustness across rest and stress acquisitions and its independence from manual interaction support its integration into automated myocardial blood flow pipelines. These results demonstrate the feasibility of reproducible, high-throughput LV segmentation for quantitative ^{82}Rb PET/CT and provide a foundation for standardized MBF analysis in clinical and research settings.

Abbreviations

CAD	Coronary artery disease
CMD	Coronary microvascular dysfunction
CNN	Convolutional neural network
CT	Computed tomography
GT	Ground truth
LV	Left ventricle
MBF	Myocardial blood flow
MPI	Myocardial perfusion imaging
OSEM	Ordered-subset expectation maximization
PET	Positron emission tomography
^{82}Rb	Rubidium-82 chloride
SUV	Standardized uptake value
TAC	Time–activity curve
SRS	Sum rest score
SSS	Sum stress score
SDS	Sum difference score (SSS-SRS)

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13550-026-01455-3.

Supplementary file 1.

Author contributions

Experimental design: (M.A., A.C., M.J. and A.D.), acquisition: (M.A., A.C., M.M.A., C.H.K., M.J.), analysis: (M.A., A.C., M.M.A., and A.D.), interpretation: (M.A., A.C., M.M.A., J.O.P., C.H.K. and A.D.), manuscript writing and review (M.A., A.C., M.M.A., R.K., R.D.K., E.M., J.O.P., C.H.K. M.J. and A.D.). All authors have been involved in the writing of the manuscript and approved the final version.

Funding

This work was partially funded by the Swiss National Science Foundation (SNSF) with the project 205320_219430, the Swiss Cancer Research foundation with the project TARGET (KFS-5549-02-2022-R), the Lundin Family Brain Tumour Research Centre at CHUV, and the Hasler Foundation with the project MSxplain number 21042.

Data availability

The patient data analyzed in this study are not publicly available as not permitted by the ethics agreement. Trained model weights, nnU-Net configurations/plans, and inference scripts generated in this work will be made available upon reasonable request to the corresponding author.

Declarations

Ethics approval and consent to participate

The study was approved by the Cantonal Ethics Commission (CER-VD, PB_2018-01513) and conducted according to the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

R.D.K. and R.K. receives royalty revenues from Rubidium-82 PET technologies licensed to Jubilant DraxImage. E.M. is an employee at Jubilant DraxImage Inc.

Received: 5 January 2026 / Accepted: 15 May 2026
Published online: 28 May 2026

References

1. Prior JO, Allenbach G, Valenta I, Kosinski M, Burger C, Verdun FR, et al. Quantification of myocardial blood flow with 82Rb positron emission tomography: Clinical validation with 15O-water. Eur J Nucl Med Mol Imaging. 2012;39(6):1037–47. https://doi.org/10.1007/S00259-012-2082-3/FIGURES/4.
2. Lortie M, Beanlands RSB, Yoshinaga K, Klein R, DaSilva JN, DeKemp RA. Quantification of myocardial blood flow with 82Rb dynamic PET imaging. Eur J Nucl Med Mol Imaging. 2007;34(11):1765–74. https://doi.org/10.1007/S00259-007-0478-2/FIGURES/7.
3. Nensa F, Bamberg F, Rischpler C, Menezes L, Poeppel TD, Fougère C, et al. Hybrid cardiac imaging using PET/MRI: a joint position statement by the European Society of Cardiovascular Radiology (ESCR) and the European Association of Nuclear Medicine (EANM). Eur Radiol. 2018;28(10):4086–101. https://doi.org/10.1007/s00330-017-5008-4.
4. Mimouni M, Bulsei J, Darlington M, Estellat C, Rouzet F, Hyafil F, et al. Cost-effectiveness of 82-Rubidium PET myocardial perfusion imaging for the diagnosis of myocardial ischemia depending on the prevalence of coronary artery disease. EJNMMI Res. 2023;13(1):1–9. https://doi.org/10.1186/S13550-023-00954-X/FIGURES/3.
5. Murthy VL, Naya M, Foster CR, Hainer J, Gaber M, Carli GD, et al. Improved cardiac risk assessment with noninvasive measures of coronary flow reserve. Circulation. 2011;124:2215–24. https://doi.org/10.1161/CIRCULATIONAHA.111.050427.
6. Taqueti VR, Everett BM, Murthy VL, Gaber M, Foster CR, Hainer J, et al. Interaction of impaired coronary flow reserve and cardiomyocyte injury on adverse cardiovascular outcomes in patients without overt coronary artery disease. Circulation. 2015;131:528–35. https://doi.org/10.1161/CIRCULATIONAHA.114.009716.
7. Nakazato R, Berman DS, Dey D, Le Meunier L, Hayes SW, Fermin JS, et al. Automated quantitative Rb-82 3D PET/CT myocardial perfusion imaging: normal limits and correlation with invasive coronary angiography. J Nucl Cardiol. 2012;19(2):265–76.
8. Wang T, Lei Y, Tang H, Harms J, Wang C, Liu T, Curran WJ, Zhou W, Li D, Yang X.: A learning-based automatic segmentation method on left ventricle in SPECT imaging. In: Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 10953, pp. 375–380. SPIE, San Diego, CA, USA (2019). https://doi.org/10.1117/12.2512554. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10953/109531M/A-learning-based-automatic-segmentation-method-on-left-ventricle-in/10.1117/12.2512554 Accessed 2025-05-01
9. Karpikov A, Tagare H, Mulnix T, Gallezot J-D, Sinusas A, Liu C, Carson RE.: Myocardial blood flow from dynamic PET using Independent Component Analysis. In: 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), pp. 2222–2226 (2012). https://doi.org/10.1109/NSSMIC.2012.6551506 . ISSN: 1082-3654. https://ieeexplore.ieee.org/abstract/document/6551506 Accessed 2025-05-01
10. Slomka PJ, Nishina H, Berman DS, Akincioglu C, Abidov A, Friedman JD, et al. Automated quantification of myocardial perfusion SPECT using simplified normal limits. J Nucl Cardiol. 2005;12:66–77. https://doi.org/10.1016/J.NUCLCARD.2004.10.006.
11. Danad I, Szymonifka J, Twisk JWR, Norgaard BL, Zarins CK, Knaapen P, et al. Diagnostic performance of cardiac imaging methods to diagnose ischaemia-causing coronary artery disease when directly compared with fractional flow reserve as a reference standard: a meta-analysis. Eur Heart J. 2017;38:991–8. https://doi.org/10.1093/EURHEARTJ/EHW095.
12. Germano G, Kiat H, Kavanagh PB, Moriel M, Mazzanti M, Su HT, et al. Automatic quantification of ejection fraction from gated myocardial perfusion SPECT. Journal of Nuclear Medicine Official Publication Society of Nuclear Medicine. 1995;36(11):2138–47.
13. Motwani M, Berman DS, Germano G, Slomka PJ. Automated quantitative nuclear cardiology methods. Cardiol Clin. 2016;34(1):47–57. https://doi.org/10.1016/j.ccl.2015.08.003.
14. Kubik T, Kałużyński K, Burger C, Passeri A, Margiacchi S, Saletti P, et al. Novel 3D heart left ventricle muscle segmentation method for PET-gated protocol and its verification. Ann Nucl Med. 2019;33:629–38.
15. American Heart Association Writing Group on Myocardial Segmentation and Registration for Cardiac Imaging, Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the CardiacImaging Committee of the Council on Clinical Cardiology of the American Heart Association. Circulation. 2002;105(4):539–42. https://doi.org/10.1161/hc0402.102975.
16. H K, Wy L, Ka A, S Y, J X, H. C. Left ventricle segmentation in cardiac MRI using data-driven priors and temporal correlations. J Cardiovasc Magn Reson. 2010;12(Suppl 1):91. https://doi.org/10.1186/1532-429X-12-S1-P91.
17. Kim SJW, Seo S, Kim HS, Kim D-Y, Kang KW, Min J-J, et al. Multi-atlas cardiac PET segmentation. Phys Med. 2019;58:32–9. https://doi.org/10.1016/j.ejmp.2019.01.003.
18. Zhang Y, Wang F, Wu H, Yang Y, Xu W, Wang S, et al. An automatic segmentation method with self-attention mechanism on left ventricle in gated PET/CT myocardial perfusion imaging. Comput Methods Programs Biomed. 2023;229:107267.
19. Bom MJ, Schumacher SP, Driessen RS, Raijmakers PG, Everaars H, Diemen PA, et al. Impact of individualized segmentation on diagnostic performance of quantitative positron emission tomography for haemodynamically significant coronary artery disease. Eur Heart J Cardiovasc Imaging. 2019;20(5):525–32. https://doi.org/10.1093/ehjci/jey201.
20. Shoaib MA, Chuah JH, Ali R, Hasikin K, Khalil A, Hum YC, et al. An overview of deep learning methods for left ventricle segmentation. Comput Intell Neurosci. 2023;2023(1):4208231. https://doi.org/10.1155/2023/4208231.
21. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18(2):203–11. https://doi.org/10.1038/s41592-020-01008-z.
22. Peng P, Lekadir K, Gooya A, Shao L, Petersen SE, Frangi AF. A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. Magn Reson Mater Phys, Biol Med. 2016;29:155–95.
23. Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W, et al. Deep learning for cardiac image segmentation: a review. Front Cardiovasc Med. 2020;7:25.
24. Irshad M, Sharif M, Yasmin M, Khan A. A Survey on Left Ventricle Segmentation Techniques in Cardiac Short Axis MRI. Current Medical Imaging Reviews. 2018;14(2):223–37. https://doi.org/10.2174/1573405613666170117124934.
25. Gigengack F, Ruthotto L, Jiang X, Modersitzki J, Burger M, Hermann S, et al. Atlas-Based Whole-Body PET-CT Segmentation Using a Passive Contour Distance. In: Menze BH, Langs G, Lu L, Montillo A, Tu Z, Criminisi A, editors. Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging. Berlin, Heidelberg: Springer; 2013. p. 82–92. https://doi.org/10.1007/978-3-642-36620-8_9.
26. Woo J, Slomka PJ, Kuo C-C, Hong B-W. Multiphase segmentation using an implicit dual shape prior: application to detection of left ventricle in cardiac MRI. Comput Vis Image Underst. 2013;117(9):1084–94. https://doi.org/10.1016/j.cviu.2012.11.012.
27. Gerard O, Billon AC, Rouet J-M, Jacob M, Fradkin M, Allouche C. Efficient model-based quantification of left ventricular function in 3-D echocardiography. IEEE Trans Med Imaging. 2002;21(9):1059–68. https://doi.org/10.1109/TMI.2002.804435.
28. Cong J, Zheng Y, Xue W, Cao B, Li S. MA-Shape: modality adaptation shape regression for left ventricle segmentation on mixed MR and CT images. IEEE Access. 2019;7:16584–93. https://doi.org/10.1109/ACCESS.2019.2892965.
29. Alam I, Tripathi RC, Raghavendra R.: Adversarial Transfer Learning for Left Ventricle Segmentation in Cardiac CT Images. In: 2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC), pp. 1–6 (2024). https://doi.org/10.1109/ICOCWC60930.2024.10470920. https://ieeexplore.ieee.org/abstract/document/10470920 Accessed 2025-05-01

30. Ronneberger O, Fischer P, Brox T.: U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI). Springer, Munich, Germany (2015). http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a Accessed 2022-12-06
31. Wang Y, Zhang W. A dense RNN for sequential four-chamber view left ventricle wall segmentation and cardiac state estimation. Front Bioeng Biotechnol. 2021. https://doi.org/10.3389/fbioe.2021.696227.
32. Wang H, Xian M, Vakanski A. Ta-net: Topology-aware network for gland segmentation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2022. p. 1556–64.
33. Zhao C, Shi S, He Z, Malhotra S, Wang C, Zhao Z, et al. Spatial-temporal V-Net for automatic segmentation and quantification of right ventricle on gated myocardial perfusion SPECT images. Med Phys. 2023;50(12):7415–26. https://doi.org/10.1002/mp.16805.
34. Poitrasson-Rivière A, Vanderver MD, Hagio T, Arida-Moody L, Moody JB, Renaud JM, et al. Automated deep learning segmentation of cardiac inflammatory FDG PET. J Nucl Cardiol. 2024;42:102052. https://doi.org/10.1016/j.nuclcard.2024.102052.
35. Bom MJ, Schumacher SP, Driessen RS, Raijmakers PG, Everaars H, Diemen PAV, et al. Impact of individualized segmentation on diagnostic performance of quantitative positron emission tomography for haemodynamically significant coronary artery disease. Eur Heart J Cardiovasc Imaging. 2019;20:525–32. https://doi.org/10.1093/EHJCI/JEY201.
36. Wasserthal J, Breit H-C, Meyer MT, Pradella M, Hinck D, Sauter AW, et al. Totalsegmentator: robust segmentation of 104 anatomic structures in ctimages. Radiol Artif Intell. 2023;5(5):230024. https://doi.org/10.1148/ryai.230024.
37. Miller RJ, Shanbhag A, Marcinkiewicz AM, Struble H, Gransar H, Hijazi W, et al. Ai-enabled ct-guided end-to-end quantification of total cardiac activity in 18fdg cardiac pet/ct for detection of cardiac sarcoidosis. J Nucl Cardiol. 2025;48:102195. https://doi.org/10.1016/j.nuclcard.2025.102195.
38. Auriac J, Nioche C, Hovhannisyan-Baghdasarian N, Loisel C, Seban R-D, Jehanno N, et al. Comparison of organ volumes and standardized uptake values in [^{18}F]FDG-PET/CT images using Moose and TotalSegmentator to segment CT images. Med Phys. 2025;52(10):70025. https://doi.org/10.1002/mp.70025.
39. Jiao R, Zhang Y, Ding L, Xue B, Zhang J, Cai R, et al. Learning with limited annotations: a survey on deep semi-supervised learning for medical image segmentation. Comput Biol Med. 2024;169:107840. https://doi.org/10.1016/j.compbiomed.2023.107840.
40. Zhang Y, Wang F, Wu H, Yang Y, Xu W, Wang S, et al. An automatic segmentation method with self-attention mechanism on left ventricle in gated PET/CT myocardial perfusion imaging. Comput Methods Programs Biomed. 2023;229:107267. https://doi.org/10.1016/J.CMPB.2022.107267.
41. Dekemp RA, Declerck J, Klein R, Pan X-B, Nakazato R, Tonge C, et al. Multisoftware reproducibility study of stress and rest myocardial blood flow assessed with 3D dynamic PET/CT and a 1-tissue-compartment model of ^{82}Rb kinetics. J Nucl Med. 2013;54(4):571–7. https://doi.org/10.2967/jnumed.112.112219.
42. Mosinska A, Marquez-Neila P, Kozinski M, Fua P. Beyond the pixel-wise loss for topology-aware delineation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2017. p. 3136–45. https://doi.org/10.1109/CVPR.2018.00331.
43. Yan Z, Yang X, Cheng K-TT.: A deep model with shape-preserving loss for gland instance segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11, pp. 138–146 (2018). Springer
44. Hattori S, Yatagawa T, Ohtake Y, Suzuki H.: Deep Mesh Prior: Unsupervised Mesh Restoration using Graph Convolutional Networks. (2021). https://doi.org/10.48550/arXiv.2107.02909. arXiv:2107.02909 Accessed 2025-05-29

Publisher's note