

Comparison of a Hybrid Mixture Model and a CNN for the Segmentation Myocardial Pathologies in Delayed Enhancement MRI

Markus Huellebrand^{1,2}, Matthias Ivantsits¹, Hannu Zhang¹, Peter Kohlmann²,
Jan-Martin Kuhnigk², Titus Kuehne^{1,3,4}, Stefan Schönberg⁵, and Anja
Hennemuth^{1,2,4}

¹ Charité – Universitätsmedizin Berlin, Augustenburger Pl. 1, 13353 Berlin, Germany

² Fraunhofer MEVIS, Am Fallturm 1, 28359 Bremen, Germany

³ Department of Radiology and Nuclear Medicine, University Medical Center
Mannheim, 68167 Mannheim, Germany

⁴ German Heart Institute Berlin, Augustenburger Pl. 1, 13353 Berlin, Germany

⁵ DZHK (German Centre for Cardiovascular Research), Berlin, Germany

Abstract. DE-MRI provides a reliable and accurate imaging technique for the assessment of pathological alterations in myocardial tissue. The clinically applied thresholding techniques enable the assessment of the amount of diseased tissue. To also assess distribution patterns, transmural and micro-vascular obstruction, more accurate segmentation methods are needed. We compare a hybrid CNN and mixture model approach with a two single-stage U-net segmentation: one based on the EMIDEC challenge data set, one with additional training data and could achieve DICE coefficients of 84.8%, 84.08%, and 82.95%, respectively. We hope to further improve the promising results through an extension of the training set.

Keywords: Delayed enhancement MRI · Mixture Model · CNN · U-Net.

1 Introduction

The analysis of delayed-enhancement magnetic resonance imaging provides an effective technique to analyze the state of the myocardial tissue after myocardial infarction. The analysis helps to select treatment, i.e., can give insight if a revascularization therapy will be successful. An automated, standardized way of segmenting the different areas such as the myocardium, the infarcted tissue, and the permanent microvascular obstruction will improve the diagnosis and therapeutic decision. Over the last decades, several segmentation approaches have been developed. However, the analysis of these DE-MRI is still a challenging task due to bad image contrast, image- and motion artifacts. In this challenge, we tested two different approaches: one traditional mixture-model based approach and one based on CNN for the differentiation between normal myocardium, regions with late enhancement, and no-reflow areas where neither wash-in nor wash-out of contrast agent can be observed because of microvascular obstructions.

2 Material and Methods

2.1 Image Data

The image data for the EMIDEC challenge [6] consists of 150 cases, 100 diseased patients and 50 normal cases. The data set is split into a training set with 100 cases and a testing set with 50 cases. The training set as well as the testing set contains 1/3 normal and 2/3 of pathological cases, which roughly corresponds to real life observations in clinical settings. The data was acquired on Siemens MRI scanners on 1.5T (Area) and 3T (Skyra); the in-plane resolution was $1.25 \times 1.25 \text{ mm}^2$ and $2 \times 2 \text{ mm}^2$, a slice thickness of 8 mm and a distance between slices of 8 to 13 mm . In a post-processing step the image slices were realigned to prevent any drawbacks resulting from breathing motions.

Furthermore, the EMIDEC challenge organizers provided ground-truth segmentations for the training set containing labels for the **cavity**, **myocardium**, **myocardial infarction**, and **no-reflow** areas. Figure 1 illustrates an example of a normal case and two cases with myocardial infarction and no-reflow areas.

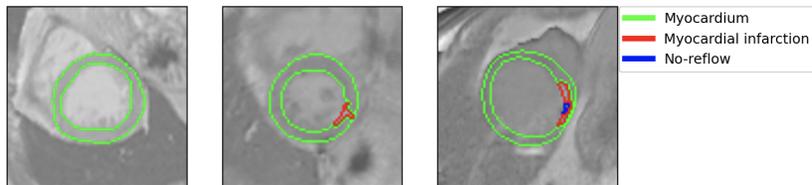


Fig. 1: **Left**: an MRI acquisition of a patient without myocardial infarction. **Center**: a pathological case with myocardial infarction. **Right**: a pathological case with myocardial infarction and no-reflow regions.

2.2 Methods

Background There have been many attempts to provide segmentation methods for a reproducible quantitative assessment of myocardial fibrosis based on late gadolinium enhancement imaging. The overview paper from the STACOM challenge by Karim et al. presented many conventional voxel classification approaches based on intensity distributions [5]. Most approaches are organized in two steps: first the myocardium is segmented to reduce the problem and provide anatomical context information such as the relative position of a voxel with regard to the endocardial border of the myocardium, then the myocardium is analyzed. More recent approaches make use of modern machine learning techniques such as CNNs [11]. Because of the lack of large training cohorts most successful approaches still use a multi-step approach, which combines different types of segmentation and classification methods. Zabihollaly et al. suggest the combination

of a myocardium segmentation U-net with an ensuing CNN-based classification of the myocardial voxels [12]. The approach by de la Rosa et al. applies a CNN for the preselection of image slices to analyze with regard to myocardial pathologies [9]. The reported DICE-values of these multi-step approaches are 88% and 77% respectively. The single-stage-CNN-approaches obviously depend more on the training set. The reported DICE coefficients in published approaches are minimum 48% [2] and can be improved through additional data sets generated with Generative Adversarial Neural Networks (GAN) [7].

CNN segmentation Over the last years, CNN based segmentation approaches often outperformed traditional approaches in medical image processing challenges. In former work[10], we already applied a 2D 4-layer u-net [8] successfully to segment the left and right ventricle on cine MRI. In this work, we tried to use that prior knowledge for the challenge.

We chose the same architecture as in our previous work and tested if transfer learning can improve the robustness and quality of the results. We used 100 additional DE-MRI data sets for which only blood-pool and myocardial labels were provided. Based on these data sets, we trained two CNNs. One u-net that was only trained with the challenge data. The other u-net was first trained on the additional DE-MRI data sets. As these images only contain labels for the left ventricular cavity and myocardium, we could not directly train our model on both data sets together, or start the training on the additional data and resume it on the challenge data directly. We had to reset the last layer and restart the training in order to incorporate the additional labels.

For training of both experiments, we chose a learning rate of 0.005, categorical cross-entropy as loss function, and used drop-out and batch normalization.

Hybrid Mixture Model-based segmentation: MM-RG Our mixture-model approach is based on [3] but is adapted to the current data, definition of the micro-vascular obstructions, and the definition of thresholds to generate binary masks. The approach uses the intensity distribution in the myocardium and the expected location and size of myocardial infarcts. In the first step, the myocardium has to be delineated. We used the same 4-layer U-net architecture for this task. As we only have to differentiate background, cavity and myocardium started with a u-net trained on cine MRI from the ACDC challenge data set [1] and in-house cine MRI data sets. In a second step, we used the additional DE-MRI data as well as the challenge data and resumed the training with that data. For the challenge data, we replaced the labels for infarction and no-reflow areas by the label for the myocardium. In a second step, we perform a mixture model fit of an expected distribution model to the myocardial histogram. For MRI data, a mixture of Rice and a Gaussian distribution or a mixture of Rayleigh and Gaussian can be assumed. During the training phase, we achieved better results using a mixture of Rayleigh and Gaussian. As a third step, the infarcted tissue is segmented, incorporating expected position and spatial connectivity. In [4], a watershed segmentation based on automatically extracted seed points in

Metric	MM-RG	Transfer-learning	EMIDEC
Dice myocardium (%)	92.07	82.43	81.00
Volume difference myocardium (mm^3)	3863.89	11766.66	13655.55
Hausdorff myocardium (mm)	9.30	15.75	16.72
Dice MI (%)	57.25	34.09	36.08
Volume difference MI (mm^3)	6048.54	5994.9	8980.5
MI volume difference ratio (%)	4.99	4.94	7.07
Dice no-reflow (%)	39.29	40.26	54.15
Volume difference no-reflow (mm^3)	1301.80	1520	1501.73
No-reflow volume difference ratio (%)	0.97	1.11	1.08

Table 1: Internal quantitative evaluation on 5-fold cross-validation

high-intensity voxels, located at the inner part of the myocardium, was used. As a binary segmentation is needed in this challenge, we chose the threshold such that resulting probabilities were optimal on the training set using a brute force optimization approach. No re-flow areas are detected by morphological closing because we assume that they are surrounded by blood-pool and/or enhanced fibrotic tissue.

3 Results

We evaluated all models based on the metrics provided for the challenge.

For the myocardial region the **DICE** index, **Hausdorff distance**, and **volume difference** are calculated. For infarcted tissue and micro-vascular obstructions the **DICE** index, **volume difference**, and **volume difference ratio, according to the myocardium**, are used.

For the pure CNN based approaches, we used 5-fold cross-validation during our experiments.

Table 1 gives an extensive overview of the conducted experiments' metrics on the internal validation. The CNN for the hybrid mixture model shows superior DICE-values on the myocardium (see figure 2). Notably, the mixture of Rayleigh and Gaussian fails to accurately infer a segmentation on the no-reflow areas within the myocardial tissue. Remarkable is the results of the randomly initialized models, which produces the top performance in the no-reflow area.

Again the mixture model of Rayleigh and Gaussian demonstrates preferable results on the Hausdorff distance (2). With a 9.3 ± 5.4 mm distance to the ground-truth surface, the results are considerably exceeding the performance of the other tested methods. Moreover, the margin of error is explicitly lower with even fewer outliers compared to the transfer-learning and randomly initialized method.

Table 2 shows the final result of the different approaches on the test set. The performance of the CNN based approaches showed a better performance on the

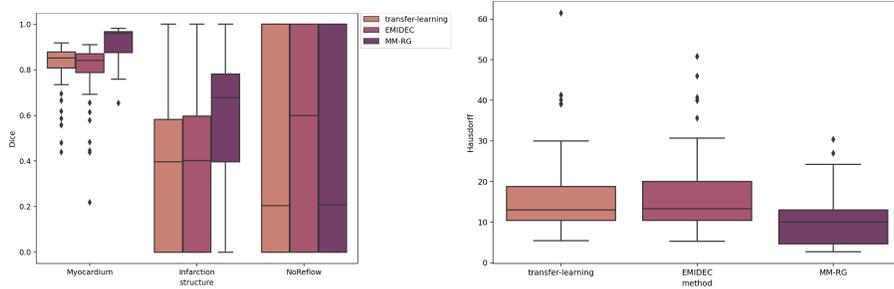


Fig. 2: An illustration of the models results. **Left:** DICE index for the myocardium infarction and no re-flow area. For the myocardium and the no re-flow areas, value zero and one in the box plots result from cases, in which i.e. regions bright regions were classified as infarct in healthy subjects. **Right:** Hausdorff distance for the different segmentation approaches. The CNN that was only trained to segment the myocardium, and was also trained on additional DE-MRI data as well as on cine MRI data, outperformed the CNN approaches that also tried to detect infarcted tissue.

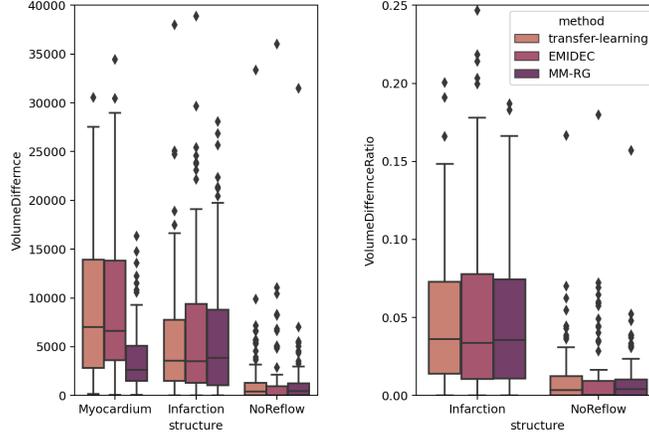


Fig. 3: **Left:** illustration of the volume difference between proposed methods and the ground truth. For the myocardium, the mixture model approach achieved the best results. In the infarction and the micro-vascular obstruction, the three approaches achieved similar results. **Right:** illustration of the volume difference ratio of infarction and micro-vascular obstructions in relation to the complete myocardial volume.

Metric	MM-RG	Transfer-learning	EMIDEC
Dice myocardium (%)	84.8	84.08	82.95
Volume difference myocardium (mm^3)	12582.76	10874.47	9060.61
Hausdorff myocardium (mm)	15.93	18.3	16.16
Dice MI (%)	35.83	37.87	37.33
Volume difference MI (mm^3)	6100.92	6166.01	6021.72
MI volume difference ratio (%)	5.17	4.93	4.87
Dice no-reflow (%)	41.25	52.25	48.39
Volume difference no-reflow (mm^3)	1090.04	953.47	990.29
No-reflow volume difference ratio (%)	0.76	0.64	0.64

Table 2: Quantitative comparison of the final results on testing set.

test set compared to the internal tests. The drop of performance of the mixture model results come from the fact, that we did not perform a cross validation and used the final model to generate the myocardium segmentation. However, the performance of the myocardium segmentation by this method still achieved the best results for DICE index and Hausdorff distance.

4 Discussion

We have demonstrated three different approaches to segment the myocardium, infarcted tissue, and micro-vascular obstructions from DE-MRI data. For pure CNN based segmentation approaches, the results on the actual testing set were better than in our internal evaluation using 5-fold cross-validation. As we re-trained our models for the submission using all data, we think that the improvement comes from the additional cases. The slightly better results for the transfer learning approach also back this hypothesis. The u-net that was additionally trained on cine data showed the best results in the segmentation of the myocardium according to the dice index and Hausdorff distance. This could also be due to the better generalization by the additional data sets. The improvement could also come from the reduction of the number of classes to be classified. The drop in the performance of the results by the hybrid approach can be explained by using the final CNN to extract the myocardial segmentation. Here we can see that the quality of the segmentation of the infarction and no-reflow areas is highly dependent on a robust segmentation of the myocardium.

Figure 5 shows several segmentation results. In the top row, healthy tissue is segmented correctly by all approaches. Row two shows an example of infarction with no re-flow area that was also correctly depicted by all methods and shows a good agreement with the ground truth. In row three, both pure CNN based approaches miss-classified brighter regions in the basal slices as infarction. In row four, an example is provided, where the mixture model approach could not correctly segment the no re-flow area.

We saw in several cases that the segmentation of the infarction was underestimated by the mixture model approach and the no-reflow areas were not surrounded by the infarction mask. Thus the no-reflow areas could not be segmented using simple morphological closing operations. Here one could try using radial closing approaches to try to overcome this limitation. To reduce the number of cases without infarction in which voxels were misclassified as infarcted tissue, one could additionally analyze the fitted distributions. One could i.e. analyze if the mean of the distributions is too close together.

In future work, it makes sense to investigate the performance of 2-step u-net approaches, also incorporating the results of our mixture model analysis. To achieve better results for the smaller regions infarction and no-reflow areas we could try to use weighted categorical cross-entropy as a loss function. Additionally, we could try if self-learning on the test set can further improve the overall performance of the proposed method.

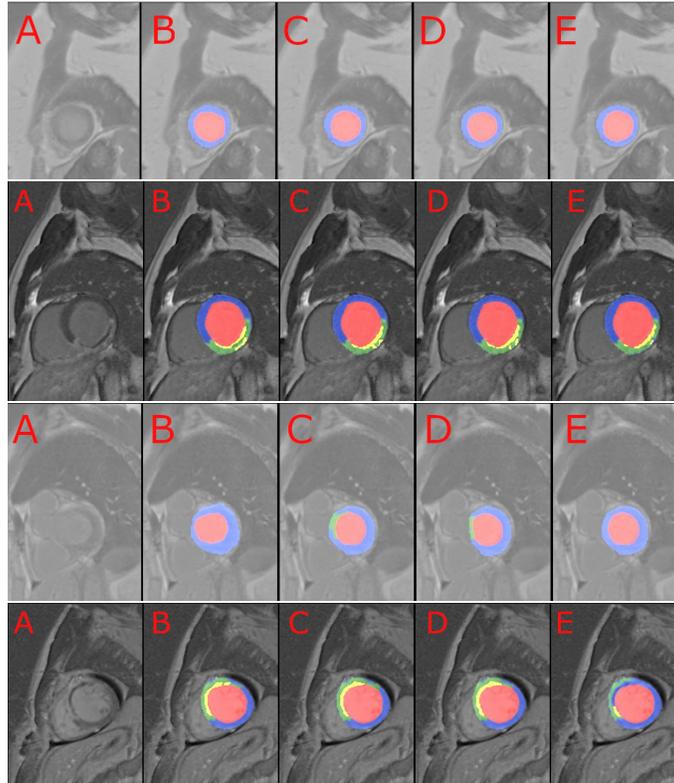


Fig. 4: Example results from the training data set. Each row shows a case on different locations in the heart. On the left, the original image is shown (A), followed by the ground truth (B), transfer learning (C), EMIDEC training (D) and the Mixture model of Rayleigh and Gaussian distribution on the right (E).

References

1. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P., et al.: Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions on Medical Imaging* **37**(11), 2514–2525 (2018)
2. Fahmy, A.S., Neisius, U., Chan, R.H., Rowin, E.J., Manning, W.J., Maron, M.S., Nezafat, R.: Three-dimensional Deep Convolutional Neural Networks for Automated Myocardial Scar Quantification in Hypertrophic Cardiomyopathy: A Multicenter Multivendor Study. *Radiology* **294**(1), 52–60 (2020)
3. Hennemuth, A., Friman, O., Huellebrand, M., Peitgen, H.O.: Mixture-model-based segmentation of myocardial delayed enhancement in MR image data. In: *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*. vol. 7746, pp. 87–96 (2013)
4. Hennemuth, A., Seeger, A., Friman, O., Miller, S., Klumpp, B., Oeltze, S., Peitgen, H.O.: A Comprehensive Approach to the Analysis of Contrast Enhanced Cardiac MR Images. *IEEE Transactions on Medical Imaging* **27**(11), 1592–1610 (2008)
5. Karim, R., Bhagirath, P., Claus, P., James Housden, R., Chen, Z., Karimaghloo, Z., Sohn, H.M., Lara Rodríguez, L., Vera, S., Albà, X., Hennemuth, A., Peitgen, H.O., Arbel, T., González Ballester, M., Frangi, A., Götte, M., Razavi, R., Schaeffter, T., Rhode, K.: Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images. *Medical Image Analysis* **30** (2016)
6. Lalande, A., Chen, Z., Decourselle, T., Qayyum, A., Pommier, T., Lorgis, L., de la Rosa, E., Cochet, A., Cottin, Y., Ginhac, D., Salomon, M., Couturier, R., Meriaudeau, F.: Emidec: A Database Usable for the Automatic Evaluation of Myocardial Infarction from Delayed-Enhancement Cardiac MRI. *Data* **5**(4), 89 (Sep 2020)
7. Lau, F., Hendriks, T., Lieman-Sifry, J., Sall, S., Golden, D.: Scargan: Chained generative adversarial networks to simulate pathological tissue on cardiovascular mr scans bt - deep learning in medical image analysis and multimodal learning for clinical decision support. pp. 343–350. Springer International Publishing (2018)
8. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. pp. 234–241. Springer International Publishing, Cham (2015)
9. de la Rosa, E., Sidibé, D., Decourselle, T., Leclercq, T., Cochet, A., Lalande, A.: Myocardial Infarction Quantification From Late Gadolinium Enhancement MRI Using Top-hat Transforms and Neural Networks (2019)
10. Tautz, L., Zhang, H., Hüllebrand, M., Ivantsits, M., Kelle, S., Kühne, T., Falk, V., Hennemuth, A.: Cardiac radiomics an interactive approach for 4d data exploration. *Current Directions in Biomedical Engineering* (2020)
11. Zabihollahy, F., Rajan, S., Ukwatta, E.: Machine Learning-Based Segmentation of Left Ventricular Myocardial Fibrosis from Magnetic Resonance Imaging. *Current Cardiology Reports* **22**(8), 65 (aug 2020)
12. Zabihollahy, F., Rajchl, M., White, J.A., Ukwatta, E.: Fully automated segmentation of left ventricular scar from 3D late gadolinium enhancement magnetic resonance imaging using a cascaded multi-planar U-Net (CMPU-Net). *Medical Physics* **47**(4), 1645–1655 (2020)