



HAL
open science

A comparative study of CNN and U-Net performance for automatic segmentation of medical images: application to cardiac MRI

Wafa Baccouch, Sameh Oueslati, Basel Solaiman, Salam Labidi

► To cite this version:

Wafa Baccouch, Sameh Oueslati, Basel Solaiman, Salam Labidi. A comparative study of CNN and U-Net performance for automatic segmentation of medical images: application to cardiac MRI. CENTERIS - International Conference on ENTERprise Information Systems., Nov 2022, Lisbonne, Portugal. pp.1089 - 1096, <10.1016/j.procs.2023.01.388>. <hal-04042038>

HAL Id: hal-04042038

<https://imt-atlantique.hal.science/hal-04042038v1>

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2022

A comparative study of CNN and U-Net performance for automatic segmentation of medical images: application to cardiac MRI

Wafa Baccouch^{a,*}, Sameh Oueslati^a, Basel Solaiman^b, Salam Labidi^a

^aUniversity of Tunis El Manar, Higher institute of Medical Technologies of Tunis, Research laboratory of Biophysics and Medical Technologies
Tunis 1006, Tunisia

^bImage & Information Processing Department (iTi), IMT-Atlantique, Technopôle Brest Iroise CS 83818, 29238 Brest Cedex France

Abstract

Medical image segmentation is one of the most challenging tasks in medical image analysis which aims to extract effective information and improve the level of clinical diagnosis. In last decades, automatic segmentation based on Deep Learning (DL) models such as U-Net and CNN architectures have been widely used to automatically extract organs or lesions contours in order to overcome manual segmentation limitations. In this paper, we performed a comparative study between CNN and U-Net performance for medical image segmentation applied to cardiac-MRI segmentation using U-Net from short-axis MRI images of ACDC database. The adopted architecture was trained and tested with and without data augmentation. The obtained results show a strong agreement between the labelled masks and the predicted ones with a mean DSC that reached 97,9% and a mean Hausdorff Distance (HD) that reached 5.318 mm. A quantitative comparison was made on two levels. The first one is an intra-comparison made between the adopted model and methods based on the same architecture and having been trained and tested using the same database which proved that our method reached the highest performance and can be considered as a promising tool for medical image segmentation. The second one is an inter-comparison made between U-Net and CNN performance which proved that U-Net is more suitable for carrying out this task since it takes less time for training as it does not have a fully connected layer and offers a fairly significant similarity to ground truth compared to CNN.

© 2023 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies 2022

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .

E-mail address: wafa.baccouch@etudiant-istmt.utm.tn

Keywords: Medical image segmentation, U-Net, CNN, Right Ventricle, Left Ventricle, data augmentation, cardiac-MRI

1. Introduction

Medical image analysis consists in highlighting certain characteristic information of an image or in classifying the image itself to solve clinical problems. Since it directly influences the diagnosis, this task still of great importance [1]. As one of the most important links in medical image analysis, is segmentation. In clinical practice, once the sequences are acquired, the radiologist examines them carefully and tries to identify the disease and possible causes. This process is tedious and time consuming as it takes from few hours to few days depending on several factors such as the severity of the pathology and clinician expertise [2]. In this context, to overcome the constraints related to manual segmentation, several methods have been proposed, in the literature, to automate this task such as Machine Learning (ML) based works [3]. However, when it comes to efficiently processing big data and extracting the most relevant image features, traditional and semi-automatic segmentation methods are increasingly considered unable of helping doctors accurately diagnose and treat diseases [4]. Recently, DL methods, specifically Convolutional Neural Networks (CNNs) and U-Net have been extensively involved in medical image segmentation. Moeskops *et al.* [5] have trained a single CNN to segment six tissues in MR images, the pectoral muscle in MR breast images, and the coronary arteries. Alakwaa *et al.* [6] have proposed an automatic method based on 3-D CNN for lung cancer detection through nodule segmentation from 3D CT scans. The proposed method reaches a test set Accuracy of 86.6%. Vardhana *et al.* [7] have presented CNN for biomedical image segmentation with hardware acceleration. The proposed design was applied for various Computed Tomography (CT) scan images of lungs, MR images of brain and X-ray images. Xie *et al.* [8] have proposed a novel framework that efficiently bridges a CNN and a Transformer (CoTr) for accurate 3D medical image segmentation. The proposed method was evaluated on the Multi-Atlas Labeling Beyond the Cranial Vault (BCV) dataset that covers 11 major human organs. A deep fully CNN was proposed in [9] for both left and right ventricle segmentation in short-axis MRI. For the first time, as mentioned, the neural network was trained end-to-end in a single learning stage from the whole image inputs and ground truths to make inference at every pixel. It is worth to note that CNN mainly focus on classification tasks, take an image as input and produce a unique class label [10]. However, biomedical images segmentation requires classification of each pixel in the image. In 2015, Ronneberger *et al.* [11], have proposed a new architecture called U-Net which is the most important semantic segmentation frameworks of CNN and works well for pixel-level prediction tasks. In the literature, several methods based on U-Net and its variants have been proposed to achieve automatic medical image segmentation. Hansch *et al.* [12] have studied the application of 2D and 3D U-net network in segmentation in order to adapt the neural network to the diversity of certain organ shapes and the low contrast of structures that surround the region of interest (ROI). The combination of these two methods is likely to solve these problems. Since the invention of U-Net, many variants have built upon that. For example, Zhou *et al.* [13], have proposed a more powerful architecture for medical image segmentation called U-Net ++. Compared to standard U-Net architecture, U-Net ++ is a neural network where the encoder and decoder sub-networks are connected through a series of nested, dense skip pathways which aim to reduce the semantic gap between the feature maps of the encoder and decoder sub-networks. In other research work [14], another modified U-Net architecture was used for automated left and right ventricular chamber segmentation in cardiac MRI images using dense FCN. U-Net has also been improved to have another variant called "Omega-Net" [15] which is a novel CNN architecture for simultaneous localization, transformation into a canonical orientation and semantic segmentation. In 2021, Chen *et al.* [16] have proposed a new variant of U-Net architecture called TransUNet. This latter employed Transformers and U-Net as a strong alternative for medical image segmentation to overcome U-Net limitations in explicitly modeling long-range dependency due to the intrinsic locality of convolution operations. In other research [17], 2D U-Net was used for the segmentation of the cartilage and the meniscus. Tao *et al.* [18], have proposed a fully automatic method for left ventricle (LV) quantification and segmentation from Cine -MRI images. Three CNNs with U-Net were trained on specific databases and tested on an independent multivendor and multicenter data set of 196 patients. This research frame is organized as follows. The proposed methodology is explained in Section.2 including data description and augmentation and the adopted architecture for the heart segmentation. In section.3, a comparative study of the obtained results with advanced state-of-the art methods is detailed.

2. Proposed methodology

2.1 Method overview

The proposed approach consists of four steps: (i) data augmentation, (ii) segmentation of the left and right ventricle (RV) based on U-Net architecture, (iii) Quantitative evaluation of the adopted U-Net performance with and without data augmentation and (iv) comparative study of CNN and U-Net performance for medical image segmentation. The overview of the proposed methodology is described in Fig.1

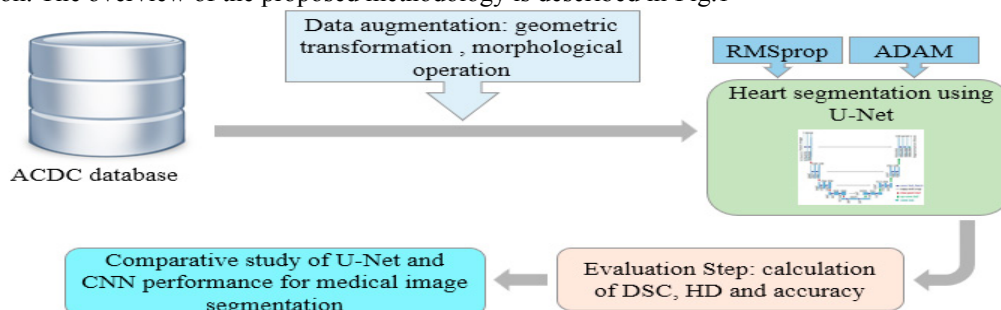


Fig.1 overview of the proposed methodology

2.2 Data description

To train and test the U-Net neural network, we used the Automated Cardiac Diagnosis Challenge (ACDC) database (MICCAI challenge, 2017) [19]. ACDC dataset contains Cine-MRI images of 150 patients acquired in the short-axis orientation. It is composed of two datasets: The training dataset which is composed of 100 patients and the testing dataset which is composed of 50 patients. For all patients of the training database, clinical information and corresponding labelled masks were provided. In our study, we converted the 3D-images already stored in NIFTI format to 2D-DICOM format using an open-source toolkit called X-MedCon. Fig.2 shows an example of the ACDC database (healthy patients) before and after conversion.

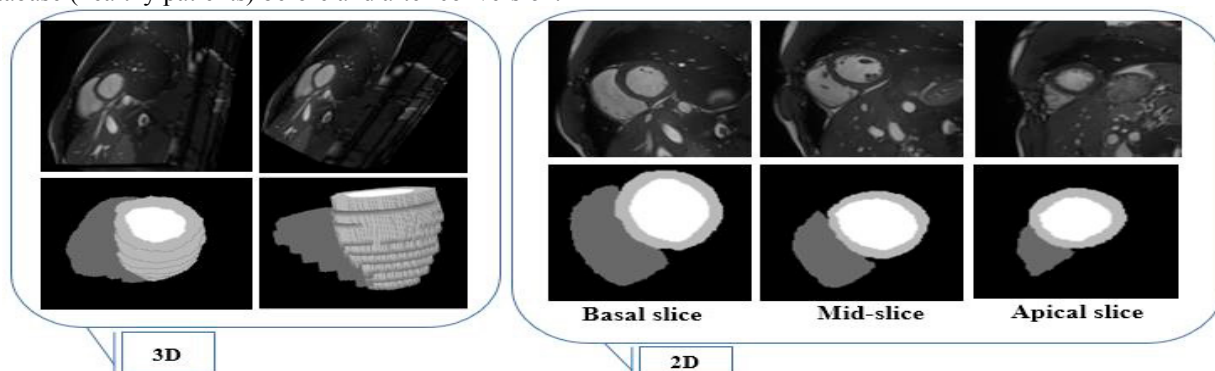


Fig. 2. MR images' visualization from ACDC dataset before (3D) and after(2D) conversion with corresponding labeled masks

2.3 Data augmentation

Based on the literature, Deep Learning techniques have greatly improved classification and segmentation accuracy thanks to their capability to handle complex conditions [20]. To achieve a good generalization performance, DL models rely heavily on big data. However, for many application areas, such as medical analysis, access to large database is limited as the data annotation is expensive and often requires significant amounts of expertise [21]. In This context, the most popular technique largely used to increase the size of the training dataset is data augmentation. In our study, we adopted some data augmentation techniques such as geometric transformations

(rotation, translation, reflection) and opening and closing morphological operation using an open-source Python library called TorchIO. The size of the generated augmented dataset is four time larger than the size of the original training data. In Fig.3, we present an example of data augmentation techniques applied to a healthy subject.

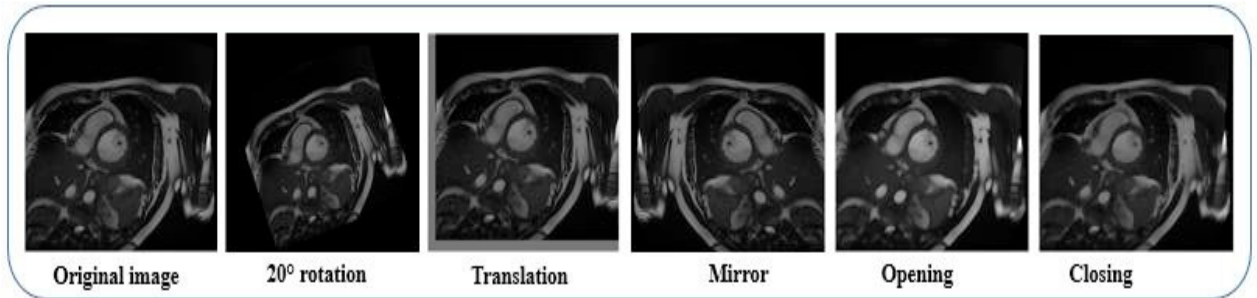


Fig.3 Data augmentation techniques applied to a healthy subject

2.4 U-Net architecture for the heart segmentation

As previously mentioned, we will perform the automatic segmentation of the heart from 2-D Cine-MRI images using U-Net architecture proposed by O. Ronneberger *et al.* [11]. The neural network is composed of the contracting path (left part) and the expansive path (right part). The first one consists in extracting the input image features. Each block of this path comprises two convolutional layers followed by Rectified Linear Unit (ReLU) activation functions and a max-pooling layer. In our study, the max-pooling operations are replaced by 2*2 Average Pooling operations to remove noise and restore main information. The second path consists in passing the feature map in a succession of blocks composed of two 3*3 convolution layers and a 2*2 up sampling layer (Up-convolution) while concatenating it with the feature map of the same level of the encoding path. It decompresses the image until its original size is reached since we used zero padding to keep the same size of the input image. The last feature map computed by the last 1*1 convolutional layer is converted to probabilistic segmentations of the foreground and background regions using soft-max pixel-wise. The two paths are linked by "skip connections", their role is to keep the information captured during the phase of features extraction.

2.5 Evaluation metrics

To evaluate the neural network performance, three evaluation metrics are calculated: the Dice Similarity Coefficient (DSC), the Hausdorff Distance (HD) [22] and accuracy calculated respectively as follows:

$$DSC = \frac{2TP}{2TP+FP+FN} * 100 \quad (1)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (2)$$

$$HD = \max (\max_{p \in C_A} d(p, C_B), \max_{q \in C_B} d(q, C_A)) \quad (3)$$

Where TP indicates an outcome where the model correctly predicts the positive class, TN indicates an outcome where the model correctly predicts the negative class, FP indicates an outcome where the model incorrectly indicates the positive class and FN indicates an outcome where the model incorrectly indicates the negative class and $d(p, C)$ is the minimum distance which separates the point p of contour C .

3. Experimental results and discussion

The proposed method was trained and tested using python version (3.7). All experiments are conducted using a single Nvidia RTX2080Ti GPU. For training, we used the Dice loss function which was minimized using Adam optimizer [23] and RMSProp. At the beginning of the experiments, we selected the hyper-parameters used for the training of the U-Net architecture. Then, our neural network was trained to segment both of LV and RV.

4.1 Hyper-parameters selection

In this sub-section, we present the experiments carried out to select hyper-parameters for training phase such as learning rate (lr), epochs and optimization algorithms. We started with a lr of $1 \cdot 10^{-2}$ then we repeat the experiments with a lr of $1 \cdot 10^{-3}$. In Table.1, we present the algorithm’s performance for a lr of 0.01 and 0.001 with mini-batch size of 4 over 10, 20, 50 and 80 epochs.

Table1. Comparison between optimization algorithms’ performance during training phase

	Epochs	ADAM %	RMSProp %
Initial learning rate = $1 \cdot 10^{-2}$	10	54.38	51.18
	20	61.49	56.34
	50	70.83	63.22
	80	76.15	50.72
learning rate = $1 \cdot 10^{-3}$	10	77.43	71.54
	20	83.17	84.92
	50	74.52	71.15
	80	92.77	87.97

As shown in Table.1, our model achieved the highest performance with ADAM optimizer (92.77%) and was trained for 80 epochs which were enough for convergence with a lr of 0.001 and a mini-batch size of 4. Fig.4 presents the mini-batch accuracy and mini-batch Dice loss during training phase with and without data augmentation.

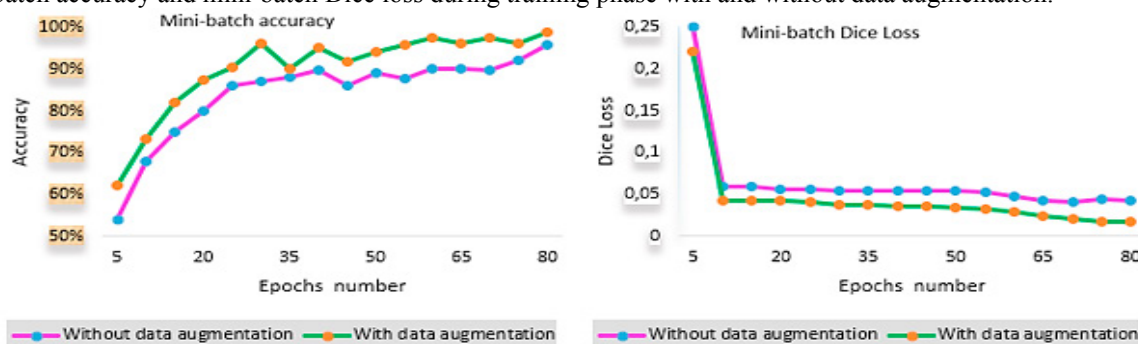


Fig.4 Mini-batch accuracy and mini-batch Dice loss during training phase with and without data augmentation

By increasing database as shown in Fig.4, the accuracy was improved from 96.1% to 98% during training and the loss function decreases from 0.042 to 0.018. This highlights the importance of the data augmentation step.

4.2 Neural network performance for left and right ventricle segmentation

In this sub-section, we present the automatic segmentation results of the LV and RV. Fig.5 presents segmentation results for two patients with data augmentation and Table.2 presents the corresponding quantitative results.

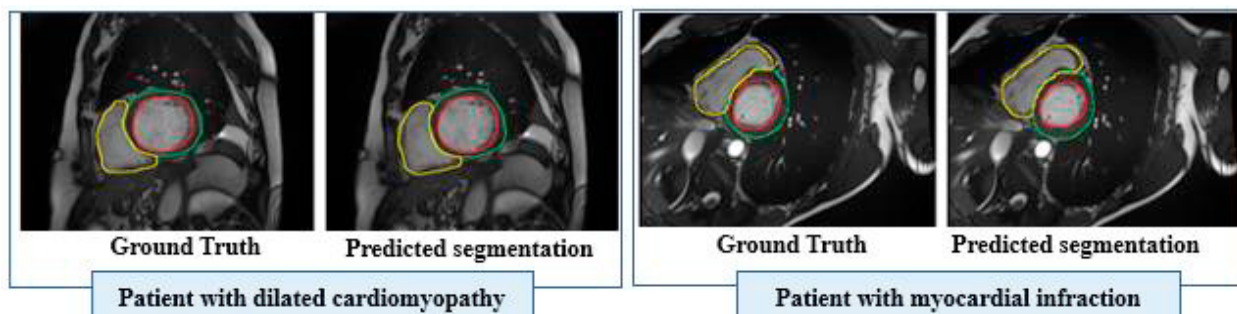


Fig.5. Automatic segmentation results for two patients with dilated cardiomyopathy and myocardial infarction respectively

Table .2. Segmentation performance of the proposed model at ED and ES phase measured as DSC, HD and accuracy

	LV						RV					
	DSC %		HD (mm)		Accuracy %		DSC %		HD (mm)		Accuracy %	
	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES
Patient with dilated cardiomyopathy	98.81	93,62	6.134	8.621	97,88	96.67	95,97	93,72	9.678	10.167	96,75	87,54
Patient with myocardial infraction	94.84	89,28	8.861	10.212	93,77	91,12	97,3	96,99	8.951	9.321	96,41	96.34

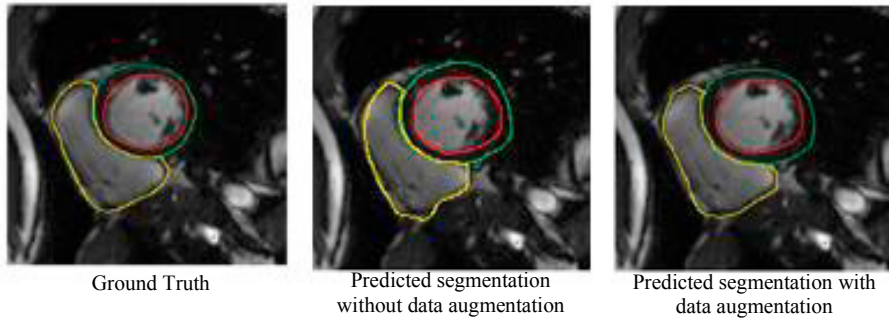


Fig.6. Automatic segmentation results for a patient with abnormal right ventricle

In Fig.6, the analysis of the segmented images shows a slight overestimation of the RV. When applying data augmentation techniques, the DSC was improved from 97.56% to 98.79%, the HD from 9.125 mm to 7.881mm and the accuracy from 96.87% to 98.34%. Fig.7 presents the segmentation results applied to four successive acquisitions of the apical slice for a patient with hypertrophic cardiomyopathy. Quantitative analysis is presented in Table.3.

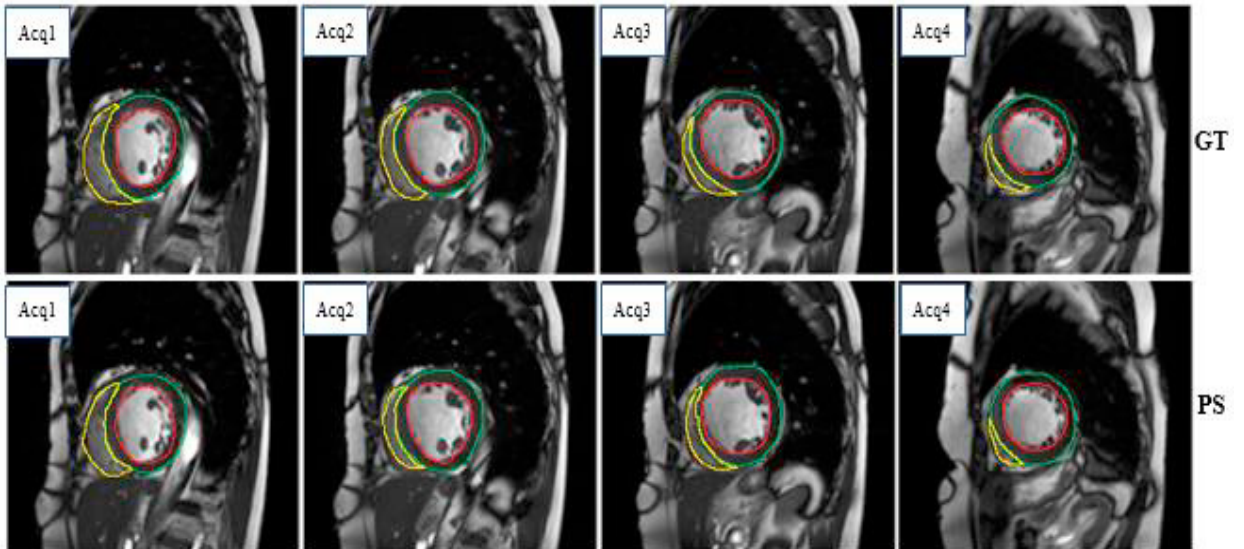


Fig. 7: Automatic contours extraction using U-Net applied on apical slices (GT: ground truth), (PS: predicted segmentation)

Table 3. Segmentation performance of the proposed model at the ED and ES phase measured as DSC, HD and accuracy applied to apical slices

	LV						RV					
	Mean DSC %		Mean HD (mm)		Mean Accuracy %		Mean DSC %		Mean HD (mm)		Mean Accuracy %	
	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES	ED	ES
Patient with hypertrophic cardiomyopathy	93.02	89,15	10.421	9.561	92,87	91.44	94,47	92,15	9.932	9.751	92,56	91,8

In Fig.7, the visual analysis of the automatically segmented images shows an underestimation of the contours, especially for the right ventricle at the level of acquisitions 2 and 4. Quantitative analysis, presented in Table.3, shows that contour extraction in apical slices was less accurate compared with basal and middle slices. This can be explained by the fact that this cut is characterized by a difficulty of segmentation due to the low contrast between the myocardium and the blood.

4.3 Comparative study with state-of-the art methods

Compared to the standard U-Net architecture, the proposed method reached an accuracy of 98.79% in terms of DSC against 78.75%. In terms of training time, it was mentioned that the standard U-Net architecture took about 10 hours while our model took almost 2h 40 min. The segmentation results proved that the proposed method is characterized by better performance compared to the standard model. To better evaluate the model performance, we made a quantitative comparison with state-of-the-art methods (Table.4).

Table.4. Performance in the MICCAI2017 test set.

Year	Database	LV				RV				
		Mean DSC (%)		Mean HD (mm)		Mean DSC (%)		Mean HD (mm)		
		ED	ES	ED	ES	ED	ES	ED	ES	
Isensee et al	2018	ACDC	96.7	92.8	5.476	6.921	95.1	90.4	8.205	11.605
Zotti et al	2019	ACDC	96.4	91.2	6.180	8.386	93.4	88.5	11.052	12.650
Zhongyu et al	2019	ACDC	94.2		6.641		90.3		13.830	
Calisto et al	2020	ACDC	95.8	90.3	5.592	8.644	93.6	88.4	10.193	12.234
Simantiris et al	2020	ACDC	96.7	92.8	6.366	7.573	93.68	88.9	13.289	14.367
Our method	2022	ACDC	97.9	94.7	5.318	7.812	95.75	92.81	8.369	10.234

Comparison with the methods having used the same database show that the proposed method reached the highest performance with a mean DSC of 97.9 % and a mean HD of 5.318 mm for LV segmentation. However, the method proposed by Isensee et al [25] reached the smallest HD (8.205mm) at the ED phase for the RV segmentation. A second quantitative comparison was made with methods based on CNNs to demonstrate which of these two models between U-Net and CNN is more suitable for medical segmentation. Results are summarized in Table.5.

Table.5. Performance of CNN and U-Net for medical image segmentation

Model	Application	Dice coefficient (%)
U-Net	Cardiac images segmentation	93
CNN	LV segmentation	91
CNN with U-Net	LV segmentation	98
FCN	Colorectal tumour segmentation	83.5
Vanilla U-Net	Mass segmentation in mammography	91
Faster RCNN	Medical image segmentation	70.2
3D-FCN	Pancreas segmentation	85.5

In the literature, it has been demonstrated that neural networks based on encoder-decoder model such as U-Net are characterized by their performance in detecting smaller anatomical structures with blurred noisy boundaries unlike CNN which imposes restrictions on the size of the input image [26] as not to complicate the calculation at the level of the fully connected layer. Even quantitatively U-Net exceeds CNN in terms of DSC in medical segmentation tasks as shown in Table.5. Thus, U-Net is more suitable for medical image segmentation.

5. Conclusion

In this paper, as part of the comparison between U-Net and CNN performance for medical image segmentation, U-Net model was adopted for left and right ventricle segmentation from short-axis MRI images with and without data augmentation. A Comparative study with the state-of-the-art methods based on U-Net and CNN architectures shows that the proposed method has achieved the highest performance in terms of dice similarity coefficient,

Hausdorff distance and accuracy. This is a significant step towards reducing manual segmentation limitations performed by clinical experts during the diagnosis of cardiac diseases. Comparative study with CNNs models shows also that U-Net is more appropriate for medical segmentation tasks.

References

- [1] Du, G., Cao, X., Liang, J., Chen, X., & Zhan, Y. (2020). Medical image segmentation based on u-net: A review. *Journal of Imaging Science and Technology*, 64(2), 20508-1.
- [2] Ummadi, V. (2022). U-Net and its variants for Medical Image Segmentation: A short review. *arXiv preprint arXiv:2204.08470*.
- [3] Jena, M., Mishra, S. P., & Mishra, D. (2018). A survey on applications of machine learning techniques for medical image segmentation. *International Journal of Engineering & Technology*, 7(4), 4489-4495.
- [4] Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. *Journal of big data*, 2(1), 1-21.
- [5] Moeskops, P., Wolterink, J. M., van der Velden, B. H., Gilhuijs, K. G., Leiner, T., Viergever, M. A., & Išgum, I. (2016, October). Deep learning for multi-task medical image segmentation in multiple modalities. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 478-486). Springer, Cham.
- [6] Alakwaa, W., Nassef, M., & Badr, A. (2017). Lung cancer detection and classification with 3D convolutional neural network (3D-CNN). *Lung Cancer*, 8(8), 409.
- [7] Vardhana, M., Arunkumar, N., Lasrado, S., Abdulhay, E., & Ramirez-Gonzalez, G. (2018). Convolutional neural network for bio-medical image segmentation with hardware acceleration. *Cognitive Systems Research*, 50, 10-14.
- [8] Xie, Y., Zhang, J., Shen, C., & Xia, Y. (2021, September). Cotr: Efficiently bridging cnn and transformer for 3d medical image segmentation. In *International conference on medical image computing and computer-assisted intervention* (pp. 171-180). Springer, Cham.
- [9] Tran, P. V. (2016). A fully convolutional neural network for cardiac segmentation in short-axis MRI. *arXiv preprint arXiv:1604.00494*.
- [10] Volpi, M., & Tuia, D. (2016). Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 881-893.
- [11] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [12] Hänsch, A., Schwier, M., Gass, T., Morgas, T., Haas, B., Dicken, V., ... & Hahn, H. K. (2018). Evaluation of deep learning methods for parotid gland segmentation from CT images. *Journal of Medical Imaging*, 6(1), 011005.
- [13] Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 3-11). Springer, Cham.
- [14] Penso, M., Moccia, S., Scafuri, S., Muscogiuri, G., Pontone, G., Pepi, M., & Caiani, E. G. (2021). Automated left and right ventricular chamber segmentation in cardiac magnetic resonance images using dense fully convolutional neural network. *Computer Methods and Programs in Biomedicine*, 204, 106059.
- [15] Vigneault, D. M., Xie, W., Ho, C. Y., Bluemke, D. A., & Noble, J. A. (2018). Ω -net (omega-net): fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks. *Medical image analysis*, 48, 95-106.
- [16] Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., & Zhou, Y. (2021). Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*.
- [17] Padoia, V., Norman, B., Mehany, S. N., Bucknor, M. D., Link, T. M., & Majumdar, S. (2019). 3D convolutional neural networks for detection and severity staging of meniscus and PFJ cartilage morphological degenerative changes in osteoarthritis and anterior cruciate ligament subjects. *Journal of Magnetic Resonance Imaging*, 49(2), 400-410.
- [18] Tao, Q., Yan, W., Wang, Y., Paiman, E. H., Shamonin, D. P., Garg, P., ... & Van Der Geest, R. J. (2019). Deep learning-based method for fully automatic quantification of left ventricle function from cine MR images: a multivendor, multicenter study. *Radiology*, 290(1), 81-88.
- [19] Bernard, O., Lalonde, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P. A., ... & Jodoin, P. M. (2018). 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.
- [20] Hesamian, M. H., Jia, W., He, X., & Kennedy, P. (2019). Deep learning techniques for medical image segmentation: achievements and challenges. *Journal of digital imaging*, 32(4), 582-596.
- [21] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- [22] Mathews, C., & Mohamed, A. (2020, January). Review of Automatic Segmentation of MRI Based Brain Tumour Using U-Net Architecture. In *2020 Fourth International Conference on Inventive Systems and Control (ICISC)* (pp. 46-50). IEEE.
- [23] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [24] Gorter, T. M., van Veldhuisen, D. J., Bauersachs, J., Borlaug, B. A., Celutkiene, J., Coats, A. J., ... & de Boer, R. A. (2018). Right heart dysfunction and failure in heart failure with preserved ejection fraction: mechanisms and management. Position statement on behalf of the Heart Failure Association of the European Society of Cardiology. *European journal of heart failure*, 20(1), 16-37.
- [25] Isensee, F., Jaeger, P. F., Full, P. M., Wolf, I., Engelhardt, S., & Maier-Hein, K. H. (2017, September). Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. In *International workshop on statistical atlases and computational models of the heart* (pp. 120-129). Springer, Cham.
- [26] Zhang, P., Zhong, Y., & Li, X. (2020). ACCL: Adversarial constrained-CNN loss for weakly supervised medical image segmentation. *arXiv preprint arXiv:2005.00328*.