

Article

Supervised Domain Adaptation for Automated Semantic Segmentation of the Atrial Cavity

Marta Saiz-Vivó¹, Adrián Colomer^{1,*} , Carles Fonfría², Luis Martí-Bonmatí^{2,3}  and Valery Naranjo¹

¹ Instituto de Investigación e Innovación en Bioingeniería, Universitat Politècnica de València, 46022 Valencia, Spain; marsaivi@etsii.upv.es (M.S.-V.); vnanranjo@i3b.upv.es (V.N.)

² Radiology Department, La Fe University and Polytechnic Hospital, 46026 Valencia, Spain; fonfría_car@gva.es (C.F.); marti_lui@gva.es (L.M.-B.)

³ Biomedical Imaging Research Group (GIBI230-PREBI), La Fe Health Research Institute, 46026 Valencia, Spain

* Correspondence: adcogra@i3b.upv.es

Abstract: Atrial fibrillation (AF) is the most common cardiac arrhythmia. At present, cardiac ablation is the main treatment procedure for AF. To guide and plan this procedure, it is essential for clinicians to obtain patient-specific 3D geometrical models of the atria. For this, there is an interest in automatic image segmentation algorithms, such as deep learning (DL) methods, as opposed to manual segmentation, an error-prone and time-consuming method. However, to optimize DL algorithms, many annotated examples are required, increasing acquisition costs. The aim of this work is to develop automatic and high-performance computational models for left and right atrium (LA and RA) segmentation from a few labelled MRI volumetric images with a 3D Dual U-Net algorithm. For this, a supervised domain adaptation (SDA) method is introduced to infer knowledge from late gadolinium enhanced (LGE) MRI volumetric training samples (80 LA annotated samples) to a network trained with balanced steady-state free precession (bSSFP) MR images of limited number of annotations (19 RA and LA annotated samples). The resulting knowledge-transferred model SDA outperformed the same network trained from scratch in both RA (Dice equals 0.9160) and LA (Dice equals 0.8813) segmentation tasks.

Keywords: supervised domain adaptation; MRI sequences; atrial geometry; semantic segmentation



Citation: Saiz-Vivó, M.; Colomer, A.; Fonfría, C.; Martí-Bonmatí, L.; Naranjo, V. Supervised Domain Adaptation for Automated Semantic Segmentation of the Atrial Cavity. *Entropy* **2021**, *23*, 898. <https://doi.org/10.3390/e23070898>

Academic Editor: Raúl Alcaraz

Received: 13 May 2021

Accepted: 10 July 2021

Published: 14 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia. It has been estimated that 6–12 million people will suffer this condition in the US by 2050 and 17.9 million people in Europe by 2060. Furthermore, this pathology is highly associated with morbidity and mortality factors such as heart failure, ischemic and hemorrhagic strokes and provokes important economic burden [1]. As in-depth knowledge of the physiopathology of the disease remains lacking, the treatment of AF is a complex area.

At this moment, the most common method for treating AF patients is catheter ablation to produce scars and electrically isolate the pulmonary veins (PV) [2]. The success rate of this procedure is significantly higher in patients with paroxysmal AF than in patients with persistent AF which develops fibrotic tissue areas that contribute to the maintenance of AF. To improve the success rate in these patients, clinicians attempt to eliminate the AF driver's areas guided by complex navigating systems such as Carto or Navex [3].

Accurate segmentation of LA is highly desirable for patient specific scar characterization to select the most appropriate ablation strategy. Furthermore, to plan or guide the procedure in patients with persistent or paroxysmal AF, patient specific 3D electro-anatomical models of both the left and right atrium (LA and RA) are required. These models are obtained by integrating the electro-anatomical information provided by the navigating systems with a 3D geometrical reconstruction of the atria [4].

Notably, 3D patient-specific reconstructions are obtained from the segmentation of MRI and CT images. Cardiac MRI is considered the gold standard for cardiac chamber evaluation including LA and RA. Different image modalities include the balanced steady-state free precession (bSSFP) sequences and late gadolinium enhanced MRI (LGE-MRI). In clinical practice, bSSFP sequences are widely used for LA size characterization due to their excellent blood to myocardium contrast and high spatial resolution [5]. On the other hand, LGE-MRI techniques provide enhanced brightness in fibrotic or infarcted regions, and they are currently used to study the extent of fibrotic tissue in the atria [6]. Furthermore, recent LGE-MRI studies of the atria have led to important advances in understanding and reversing AF [7,8].

In clinical practice, cardiac MRI segmentation is performed manually by experts which is a time-consuming, labor-intensive and error-prone method. Therefore, there is a need for computational methods for automatic image segmentation. However, high segmentation accuracy remains a challenging task, especially for LA due to its complex geometry [9]. Furthermore, given the diversity of MRI sequences used in the clinical environment, the heterogeneous domain shift is a hot-topic research field [10]. For example, as shown in Figure 1, cardiac regions appear significantly different visually in images acquired using bSSFP and LGE-MRI cardiac acquisitions.

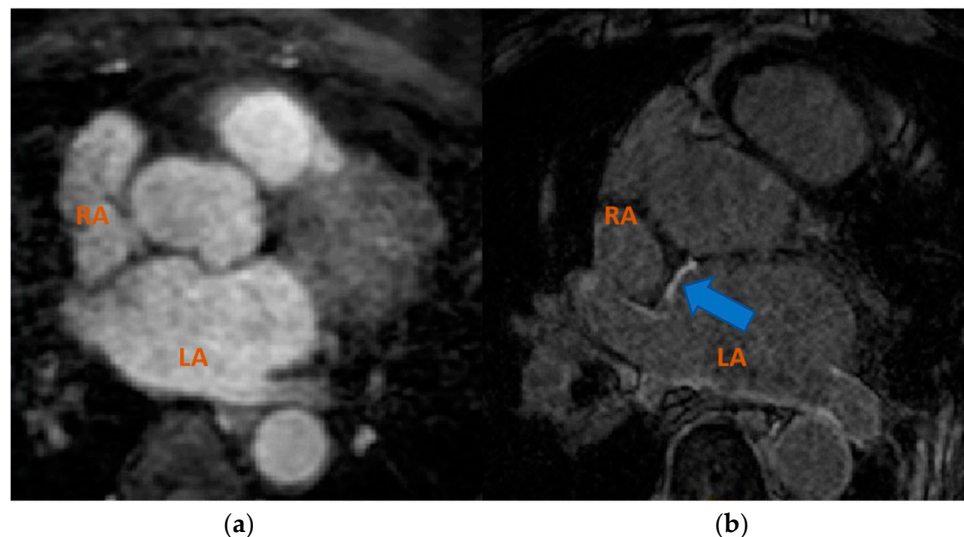


Figure 1. Axial view of different MRI sequences. (a) bSSFP and (b) LGE-MRI. The blue arrow points to the enhanced brightness of infarcted regions obtained in LGE-MRI. Atrial cavities are labelled as LA (left atrium) and RA (right atrium).

Deep learning (DL) techniques have recently gained popularity in atrial segmentation applications. However, the need for model re-training each time that an external database is provided supposes a significant limitation for their adoption in the clinical practice. Furthermore, it must be remarked upon that although most works focus on the LA segmentation task [11–13] due to its complex geometry and PV ablation strategies, the segmentation and reconstruction of the RA has an equally important role for the generation of electro-anatomical models that help to guide and plan ablation procedures, especially in patients with persistent AF [4]. Therefore, there is a growing importance in obtaining models that accurately segment both atrial cavities.

In this work, we present a novel supervised domain adaptation (SDA) framework for LA and RA segmentation tasks. Our approach transfers the knowledge acquired from an LGE-MRI source domain, in which a large, labelled database is available to a b-SSFP-MRI target domain characterized by just a few labelled samples.

The aim is to introduce feature transfer learning techniques in a two-stage 3D U-Net [14] to obtain high accuracy 3D MRI segmentation models capable of segmenting

the LA and RA in a target domain composed of few labelled examples. To accomplish this, the supervised domain adaptation technique is formulated in two different scenarios: “different domain, same task” and “different domain, different task” occurrences. Note that this paper extends the preliminary results obtained in [15]. Compared to [15], in this work, we formally formulate the proposed SDA framework, and we extend the validation of our methodology, including an external bSSFP composed of 19 patients with RA annotations. These annotations were manually carried out by the radiologists involved in this work (see Data Availability statement).

The remainder of the paper is organized as follows: In Section 2, we provide an overview of automatic medical image segmentation methods and supervised domain adaptation techniques. In Section 3, the proposed atrial segmentation framework is presented by describing the materials and methodology used. In Section 4, the extensive experiments to validate the proposed approach are described, and results are presented. In Section 5, we provide a discussion of our results, summarize the contributions of this paper and provide some future research directions.

2. Related Work

2.1. Automatic Medical Image Segmentation

In the field of medical image segmentation, early automated algorithms consisting of thresholding methods and region growing approaches, such as watershed methods, have been widely used. For example, Huang et al. [16] applied watershed algorithms for automatic contouring of breast tumors in US images, and more recently, Masoumi et al. [17] applied watershed algorithms in MRI images with artificial neural networks for liver segmentation. However, the performance of these methods is highly sensitive to boundary contrast and image noise.

By contrast, other region-based methods such as active contours have demonstrated good performance in challenging segmentation problems [18]. For example, Zhao et al. [18] applied active contours with different types of region information for blood vessel segmentation achieving promising results. Ciecholewski et al. [19] used active contours for semi-automatic segmentation of corpus callosum in MR images.

More recently, superpixel-based methods have been proposed to improve the efficiency and robustness of the segmentation procedures. For example, Tian et al. [20] presented a superpixel-based graph-cut model which is iteratively used along with an active contour model to obtain smoother segmentations in prostate MR images. Nguyen et al. [21] applied superpixel methods with multi-atlas segmentation to optimize the image registration process. The above methods are characterized for providing an alternative to machine learning (ML)-based segmentation techniques.

Furthermore, multi-atlas-based segmentation and, more recently, ML algorithms such as support vector machines as statistical classifiers, have also gained interest in the field of cardiac image segmentation and classification. Despite the promising results shown by both the non-ML and ML methods mentioned above, their ad hoc nature and reliability on good initialization have limited their widespread adoption in clinical practice [22].

In recent years, the development of hardware for higher computational power and the growth of clinical databases has enabled the expansion of machine learning (ML) methods into deep learning (DL) algorithms, capable of automatic feature learning and extraction, in the field of image classification and segmentation. Convolutional neural networks (CNNs) with encoder-decoder architectures such as U-Net [23] are widely used in medical image segmentation. Furthermore, 3D extensions of the U-Net [24] have gained popularity as they provide better spatial representation, fully exploiting inter-slice continuity during training [14]. In atrial segmentation tasks, several 3D U-Net variants have been proposed achieving very promising results. The authors of [11] extended 3D U-Net with hierarchical aggregation to obtain better spatial fusion information. The authors of [12] proposed a 3D U-Net with dilated convolutions in the lowest level of the network to extract features spanning a wider spatial range. Furthermore, they cropped the input images around the region of

interest as a pre-processing step. For the STACOM 2018 Atrial Segmentation Challenge [25], the authors of [26,27] proposed a 3D U-Net with a dual strategy for localization and segmentation of LA cavity achieving high-accuracy results. More recently, the authors of [13] extended the 3D Dual U-Net strategy by predicting density maps for LA localization with promising results. However, the aforementioned methods [13,26,27] were trained with a large and labelled dataset (the STACOM 2018 challenge dataset), and the authors of [26,27] do not validate their results with an external dataset.

2.2. Supervised Domain Adaptation

Given a source domain D_S with a corresponding source task T_S and a target domain D_T with a corresponding task T_T , transfer learning is the process of improving the target predictive function $f_T(\cdot)$ by using the related information from D_S and T_S , where $D_S \neq D_T$ or $T_S \neq T_T$ [28]. Image domains are expressed with a feature space S and a probability distribution $P(x)$ where $X = \{x_1, x_2, \dots, x_n\} \in S$. In this context, domain adaptation (DA) refers to the “different domain, same or related task” scenario where $D_S \neq D_T$ due to differences in the samples distribution $P(x)$, for example, when data are acquired from different scanning protocols [29]. The aim of DA is to leverage knowledge learned on a source domain normally consisting of large amounts of data and apply it to a different but related target domain with few labelled data [30]. This work falls in the category of supervised domain adaptation (SDA). Given a source training dataset with pairs $D_S = \{(x_i^S, y_i^S)\}_{i=1}^N$ with random variable $x_i^S \in X^S$ and label $y_i^S \in Y^S$ and a target training dataset $D_T = \{(x_i^T, y_i^T)\}_{i=1}^M$ with $x_i^T \in X^T$ and label $y_i^T \in Y^T$, the goal of SDA is to improve $f_T(\cdot)$ where $M \ll N$ and $D_S \neq D_T$ [31].

Many DL applications require large amount of labelled data representative of the target task for their optimization. In medical image segmentation, large amounts of labelled target data are often not available or come with high-acquisition and labelling costs. In this case, the typical approach is to use the available datasets (source data) representative of a closely related task but often of different domain due to different scanners or acquisition protocols. SDA techniques were developed to improve the performance of this approach [31].

DA is a widely used technique in transfer-learning scenarios for medical image applications due to its ability to transfer knowledge from different image domains. A comparison of DA approaches for medical image classification and segmentation application can be observed in Table 1. Existing DA approaches include instance transfer and weighting DA strategies where the goal is to align the source and target domains by a transformation of the feature space [32–34]. However, these strategies are limited to the “same task” scenario application. With the development of deep-learning methods, feature transfer DA approaches have become increasingly popular due to their ability to transfer information from different domains and different tasks. This is achieved by pre-training a network with a large and available source dataset and fine-tuning the network for a target task [29]. Several works used this approach for image classification tasks [35,36].

More recently, DA approaches have been also applied to medical image segmentation (pixel-wise classification) tasks achieving promising results. The authors of [37] explored unsupervised DA (UDA) with adversarial networks for brain segmentation achieving promising results; however, adversarial training requires subtle parameter tuning. Furthermore, UDA approaches are consistently outperformed with SDA methods which take advantage of a few labelled target samples. Existing works focusing on feature transfer-SDA for medical image segmentation are limited and mostly focus on brain segmentation. For example, Reference [38] explores SDA methods for sub-cortical brain structure segmentation by comparing pre-trained networks fine-tuned with small target samples with fully trained networks and achieving similar accuracy results. Reference [39] compares the performance for brain lesion segmentation in a domain-adapted network and a network trained with the same examples “from scratch” achieving significantly higher accuracy results in the domain adapted network. In the field of cardiac MRI segmentation, the authors of [10] pre-trained a network with bSSFP MR images and fine-tuned the resulting

models for LGE-MRI target domain with a few target samples for ventricular segmentation task outperforming the same network without DA. However, in this work, the multi-class segmentation task is kept the same across the different domain instances. As can be observed in Table 1, previous segmentation works focusing on feature transfer with network pre-training have limited their application to same source and target task scenario.

Table 1. Overview of DA applications. UDA: unsupervised domain adaptation, SDA: supervised domain adaptation. Transfer type column refers to DA transfer technique, i.e., feature adaptation, instance transfer (weighting or aligning samples), or feature transfer (pre-training on same or different auxiliary task).

First Author	Application	Task	Transfer Type	Pros/Cons
UDA				
Kamnistas et al. [37]	Segmentation of abnormalities	Same	Feature adaptation	No labels/subtle parameter tuning
SDA				
Conjeti et al. [32]	Tissue classification	Same	Instance transfer, align	Considers local and global domain/only align similar features
Wachinger et al. [34]	AD classification	Same	Instance transfer, weights	Good accuracy with small target labels/weight selection
Menegola et al. [35]	Melanoma classification	Different	Feature transfer, pre-training	Fine-tuned outperforms baseline models/relies on size of source dataset
Ribeiro et al. [36]	Polyp classification	Different	Feature transfer, pre-training	Comparative study/conclusions not always hold
Kushibar et al. [38]	Brain segmentation	Same	Feature transfer, pre-training	Fine-tune with small target sample/only applied to same task
Ghafoorian et al. [39]	Brain lesion segmentation	Same	Feature transfer, pre-training	Last dense layers fine-tuning/only applied to same task
Vesal et al. [10]	Ventricle segmentation	Same	Feature transfer, pre-training	Good results in multi-class segmentation/only applied to same task

To the best of the authors' knowledge, no previous works have explored the field of SDA for both LA (T_S) and RA (T_T) atrial segmentation tasks. Inspired by this, our purpose here is to formulate and give a solution to the two-fold SDA scenarios, i.e., $D_S \neq D_T$ and $T_S = T_T$ and $D_S = D_T$ and $T_S \neq T_T$, improving the ability of generalization of atrial segmentation models making steps forward in the adoption of computer-aided diagnosis systems based on deep learning in the clinical practice.

3. Materials and Methodology

3.1. Materials

Three public databases coming from different sources were used to develop and evaluate the proposed SDA framework for LA and RA segmentation.

Database 1 was a result of the STACOM 2018 Atrial Segmentation Challenge. It contains 100 3D LGE-MRIs images from patients with AF and their corresponding LA binary segmentations. The data resolution is $0.625 \times 0.625 \times 0.625 \text{ mm}^3$. All volumes are characterized by the same number of slices (88), and there is little variation in terms of image contrast among samples [25].

Database 2 was publicly released by the Multi-Modality Whole Heart Segmentation (MM-WHS) challenge, in conjunction with MICCAI 2017 [40]. It contains 20 volumetric MRI images from human patients with a data resolution of $0.8\text{--}1 \times 0.81 \times 1\text{--}1.6 \text{ mm}^3$ and the corresponding binary segmentation masks of both RA and LA. All volumes had between 256 and 512 slices each. In this work 19 volumetric images were used due to quality controls. This dataset was acquired with a navigator-gated 3D b-SSFP sequence; therefore, it is considered as a dataset of different domain to Database 1 [41].

Database 3 was provided by STACOM 2013 Left Atrial Segmentation Challenge in collaboration with MICCAI13 [42]. It contains 30 volumetric MRI images also acquired using 3D bSSFP sequence with a voxel resolution of $1.25 \times 1.25 \times 2.7 \text{ mm}^3$ and the corresponding ground-truth binary segmentation of the LA cavity. All volumes had between 300 and 400 slices each. This dataset contains a variety of quality levels in the following proportions: 19 high and moderate quality and 11 with local artefacts and high noise [43]. In this work, a quality control protocol was applied selecting the 19 patients with good and moderate quality.

In addition, to further extend the validation of the RA segmentation models, a RA ground-truth was generated from this database. The RA ground-truth was obtained by manual segmentation performed by a radiologist (CF) (see Data Availability statement).

3.2. Methodology

3.2.1. Domain Adaptation

The scarcity of large and pixel-wise labelled datasets is one of the main problems of deep-learning applications in the field of medical image segmentation. Domain adaptation paradigm enables the transfer of knowledge from a large and labelled source dataset to a target dataset of few examples and different domain [29]. In this work, an SDA technique is formulated for two different scenarios. On the one hand, the segmentation of LA cavity in the sample-constrained bSSFP domain (i.e., Database 2) is faced by transferring the knowledge acquired on the 100 labelled LA instances from the source LGE-MRI domain (i.e., Database 1) under a “different domain, same task” premise. On the other hand, a “different domain, different task” scheme is proposed taking profit from the knowledge encoded during the LA segmentation task on the source LGE-MRI domain (i.e., Database 1) and transferring it to solve a RA segmentation task in a different bSSFP target domain (i.e., Database 2).

Let θ_{LA}^{LGE} be a source model associated with $D_S \equiv D_{LGE} = \{S^S, P(X^S)\}$ as LGE-MRI image domain with feature space S^S and probability distribution of $P(X^S)$ where $X^S = \{x^S_1, x^S_2, \dots, x^S_n\} \in S^S$, T_S a LA segmentation learning task and Y^S label space of binary LA segmentation masks. In supervised learning, θ_{LA}^{LGE} can be optimized by using a pair of samples $\{x^S_i, y^S_i\}_{i=1}^N$ where $x^S_i \in X^S$ and $y^S_i \in Y^S$. After the learning process, the model can be used for solving the T_S task in D_{LGE} domain.

Inspired by the domain adaptation paradigm, we can define two target models, each responsible of solving different associated learning tasks (T_T), under $D_T \equiv D_{bSSFP} = \{S^T, P(X^T)\}$ image domain and a small label space Y^T with both RA and LA segmentation masks available. One model is tasked with LA segmentation (ψ_{LA}^{bSSFP}) and the other with RA segmentation (ψ_{RA}^{bSSFP}). The idea is to transfer the learned weights from θ_{LA}^{LGE} to improve the predictive function of θ_{LA}^{bSSFP} and θ_{RA}^{bSSFP} making use of the $\{x^T_i, y^T_i\}_{i=1}^M$ set of samples, where $M \ll N$. Note that in the first case we have defined an SDA scenario in which $D_S \neq D_T$ and $T_S = T_T$, whilst in the second, $D_S \neq D_T$ and $T_S \neq T_T$.

In this work, the proposed SDA framework was built upon the dual 3D U-Net proposed by [14] (see Section 3.2.2 for more details). First, this encoder–decoder CNN was pre-trained on the large, labelled source LGE-MRI dataset giving place to the θ_{LA}^{LGE} model.

The knowledge acquired on the source domain during the learning procedure was transferred, in a second training stage, to the target domain via weight initialization. In particular, the target encoder, bottleneck and decoder of the dual 3D U-Net were initialized to the resulting θ_{LA}^{LGE} weights. Then, the network was fine-tuned using the target bSSFP dataset of different domain and few labelled samples for both LA and RA tasks individually. The proposed SDA pipeline is illustrated in Figure 2, where the final bSSFP adapted models are denoted as ψ_{LA}^{bSSFP} and ψ_{RA}^{bSSFP} for LA and RA segmentation tasks, respectively. Note that for both segmentation tasks, the loss function, data augmentation scheme and hyperparameters were kept the same.

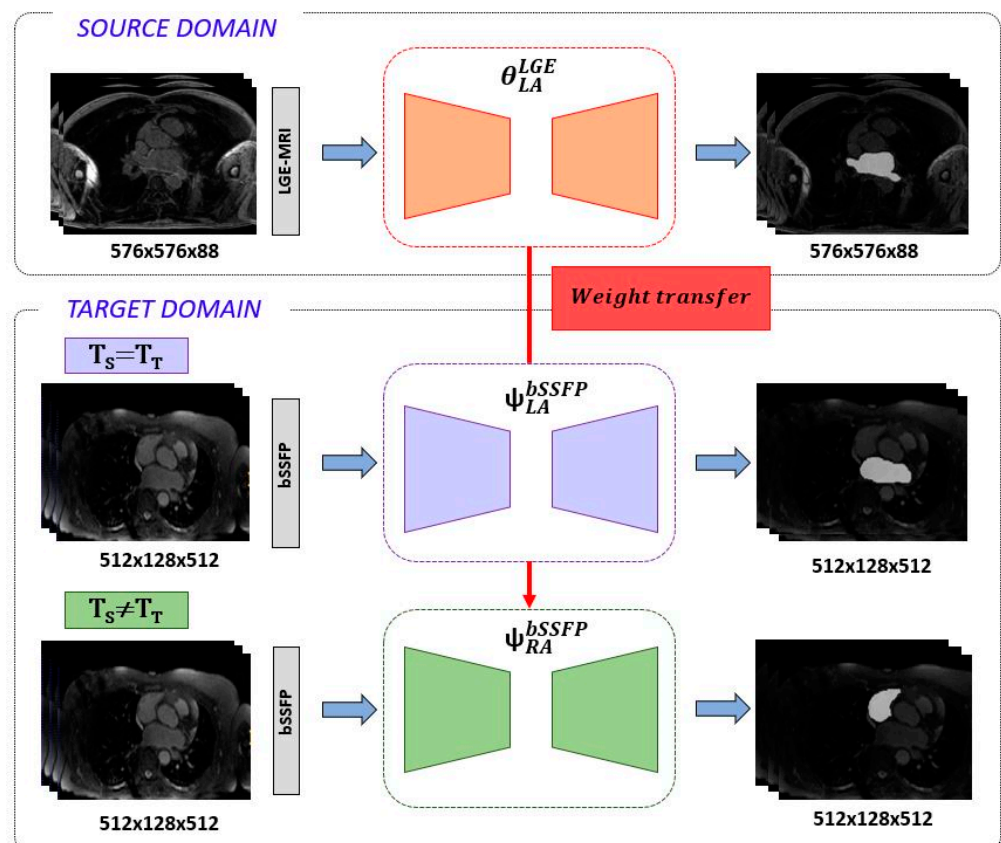


Figure 2. Overview of feature transfer framework. The 3D U-Net is first trained under the LGE-MRI source domain for exclusively LA segmentation. In a second stage, the network is retrained and fine-tuned on the bSSFP target domain for LA and RA segmentation tasks. The red arrow refers to the weight transfer from source model to target model for LA and RA segmentation tasks.

3.2.2. Dual 3D U-Net

As we mentioned in Section 3.2.1, the proposed SDA framework is built upon the dual 3D U-Net architecture proposed by [14] and illustrated in Figure 3. The 3D U-Net is a specialization of the net proposed by [24] and the implementation follows the work of [44]. It consists of a two-stage network. In the first stage, the network is tasked with locating the region of interest (in this work LA/RA). This preliminary segmentation output is then passed through a processing function that computes the spatial location of the atria and crops the image and masks with a cuboid centered around the region of interest. In the second stage, the training is performed with the cropped images at full resolution. This two-stage approach removes background noise and enables for a more precise segmentation.

The 3D U-Net is a five-level depth encoder–decoder architecture. Each level in the encoder consists of a context module. Each context module is a pre-activation residual block with two $3 \times 3 \times 3$ convolutional layers with a dropout layer ($p = 0.3$) in between. Context modules are connected by $3 \times 3 \times 3$ convolutions with a stride = 2 to reduce the dimensions of the feature maps. In the decoder the low-resolution feature maps are up-sampled by repeating the feature voxels twice in the spatial dimension followed by a $3 \times 3 \times 3$ convolution that halves the number of feature maps. To reduce stochasticity induced by small batch size, instance normalization layers are used instead of the more traditional batch normalization. Skip connections propagate same level features from encoder to decoder. In the decoder, segmentation layers at different levels are integrated via concatenation to reduce the coarseness of the final segmentation output. The activation function used throughout the network is the leaky ReLU with negative slope of 0.01 [44].

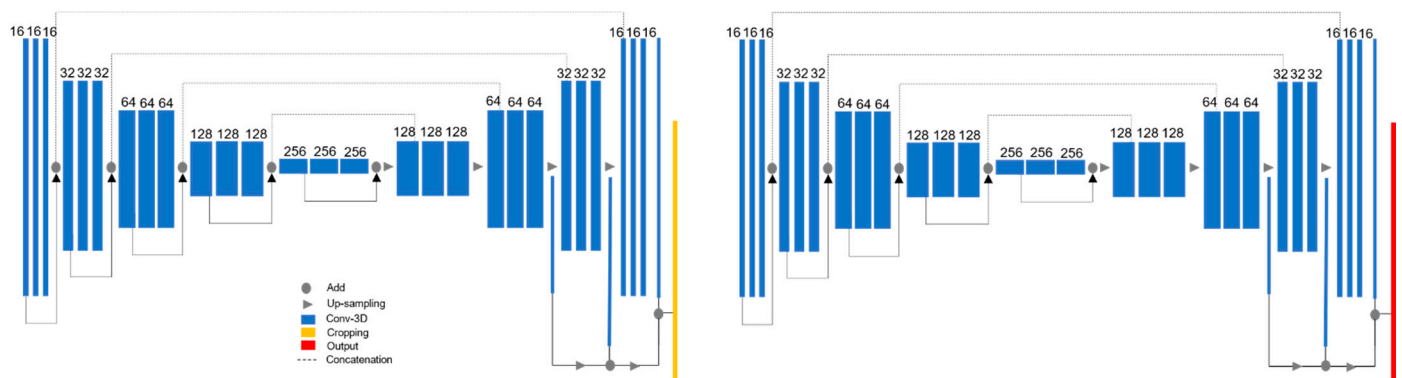


Figure 3. 3D Dual U-Net structure proposed by [14]. Blue blocks represent 3D features; orange refers to the cropping interface to crop the region of interest of the first U-Net prediction. Image adapted from [14].

4. Results

4.1. Experimental Configuration

The network was trained several times with fixed experimental settings. The input of the network consisted of volumetric images resized to (224, 144, and 96) with a batch size equal to 1. The number of base filters is 16, and at each level (0–4), the number of filters is computed by $(2^{\text{level}}) \times 16$.

Training was performed using the Adam optimizer with an initial learning rate equal to 5×10^{-4} , which reduces by half after 10 epochs if the validation loss is not improving. The convergence is defined as not improving after 50 epochs, and the maximum number of epochs is 500 with 200 steps per epoch. The training–validation split ratio equals 0.8. In addition, data augmentation techniques were used during training to reduce overfitting. Geometric transformations such as image rotations (10 degrees of rotation range), image shifts (in the range of 0.08) and zoom operations (with range equal to 0.2) were applied. Furthermore, to avoid the issue of class imbalance a modified version the Dice loss function introduced by [45] was used for training and validation. The binary class Dice coefficient loss function is defined as follows:

$$L_{DC} = -\frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i + \sum_i^N g_i + \sigma} \quad (1)$$

where i refers to each voxel of the predicted (p_i) and ground-truth (g_i) binary segmentation volumes of size N . The smoothness factor σ equals 0.00001 and was established to avoid the division by 0.

The proposed framework was developed using TensorFlow and Keras, open-source DL libraries for Python, and it was trained on a NVIDIA A100 Tensor with 40 GB RAM. The training time of the network trained with Database 1 (100 LGE-MRI volumes) was 24 h, whilst the network trained with Database 2 (19 bSSFP volumes) was comparatively faster, with the learning process taking 6 h. The evaluation of the predicted segmentations was performed via Dice coefficient and the average surface distance (ASD). The Dice coefficient measures the degree of overlap between the prediction and ground-truth segmentations. The ASD is defined as the average of the minimum distances (voxel-wise) between the ground-truth and prediction object boundaries. The ASD is evaluated using the shortest Euclidean distance of a voxel v to a point P , defined as:

$$\bar{d}(v, P) = \min_{p \in P} \|v - p\| \quad (2)$$

4.2. Source Domain Model

Firstly, the dual 3D U-Net was trained and tested with a large number of labelled images from LGE-MRI (source) domain. For this, Database 1 was employed which contains 100 volumes each of them composed of 88 slices. The aim of this experiment was to pre-

train the network and obtain a high-performance model from which knowledge can be inferred. The mean Dice coefficient and standard deviation obtained during evaluation across the 20 testing volumes were 0.9155 ± 0.0270 . The segmentation results for three representative patients and slices of Database 1 can be observed in Figure 4.



Figure 4. Axial slices of LA segmentations for three representative patients (a–c) obtained through the LGE-MRI source model. Ground-truth mask in white, prediction mask in red and intersection in pink.

4.3. Supervised Domain Adaptation Models

In a second stage, the learned knowledge θ_{LA}^{LGE} in D_{LGE} was transferred under the proposed SDA framework via weight initialization. In this line, the dual 3D U-Net was adapted for D_{bSSFP} by retraining the network with few labelled samples of bSSFP. For this purpose, Database 2 was employed which consists of 19 bSSFP volumetric samples. The SDA was performed twice, once for LA segmentation task and the other for RA segmentation task; thus, two domain-adapted networks were obtained.

Table 2 shows the Dice and ASD metrics of three LA segmentation models, each with a different training strategy and evaluated with bSSFP test samples from Database 2. The training strategies include the source model trained with LGE-MRI images and the models trained with bSSFP samples (Database 2) with and without SDA. This table provides a performance comparison of the domain-adapted network in the $D_S \neq D_T$ and $T_S = T_T$ scenario. We use boldface to highlight the most important results in the tables.

Table 3 shows the Dice and ASD metrics for RA segmentation models with and without SDA approach ($D_S \neq D_T$ and $T_S \neq T_T$ scenario).

Table 2. Dice coefficient and ASD for LA segmentation on bSSFP (Database 2) test data set.

Methods	Dice				ASD [mm]			
	Patient 1	Patient 2	Patient 3	Mean $\pm \sigma$	Patient 1	Patient 2	Patient 3	Mean $\pm \sigma$
LGE-MRI	0.7155	0.8030	0.7275	0.7487 ± 0.0474	2.3120	1.8907	4.5210	2.9079 ± 1.4128
w/o SDA	0.8693	0.8756	0.8736	0.8729 ± 0.0032	1.3614	1.2137	2.3517	1.6422 ± 0.6188
W-SDA	0.8781	0.8948	0.8707	0.8813 ± 0.0124	1.2522	0.9726	2.4132	1.5460 ± 0.7639

Table 3. Dice coefficient and ASD for RA segmentation on bSSFP (Database 2) test data set.

Methods	Dice				ASD [mm]			
	Patient 1	Patient 2	Patient 3	Mean $\pm \sigma$	Patient 1	Patient 2	Patient 3	Mean $\pm \sigma$
w/o SDA	0.9162	0.8773	0.8892	0.8942 ± 0.0199	0.9683	1.7313	2.1606	1.6201 ± 0.6039
W-SDA	0.9273	0.9043	0.9166	0.9160 ± 0.0115	0.7278	1.5357	1.6438	1.3024 ± 0.5006

4.4. External Database Validation

With the aim of providing an in-depth evaluation of the proposed SDA framework, an external validation was carried out. The idea behind this experiment is to corroborate the ability of generalization of the domain-adapted models. To this end, the trained models for

bSSFP domain with and without adaptation were further validated with a testing database of 19 volumetric bSSFP samples with 300–400 slices each (Database 3).

Tables 4 and 5 show the average inferred Dice coefficient and ASD metric for LA and RA segmentation tasks, respectively. Note that in the case of LA segmentation the source LGE-MRI trained model is included as baseline indicator (Table 3).

Table 4. Average Dice coefficient and ASD for LA prediction segmentation on bSSFP (Database 3) test data set.

Methods	Dice	ASD [mm]
LGE-MRI	0.7552 ± 0.0640	2.9208 ± 1.6903
w/o SDA	0.8186 ± 0.0703	2.6311 ± 1.4620
w-SDA	0.8405 ± 0.0614	2.2574 ± 1.5047

Table 5. Average Dice coefficient and ASD for RA prediction segmentation on bSSFP (Database 3) test data set.

Methods	Dice	ASD [mm]
w/o SDA	0.7724 ± 0.0996	3.2821 ± 1.8413
w-SDA	0.7993 ± 0.0928	2.8816 ± 1.6158

To further assess the evaluation performance of each model, qualitative results of the segmentation predictions were obtained.

Figure 5 shows the axial view of predictions obtained for three representative patients (and slices) of the testing cohort from Database 3 with the different trained models for LA segmentation.

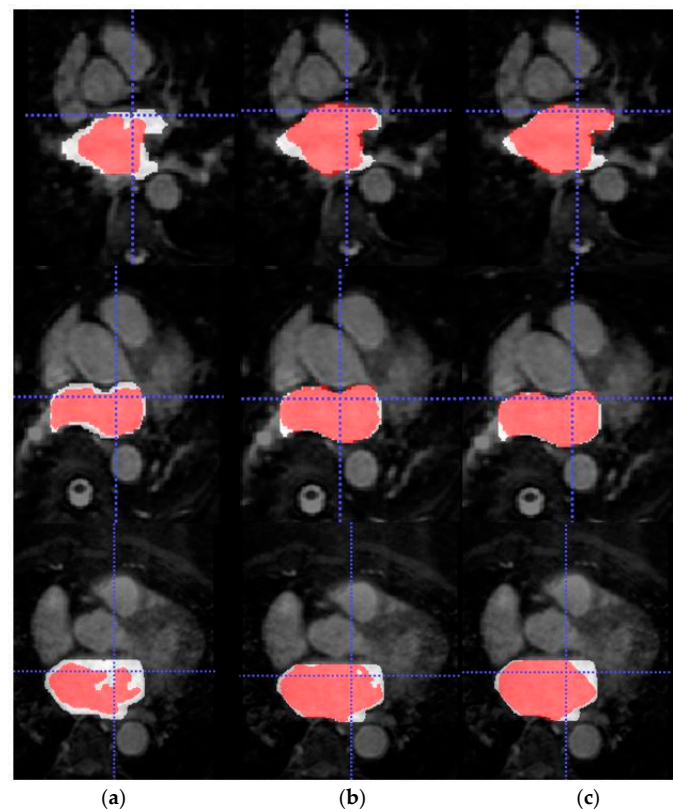


Figure 5. From top to bottom, axial slice of LA segmentations for three patients obtained through models (a) LGE-MRI, (b) w/o SDA and (c) w-SDA. Ground-truth mask in white, prediction mask in red and intersection in pink.

Figure 6 shows the axial view of predictions obtained for three representative testing patients (and slices) of Database 3 with the different trained models for RA segmentation.

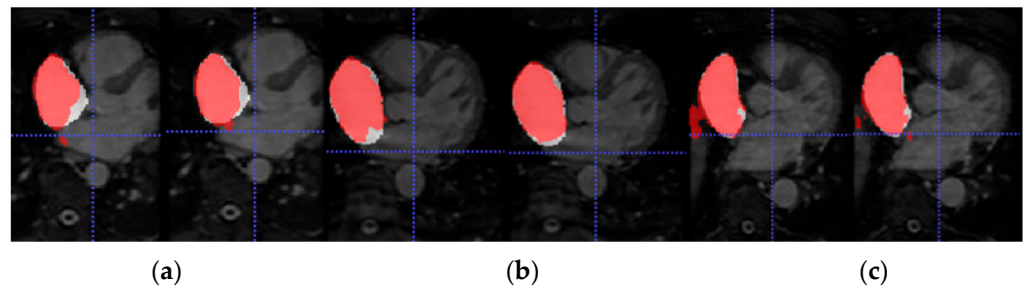


Figure 6. From left to right, axial slice of RA segmentations in networks w/o SDA (left) and w-SDA (right) for 3 patients (a–c).

Furthermore, 3D reconstructions of the atrial cavities were obtained from the testing patients of Database 3 to assess the geometry of the predicted segmentations.

Figure 7 illustrates the 3D reconstruction of the LA segmentation obtained by LGE-MRI, bSSFP and bSSFP with SDA-trained models for two representative patients of Database 3 as well as the ground-truth mask reconstruction for comparison.

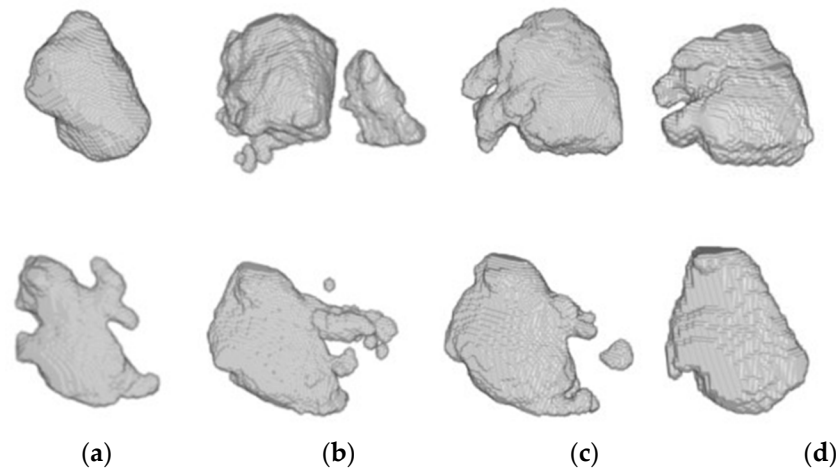


Figure 7. From top to bottom, 3D reconstruction of LA segmentations for two representative patients obtained through the models (a) LGE-MRI, (b) w/o SDA, (c) SDA and (d) ground-truth.

Figure 8 shows the 3D reconstruction of both the LA and RA cavities obtained by the trained models bSSFP with and without domain adaptation for two representative patients from Database 3.

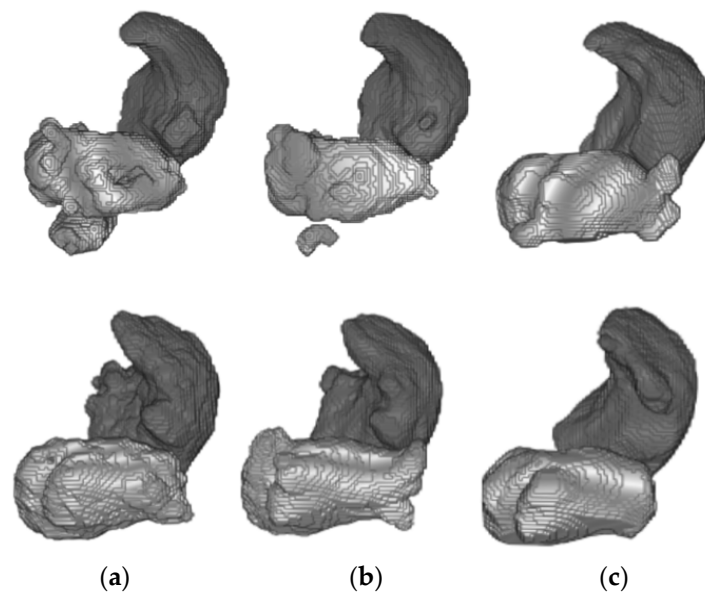


Figure 8. From top to bottom, 3D reconstruction of LA (light gray) and RA (dark gray) segmentations for two representative patients of Database 3 through the models (a) w/o SDA, (b) SDA and (c) ground-truth.

5. Discussion

In this section, we review the results obtained in the different experiments presented in Section 4. Firstly, the source network was trained with a large number of annotated LGE-MRI samples (Database 1). As can be observed, the model achieved very high accuracy segmentation results. This is expected as Database 1 corresponds to the data provided by the 2018 LA segmentation challenge [25], and the network proposed by [14] was specifically designed to face this challenge. Therefore, by keeping the same learning parameters, the Dice obtained coincides with the Dice reported in [14] (0.91–0.92). The segmentation masks shown in Figure 5 further highlight the high-accuracy performance of this model.

As we can observe in Tables 2 and 3, for both LA and RA segmentation the domain adaptation improved the segmentation performance in terms of Dice and ASD metrics. Remarkably, with the proposed SDA methodology the RA segmentation average Dice increased to 91%, and the ASD was reduced to 1.3024 mm. This is particularly interesting as the knowledge was transferred from a source model pre-trained for a different segmentation task (LA), which demonstrates the robustness of SDA approaches applied in a $T_S \neq T_T$ scenario. With this approach, models could be trained with a few labelled examples and yield high-accuracy segmentation results regardless of the target segmentation task.

Furthermore, as we can observe in Table 2, the source LGE-MRI model obtained a considerably lower Dice score (from 0.8813 to 0.7773) and higher ASD score (from 1.5460 to 2.9079 mm) when evaluated with Database 2. This illustrates how the performance of a U-Net trained from a particular source domain, when transferred to a different target domain, can drop unexpectedly, otherwise known as the domain shift problem. Compared to the Dice average of evaluated methods presented in the Multi-Modality Whole Heart Segmentation Challenge with the same database (0.815 for LA segmentation and 0.831 for RA segmentation) [41], the Dice obtained with our proposed method for both LA and RA segmentation is considerably higher; note that in the RA segmentation task, the Dice increment is almost 10%, which suggests that supervised DA and transfer learning is an effective method for obtaining good quality LA and RA segmentation. Although each testing volume contains between 100 and 200 image slices, to obtain more representative results, an increased number of testing patients are needed. For this reason, the proposed SDA framework is also assessed using an external database.

Tables 4 and 5 summarize the comparison results, in terms of Dice and ASD scores, of the trained models evaluated with an external bSSFP database for LA and RA segmentation, respectively. As we can observe, the SDA models achieved superior segmentation performance for both segmentation tasks, obtaining an average Dice and ASD score of 0.8405 and 2.2574 mm, respectively, for LA segmentation and an average Dice and ASD score of 0.7993 and 2.8816 mm, respectively, for RA segmentation. These results suggest that domain adaptation also improves the generalization capability of segmentation models. The results in Table 4 show that both bSSFP-trained models outperformed the source LGE-MRI trained model, achieving a 10% of gain in Dice accuracy from the LGE-MRI source model to the SDA-bSSFP model. This illustrates the importance of adapting the domain of the models to the target segmentation. Furthermore, it should be noted that the gain in Dice accuracy between bSSFP models with and without SDA is similar in both RA and LA target segmentation tasks (2–3%), further proving the robustness of SDA methods. Regarding the ASD metrics, an increased number of samples in the test set would be needed to properly assess the statistical significance of the results.

Due to the absence of previous works that employ Database 3 solely as an external validation dataset, we are not able to compare the generalization ability of our proposed model. On the other hand, a visual representation of the increased performance of the models with SDA can be seen in Figures 5 and 6 for LA and RA segmentation tasks, respectively. As we can observe, the blood pool prediction segmentation of the LA and RA performs best in the SDA model.

In addition, Figures 7 and 8 depict the 3D reconstruction of the prediction segmentation obtained by each trained model. Specifically, Figure 8 shows the increased performance of LA segmentation from the source model to the SDA bSSFP model as the geometry of the cavity becomes more similar in shape to the ground-truth reconstruction. Furthermore, the domain shift problem can be qualitatively assessed as the LGE-MRI source model clearly fails to generate an atrial geometry anatomically similar to the ground-truth. Figure 7 shows the combined reconstruction of LA and RA predictions in bSSFP models with and without SDA. As we can observe with SDA, the overall geometry of the atrial cavities is closer to the ground-truth.

In light of the results, transfer learning with a supervised DA approach enables that a model trained with a few labelled samples could yield improved segmentation accuracy by successfully leveraging knowledge from different image domains and segmentation tasks. Furthermore, SDA models demonstrate a greater generalization capacity when evaluated with external databases without retraining.

The limitations of this work remain similar to those presented in the previous work [15]. Specifically, more testing image data would be required to further validate the generalization capability of the obtained model and to adequately evaluate the statistical significance of the results. Given that the results promisingly show that the SDA-adapted models show good performance in a bSSFP image domain without the need for retraining, image datasets of more different domains would be needed to test if the improved generalization ability can be extended to other domains. Due to the limited amount of public cardiac datasets with LA and RA annotations, this constraint remains challenging to resolve; nevertheless, since the last presented work [15], a dataset with RA annotations performed manually by a radiologist has been generated as an attempt to reduce this main limitation.

6. Conclusions and Future Work

In conclusion, supervised DA approaches enable deep neural networks to be trained with a few labelled examples to improve segmentation accuracy results by leveraging the knowledge of a larger source dataset of a different domain. Furthermore, our proposed framework has demonstrated that leveraging knowledge from different segmentation tasks (LA to RA) could be useful for improving segmentation accuracy.

These results are the first step toward the integration of automatic segmentation algorithms in the clinical practice for AF management where accurate segmentation models of

both atrial cavities are needed to ensure the success of AF ablation procedures. Furthermore, the amount of manually labelled samples to train automatic computational models is limited in clinical practice, especially for the RA. Therefore, it is extremely useful to be able to leverage knowledge from models trained on LA segmentation tasks for RA segmentation. In addition, as acquisition protocols vary in different centers, there is a need to take advantage of the available clinical datasets even from a different domain for computation learning. Our proposed method has proven to have greater generalization capacity when trained with an external testing database, thus potentially eliminating the need for retraining for each acquisition protocol employed.

The trained models should be further evaluated with external databases of different domains and a large sample size to explore the role of SDA approach in increasing segmentation accuracy and the generalization ability of the network. Furthermore, it would be interesting to evaluate the obtained SDA models with RA segmentations manually generated by other trained experts to assess the effect of inter-user variability on the results obtained.

As a future line of research and given the promising results obtained by SDA techniques, we propose extending this framework to generate multi-organ segmentation models. Finally, the extent of the differences between source and image domains and source and image tasks for SDA to remain effective should be further investigated (i.e., the effect of introducing noise to the source domain, also known as adversarial attacks).

Author Contributions: Conceptualization, A.C. and V.N.; methodology, M.S.-V., A.C. and V.N.; software, M.S.-V.; validation, M.S.-V. and A.C.; formal analysis, M.S.-V., A.C. and V.N.; investigation, M.S.-V.; resources, M.S.-V., A.C. and V.N.; data curation, C.F. and L.M.-B.; writing—original draft preparation, M.S.-V.; writing—review and editing, A.C. and V.N.; supervision, A.C. and V.N.; project administration, V.N.; funding acquisition, V.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the GVA through project PROMETEO/2019/109. The equipment used for this research has been funded by the European Union within the operating Program ERDF of the Valencian Community 2014–2020 with the grant number IDIFEDER/2020/030.

Data Availability Statement: The manual annotations of the RA of dataset [40] generated in this study can be found in <https://github.com/marsaivi/SDA-RA-Dataset> (accessed on 14 May 2021).

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Lippi, G.; Sanchis-Gomar, F.; Cervellin, G. Global epidemiology of atrial fibrillation: An increasing epidemic and public health challenge. *Int. J. Stroke* **2021**, *16*, 217–221. [[CrossRef](#)] [[PubMed](#)]
2. Li, L.; Wu, F.; Yang, G.; Xu, L.; Wong, T.; Mohiaddin, R.; Firmin, D.; Keegan, J.; Zhuang, X. Atrial scar quantification via multi-scale CNN in the graph-cuts framework. *Med. Image Anal.* **2020**, *60*, 101595. [[CrossRef](#)] [[PubMed](#)]
3. Zipes, D.P.; Jalife, J. (Eds.) *Cardiac Electrophysiology: From Cell to Bedside: Seventh Edition*, 4th ed.; W.B. Saunders: Philadelphia, PA, USA, 2004; p. 1155. [[CrossRef](#)]
4. Rolf, S.; Hindricks, G.; Sommer, P.; Richter, S.; Arya, A.; Bollman, A.; Kosiuk, J.; Koutalas, E. Electroanatomical mapping of atrial fibrillation: Review of the current techniques and advances. *J. Atr. Fibrillation* **2014**, *7*, 57–68. [[CrossRef](#)]
5. Guglielmo, M.; Baggiano, A.; Muscogiuri, G.; Fusini, L.; Andreini, D.; Mushtaq, S.; Conte, E.; Annoni, A.; Formenti, A.; Mancini, E.M.; et al. Multimodality imaging of left atrium in patients with atrial fibrillation. *J. Cardiovasc. Comput. Tomogr.* **2019**, *13*, 340–346. [[CrossRef](#)]
6. McGann, C.; Akoum, N.; Patel, A.; Kholmovski, E.; Revelo, P.; Damal, K.; Wilson, B.; Cates, J.; Harrison, A.; Ranjan, R.; et al. Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI. *Circ. Arrhythmia Electrophysiol.* **2014**, *7*, 23–30. [[CrossRef](#)] [[PubMed](#)]
7. Hansen, B.J.; Zhao, J.; Csepe, T.A.; Moore, B.T.; Li, N.; Jayne, L.A.; Kalyanasundaram, A.; Lim, P.; Bratasz, A.; Powell, K.A.; et al. Atrial fibrillation driven by micro-anatomic intramural re-entry revealed by simultaneous sub-epicardial and sub-endocardial optical mapping in explanted human hearts. *Eur. Heart J.* **2015**, *36*, 2390–2401. [[CrossRef](#)] [[PubMed](#)]
8. Zhao, J.; Hansen, B.J.; Wang, Y.; Csepe, T.A.; Sul, L.V.; Tang, A.; Yuan, Y.; Li, N.; Bratasz, A.; Powell, K.A.; et al. Three-dimensional integrated functional, structural, and computational mapping to define the structural “fingerprints” of heart-specific atrial fibrillation drivers in human heart ex vivo. *J. Am. Heart Assoc.* **2017**, *6*. [[CrossRef](#)]

9. Kurzendorfer, T.; Forman, C.; Schmidt, M.; Tillmanns, C.; Maier, A.; Brost, A. Fully automatic segmentation of left ventricular anatomy in 3-D LGE-MRI. *Comput. Med. Imaging Graph.* **2017**, *59*, 13–27. [[CrossRef](#)] [[PubMed](#)]
10. Vesal, S.; Ravikumar, N.; Maier, A. *Automated Multi-Sequence Cardiac MRI Segmentation Using Supervised Domain Adaptation*; Springer: Cham, Switzerland, 2019; Volume 1. [[CrossRef](#)]
11. Li, C.; Tong, Q.; Liao, X.; Si, W.; Sun, Y.; Wang, Q.; Heng, P.A. Attention Based Hierarchical Aggregation Network for 3D Left Atrial Segmentation. In *International Workshop on Statistical Atlases and Computational Models of the Heart*; Springer: Cham, Switzerland, 2019; pp. 255–264. [[CrossRef](#)]
12. Vesal, S.; Ravikumar, N.; Maier, A. Dilated Convolutions in Neural Networks for Left Atrial Segmentation in 3D Gadolinium Enhanced-MRI. In *International Workshop on Statistical Atlases and Computational Models of the Heart*; Springer: Cham, Switzerland, 2019; pp. 319–328. [[CrossRef](#)]
13. Vesal, S.; Maier, A.; Ravikumar, N. Fully Automated 3D Cardiac MRI Localisation and Segmentation Using Deep Neural Networks. *J. Imaging* **2020**, *6*, 65. [[CrossRef](#)]
14. Jia, S.; Despinasse, A.; Wang, Z.; Delingette, H.; Pennec, X.; Jaïs, P.; Cochet, H.; Sermesant, M. Automatically Segmenting the Left Atrium from Cardiac Images Using Successive 3D U-Nets and a Contour Loss. *Lect. Notes Comput. Sci.* **2019**. [[CrossRef](#)]
15. Saiz-Vivó, M.; Colomer, A.; Naranjo, V. Deep convolutional encoder-decoder network for semantic segmentation of atrial cavity. In Proceedings of the XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020), Valladolid, Spain, 25–27 November 2020; pp. 443–446.
16. Huang, Y.L.; Chen, D.R. Watershed segmentation for breast tumor in 2-D sonography. *Ultrasound Med. Biol.* **2004**, *30*, 625–632. [[CrossRef](#)]
17. Masoumi, H.; Behrad, A.; Pourmina, M.A.; Roosta, A. Automatic liver segmentation in MRI images using an iterative watershed algorithm and artificial neural network. *Biomed. Signal Process. Control* **2012**, *7*, 429–437. [[CrossRef](#)]
18. Zhao, Y.; Rada, L.; Chen, K.; Harding, S.P.; Zheng, Y. Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images. *IEEE Trans. Med. Imaging* **2015**, *34*, 1797–1807. [[CrossRef](#)]
19. Ciecholewski, M.; Spodnik, J.H. Semi-automatic corpus callosum segmentation and 3d visualization using active contour methods. *Symmetry* **2018**, *10*, 589. [[CrossRef](#)]
20. Tian, Z.; Liu, L.; Zhang, Z.; Fei, B. Superpixel-based segmentation for 3D prostate MR images. *IEEE Trans. Med. Imaging* **2015**, *35*, 791–801. [[CrossRef](#)]
21. Nguyen, D.C.T.; Benameur, S.; Mignotte, M.; Lavoie, F. Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images. *Med. Image Anal.* **2018**, *48*, 58–74. [[CrossRef](#)]
22. Jamart, K.; Xiong, Z.; Maso Talou, G.D.; Stiles, M.K.; Zhao, J. Mini Review: Deep Learning for Atrial Segmentation From Late Gadolinium-Enhanced MRIs. *Front. Cardiovasc. Med.* **2020**. [[CrossRef](#)] [[PubMed](#)]
23. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. *Lect. Notes Comput. Sci.* **2015**, *9351*, 234–241. [[CrossRef](#)]
24. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-net: Learning dense volumetric segmentation from sparse annotation. *Lect. Notes Comput. Sci.* **2016**, 424–432. [[CrossRef](#)]
25. 2018 Atrial Segmentation Challenge. Available online: <http://atriaseg2018.cardiacatlas.org/> (accessed on 10 March 2020).
26. Xia, Q.; Yao, Y.; Hu, Z.; Hao, A. Automatic 3D atrial segmentation from GE-MRIs using volumetric fully convolutional networks. In *International Workshop on Statistical Atlases and Computational Models of the Heart*; Springer: Cham, Switzerland, 2018; pp. 211–220.
27. Yang, X.; Wang, N.; Wang, Y.; Wang, X.; Nezafat, R.; Ni, D.; Heng, P.A. Combating uncertainty with novel losses for automatic left atrium segmentation. In *International Workshop on Statistical Atlases and Computational Models of the Heart*; Springer: Cham, Switzerland, 2018; pp. 246–254.
28. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
29. Cheplygina, V.; de Bruijne, M.; Pluim, J.P.W. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. *Med. Image Anal.* **2019**, *54*, 280–296. [[CrossRef](#)] [[PubMed](#)]
30. Patel, V.M.; Gopalan, R.; Li, R.; Chellappa, R. Visual Domain Adaptation: A survey of recent advances. *IEEE Signal Process Mag.* **2015**, *32*, 53–69. [[CrossRef](#)]
31. Motiian, S.; Piccirilli, M.; Adjero, D.A.; Doretto, G. Unified Deep Supervised Domain Adaptation and Generalization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017. [[CrossRef](#)]
32. Conjeti, S.; Katouzian, A.; Roy, A.G.; Peter, L.; Sheet, D.; Carlier, S.; Laine, A.; Navab, N. Supervised domain adaptation of decision forests: Transfer of models trained in vitro for in vivo intravascular ultrasound tissue characterization. *Med. Image Anal.* **2016**, *32*, 1–17. [[CrossRef](#)]
33. Hofer, C.; Kwitt, R.; Holler, Y.; Trinka, E.; Uhl, A. Simple domain adaptation for cross-dataset analyses of brain MRI data. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 441–445. [[CrossRef](#)]
34. Wachinger, C.; Reuter, M. Domain adaptation for Alzheimer’s disease diagnostics. *Neuroimage* **2016**, *139*, 470–479. [[CrossRef](#)] [[PubMed](#)]

35. Menegola, A.; Fornaciali, M.; Pires, R.; Bittencourt, F.V.; Avila, S.; Valle, E. Knowledge transfer for melanoma screening with deep learning. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 297–300. [[CrossRef](#)]
36. Ribeiro, E.; Häfner, M.; Wimmer, G.; Tamaki, T.; Tischendorf, J.J.W.; Yoshida, S.; Tanaka, S.; Uhl, A. Exploring texture Transfer Learning for Colonic Polyp Classification via Convolutional Neural Networks. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, VIC, Australia, 18–21 April 2017; pp. 1044–1048. [[CrossRef](#)]
37. Kamnitsas, K.; Baumgartner, C.; Ledig, C.; Newcombe, V.; Simpson, J.; Kane, A.; Menon, D.; Nori, A.; Criminisi, A.; Rueckert, D.; et al. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In *International Conference on Information Processing in Medical Imaging*; Springer: Cham, Switzerland, 2017. [[CrossRef](#)]
38. Kushibar, K.; Valverde, S.; González-Villà, S.; Bernal, J.; Cabezas, M.; Oliver, A.; Lladó, X. Supervised Domain Adaptation for Automatic Sub-cortical Brain Structure Segmentation with Minimal User Interaction. *Sci. Rep.* **2019**, *9*, 1–15. [[CrossRef](#)] [[PubMed](#)]
39. Ghafoorian, M.; Mehrtash, A.; Kapur, T.; Karssemeijer, N.; Marchiori, E.; Pesteie, M.; Guttmann, C.R.G.; de Leeuw, F.E.; Tempny, C.M.; van Ginneken, B.; et al. Transfer learning for domain adaptation in MRI: Application in brain lesion segmentation. *Lect. Notes Comput. Sci.* **2017**. [[CrossRef](#)]
40. MM-WHS: Multi-Modality Whole Heart Segmentation. Available online: <http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/mmwhs/> (accessed on 10 March 2020).
41. Zhuang, X.; Li, L.; Payer, C.; Štern, D.; Urschler, M.; Heinrich, M.P.; Oster, J.; Wang, C.; Smedby, Ö.; Bian, C.; et al. Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge. *Med. Image Anal.* **2019**, *58*, 101537. [[CrossRef](#)] [[PubMed](#)]
42. Cardiac Atlas Project. Left Atrium Segmentation Challenge. Available online: <https://www.cardiacatlas.org/challenges/left-atrium-segmentation-challenge/> (accessed on 8 March 2020).
43. Tobon-Gomez, C.; Geers, A.J.; Peters, J.; Weese, J.; Pinto, K.; Karim, R.; Ammar, M.; Daoudi, A.; Margeta, J.; Sandoval, Z.; et al. Benchmark for Algorithms Segmenting the Left Atrium from 3D CT and MRI Datasets. *IEEE Trans. Med. Imaging* **2015**, *34*, 1460–1473. [[CrossRef](#)]
44. Isensee, F.; Kickingereder, P.; Wick, W.; Bendszus, M.; Maier-Hein, K.H. Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge. *Lect. Notes Comput. Sci.* **2018**. [[CrossRef](#)]
45. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571. [[CrossRef](#)]