

---

# 6 Left Atrial Scarring Segmentation from Delayed-Enhancement Cardiac MRI Images: A Deep Learning Approach

*Guang Yang<sup>1</sup>, Xiahai Zhuang<sup>2</sup>, Habib Khan<sup>3</sup>,  
Eva Nyktar<sup>3</sup>, Shouvik Haldar<sup>3</sup>, Lei Li<sup>4</sup>, Rick Wage<sup>3</sup>,  
Xujiong Ye<sup>5</sup>, Greg Slabaugh<sup>6</sup>, Raad Mohiaddin<sup>1</sup>,  
Tom Wong<sup>1</sup>, Jennifer Keegan<sup>1</sup>, David Firmin<sup>1</sup>*

<sup>1</sup>Royal Brompton Hospital and Imperial College London

<sup>2</sup>Fudan University

<sup>3</sup>Royal Brompton Hospital

<sup>4</sup>Shanghai Jiao Tong University

<sup>5</sup>University of Lincoln

<sup>6</sup>City University London

## CONTENTS

6.1	Introduction .....	110
6.1.1	Background.....	110
6.1.2	Related Work .....	111
6.1.3	Our Contributions .....	112
6.2	Method.....	112
6.2.1	Study Population.....	112
6.2.2	MRI Protocol .....	113
6.2.3	Multi-Atlas Whole Heart Segmentation (MA-WHS).....	113
6.2.4	Image Over-Segmentation Using SLIC Super-Pixels .....	115
6.2.5	Atrial Scarring Segmentation Based on Super-Pixels Classification .....	116
6.2.5.1	Training Data Construction and Ground Truth Definition .....	116
6.2.5.2	Deep Learning Using Stacked Sparse Auto-Encoders .....	117
6.2.5.3	Hyper-Parameters Settings .....	119
6.2.6	Validation Approaches.....	120
6.2.6.1	Validation for the Whole Heart Segmentation.....	120
6.2.6.2	Validation for the Atrial Scarring Validation .....	120
6.3	Results.....	121
6.3.1	Whole Heart Segmentation.....	121
6.3.2	Atrial Scarring Segmentation.....	121

6.4 Discussion.....	122
6.5 Conclusion.....	124
Acknowledgments.....	125
References.....	125

## 6.1 INTRODUCTION

### 6.1.1 BACKGROUND

Atrial fibrillation (AF) is the most common arrhythmia of clinical significance, and it affects approximately 1–2% of the population, a figure that is rising fast with ageing [1], [2]. AF occurs when chaotic and disorganized electrical activity develops in the atria, causing muscle cells to contract irregularly and rapidly. Moreover, it is associated with structural remodelling, including fibrotic changes in the left atrial substrate [3]. AF can cause increased morbidity, especially stroke and heart failure, and result in poor mental health, dementia, and increased mortality [2], [4], [5].

Pharmacological treatment of AF aims to restore and maintain sinus rhythm [6]. However, AF recurrence, side effects of antiarrhythmic drugs, and risks of proarrhythmia may offset the benefits of pharmacological treatment [7]. Consequently, there have been increasing efforts to develop non-pharmacological methods to treat AF patients such as percutaneous catheter ablation (CA) and surgical ablation (SA). Since the importance of pulmonary vein (PV) triggers in the initiation of AF was found [8], CA, which electrically isolates the PVs, has developed into an important interventional therapy [9]. However, despite efforts to improve targeting and delivery of CA, the success rate for a single procedure is just 30–50% at 5 years follow-up [10], [11]. Thoracoscopic SA has shown higher long-term success rates for a single procedure [12]–[14], but this comes with a procedural major adverse event rate of 23% [15].

The high failure rate of ablation in AF patients can be attributed to: (1) inadequate understanding of the arrhythmia mechanisms and arrhythmogenic substrates; (2) difficulty in identifying the potential non-responders of ablation; (3) inability to establish the ideal ablation strategy for each patient; (4) inadequate information of lesion integrity and longevity; and (5) limitations in the information about the location and extent of the ablation-induced scarring during and/or after the procedure [16], [17]. These have motivated researchers to develop better fibrosis imaging and assessment techniques to provide accurate guidance of the pre- and post-ablation procedures and improve their performance.

The current clinical gold standard for assessment of atrial scarring is electro-anatomical mapping (EAM), performed during an electrophysiological (EP) study [17]. The electrical activity of the left atrium (LA) is recorded using a mapping catheter prior to CA, with regions of scarring being associated with low voltage ( $<0.5\text{mV}$ ). The main limitations of this technique are its invasiveness, the use of ionizing radiation, and the suboptimal accuracy, with reported errors of up to 10 mm in the localization of scar tissue [18], [19].

Late gadolinium enhancement (LGE) magnetic resonance imaging (MRI) is an established noninvasive technique for detecting myocardial scar tissue [20]. With this technique, healthy and scar tissues are differentiated by their altered wash-in and wash-out contrast agent kinetics, which result in scar tissue being seen as a region of enhanced or high signal intensity while healthy tissue is nulled. While 2D breath-hold LGE MRI is well-established for ventricular imaging, there is a growing interest in imaging the thinner walled atria for identification of native and ablation scarring in AF patients [21]–[24]. This requires higher spatial resolution and contiguous coverage, and data are best acquired as a 3D volume during free-breathing with diaphragmatic respiratory-gating. Atrial 3D LGE imaging has been used to: (1) assess patient suitability for AF ablation by identifying potential non-responders [22], [25]–[30], and (2) define the most appropriate ablation approach [26], [27], [31]. In addition, visualization and quantification of native and post-ablation atrial scarring derived from LGE MRI has been used to guide initial and follow-up ablation procedures [27],

[28], [32]–[35]. Histopathological studies have validated LGE MRI for quantification of native AF fibrosis [30], [36] and for characterization of AF ablation-induced wall injury [37].

Visualization and quantification of atrial scarring requires objective, robust, and accurate segmentation of the enhanced scar regions from the LGE MRI images. Essentially, there are two segmentations required: one showing the cardiac anatomy (geometry), particularly the LA wall and PVs, the other delineating the enhanced scar regions. The former segmentation is required to rule out confounding enhanced tissues from other parts of the heart (e.g., the mitral valve and aorta), or the enhancement from non-heart structures, while the latter is a prerequisite for insightful visualization and meaningful quantification. Segmentation of the atrial scarring from LGE MRI images is a very challenging problem. Firstly, the LA wall is very thin, and scarring is hard to distinguish even by experienced expert cardiologists specialized in cardiac MRI. Secondly, residual respiratory motion, heart rate variability, low signal-to-noise ratio (SNR), and contrast agent wash-out during the long acquisition (current scanning time  $\approx 10$ mins) frequently result in image quality being poor. In addition, artifactual enhanced signal from surrounding tissues may result in a large number of false positives.

### 6.1.2 RELATED WORK

Oakes et al. [22] quantified the enhanced atrial scarring by analyzing the intensity histogram of the manually segmented LA wall. Perry et al. [38] applied k-means clustering to quantitatively assess normal and scarred tissue from manual LA wall segmentation. A grand challenge was carried out for evaluation and benchmarking of various atrial scarring segmentation methods, including histogram analysis, simple and advanced thresholding, k-means clustering, and graph-cuts [39]. Although these pioneering studies have shown promising results on the segmentation and quantification of atrial scarring using LGE MRI images, most have relied on manual segmentation of the LA wall and PVs. This has several drawbacks: (1) it is a time-consuming task; (2) there are intra- and inter-observer variations; (3) it is less reproducible for a multi-center and multi-scanner study. Moreover, some previous studies have assumed a fixed thickness of the LA wall, while there is no evidence that this is the case, and re-orientation and interpolation of the MR images can result in partial volume effects and increase the variance of the wall thickness. Inaccurate manual segmentation of the LA wall and PVs can further complicate the delineation of the atrial scarring, and its quantification can be error-prone. This could be one of the major reasons that there are ongoing concerns regarding the correlation between atrial scarring identified by LGE MRI (enhanced regions) and the gold standard EAM (low voltage regions) [30], [40].

The LA and PVs would ideally be segmented from the cardiac and respiratory-gated LGE MRI dataset. However, this is difficult as normal tissue is nulled and only scar tissue is seen with high signal. Other options are to segment them from a separately acquired breath-hold magnetic resonance angiogram (MRA) study [29], [41], [42] or from a respiratory and cardiac gated 3D Roadmap acquisition, that is, using a balanced steady state free precession (b-SSFP) sequence [43]. While MRA shows the LA and PVs with high contrast, these acquisitions are generally un-gated and acquired in an inspiratory breath-hold. The anatomy extracted from MRA therefore can be highly deformed compared to that in the LGE MRI study. Although the 3D Roadmap acquisition takes longer to acquire, it is in the same respiratory phase as the LGE MRI, and the extracted anatomy can be better matched. Ravanelli et al. [29] proposed to manually segment the LA wall and PVs using MRA images in 3D, for which both efficiency and accuracy have been claimed. The segmented LA and PVs were then mapped to LGE MRI and this was followed by a thresholding-based segmentation of the atrial scarring [29]. Recently, Tao et al. [42] combined atlas-based segmentation of LGE MRI and MRA to define the cardiac anatomy. After image fusion of the LGE MRI and MRA, accurate LA chamber and PVs segmentation has been achieved by a level set based local refinement, based on which an objective atrial scarring assessment is envisaged in future development [42]. Instead of using MRA, Karim et al. [43] used b-SSFP whole-heart acquisition to define the cardiac anatomy.

**TABLE 6.1****Previous Studies on Atrial Scarring.**

Methods	Subjects No.	LA Wall Segmentation	Scarring Segmentation	Evaluation Methods for Scarring Segmentation
Oakes et al., 2009 [22]	81	Manual	2-4 SD	Scar Percentage
Knowles et al., 2010 [41]	7	Semi-Auto	Maximum Intensity Projection	Scar Percentage
Perry et al., 2012 [44]	34	Manual	k-means Clustering	Dice Score
Ravanelli et al., 2014 [29]	10	Manual	4 SD	Dice Score
Karim et al., 2014 [43]	15	Semi-Auto	Graph-Cuts	Dice Score, ROC, Scar Percentage
Tao et al., 2016 [42]	46	Automatic Atlas	Maximum Intensity Projection	Qualitative Visualization
Veni et al., 2017 [45]	72	Automatic ShapeCut	Threshold/k-Means Clustering	Scar Percentage
Ours	20	Automatic Atlas	Super-Pixel and Deep Learning	Multiple Quantitative Metrics

The cardiac anatomy was resolved using a statistical shape model, and the atrial scarring was then segmented using a graph-cut model assuming that the LA wall is  $\pm 3$ mm from the endocardial border obtained from the LA geometry extraction [43]. Table 6.1 provides a summary of previously published methods on atrial scarring segmentation using LGE MRI.

### 6.1.3 OUR CONTRIBUTIONS

In this chapter, we present a novel fully automatic segmentation and objective assessment of atrial scarring for longstanding persistent AF patients scanned by LGE MRI. The LA chamber and PVs are defined using a multi-atlas based whole heart segmentation (MA-WHS) method on Roadmap MRI images, which are acquired using a respiratory and cardiac gated 3D b-SSFP sequence. LA and PVs geometry is resolved by mapping the segmented Roadmap anatomy to LGE MRI using the DICOM header data, and is further refined by affine and nonrigid registration steps. The LGE MRI images are over-segmented by a novel Simple Linear Iterative Clustering (SLIC) based super-pixels method [46]. Then a fully automatic supervised deep learning classification method is applied to segment the atrial scarring within the segmented LA and PVs geometry. In this study, two validation steps have been performed: one for the LA chamber and PVs segmentation; and one for the atrial scarring segmentation—both against established ground truths from manual segmentations by experienced expert-cardiologists specialized in cardiac MRI. This chapter is based on our previous work on atrial scarring segmentation [47]–[49].

The rest of this manuscript is organized as follows. Section 2 details the materials and main methods of this study. Section 3 demonstrates our experimental results, which are followed by discussions and a conclusion (Sections 4 and 5).

## 6.2 METHOD

### 6.2.1 STUDY POPULATION

Cardiac MRI was performed in longstanding persistent AF patients between 2011–2013 in agreement with the local regional ethics committee. A Likert-type scale was applied to score the image quality of each LGE MRI scan, e.g., 0 (non-diagnostic), 1 (poor), 2 (fair), 3 (good), and 4 (very good) depending on the level of SNR, appropriate inversion time, and the existence of navigator beam and ghost artefacts.

Ten pre-ablation scans with image quality  $\geq 2$  have been retrospectively entered into this study (~60% of all the scanned pre-ablation cases). To make a balanced dataset, we randomly selected 10 post-ablation cases from all the 26 post-ablation scans with image quality  $\geq 2$  (~92% of all the scanned post-ablation cases).

### 6.2.2 MRI PROTOCOL

Cardiac MR data were acquired on a Siemens Magnetom Avanto 1.5T scanner (Siemens Medical Systems, Erlangen, Germany).

Transverse navigator-gated 3D LGE MRI [21], [22], [50] was performed using an inversion prepared segmented gradient echo sequence (TE/TR 2.2ms/5.2ms) 15 minutes after gadolinium (Gd) administration (Gadovist—gadobutrol, 0.1mmol/kg body weight, Bayer-Schering, Berlin, Germany) when a transient steady-state of Gd wash-in and wash-out of normal myocardium had been reached [51]. Detailed scanning parameters are: 30–34 slices at  $1.5 \times 1.5 \times 4 \text{ mm}^3$ , reconstructed to 60–68 slices at  $0.75 \times 0.75 \times 2 \text{ mm}^3$ , field-of-view  $380 \times 380 \text{ mm}^2$ , acceleration factor of 2 using generalized auto-calibrating partially parallel acquisition (GRAPPA), acquisition window 125ms positioned within the subject-specific rest period, single R-wave gating, chemical shift fat suppression, flip angle  $20^\circ$ . Data were acquired during free-breathing using a crossed-pairs navigator positioned over the dome of the right hemi-diaphragm with navigator acceptance window size of 5mm and CLAWS respiratory motion control [52]. The nominal acquisition duration was 204–232 cardiac cycles assuming 100% respiratory efficiency.

Coronal navigator-gated 3D b-SSFP (TE/TR 1ms/2.3ms) Roadmap data were acquired with the following parameters: 80 slices at  $1.6 \times 1.6 \times 3.2 \text{ mm}^3$ , reconstructed to 160 slices at  $0.8 \times 0.8 \times 1.6 \text{ mm}^3$ , field-of-view  $380 \times 380 \text{ mm}^2$ , acceleration factor of 2 using GRAPPA, partial Fourier 6/8, acquisition window 125ms positioned within the subject-specific rest period, single R-wave gating, chemical shift fat suppression, flip angle  $70^\circ$ . Data were acquired during free-breathing using a crossed-pairs navigator positioned over the dome of the right hemi-diaphragm with navigator acceptance window size of 5mm and CLAWS respiratory motion control [52]. The nominal acquisition duration was 241 cardiac cycles assuming 100% respiratory efficiency. More details of the acquisition parameters can be found in Table 6.2.

### 6.2.3 MULTI-ATLAS WHOLE HEART SEGMENTATION (MA-WHS)

A multi-atlas approach [53], [54] was developed to derive the whole heart segmentation of the Roadmap acquisition and then mapped to LGE MRI [55], [56]. This segmentation consists of two major steps: (1) atlas propagation based on image registration algorithms and (2) label fusion from multi-atlas propagated segmentation results.

First we obtained 30 MRI Roadmap studies from the Left Atrium Segmentation Grand Challenge organized by King's College London [57] together with manual segmentations of the left atrium,

**TABLE 6.2**

**MRI Sequence Details for the 3D Roadmap and 3D LGE MRI Acquisitions.**

Sequence	ST (mm)	FOV (cm)	NEX	AM	RM	TR/TE/FA
3D Roadmap	1.6	$38 \times 38$	1	$256 \times 256$	$512 \times 512$	2.3ms/1ms/ $70^\circ$
3D LGE MRI	2	$38 \times 38$	1	$256 \times 256$	$512 \times 512$	5.2ms/2.2ms/ $20^\circ$

ST = reconstructed slice thickness; FOV = field of view; NEX = number of excitations; AM = acquisition matrix; RM = reconstruction matrix; TR/TE/FA = repetition time/echo time/flip angle.

pulmonary veins, and appendages. In these, we further labelled the right and left ventricles, the right atrium, the aorta, and the pulmonary artery, to generate 30 whole heart atlases. These 30 MRI Roadmap studies were employed only for building an independent multi-atlas dataset, which will then be used for segmenting our Roadmap studies that linked with the LGE MRI scans for the AF patients.

Let  $I$  be the target image to be segmented,  $\{(\mathcal{I}_a, \mathcal{L}_a) \mid a = 1, \dots, N\}$  be the set of atlases, where  $N = 30$ ,  $\mathcal{I}_a$  and  $\mathcal{L}_a$  are respectively the intensity image and corresponding segmentation label image of the  $a$ -th atlas. For each atlas, MA-WHS performs an atlas-to-target registration, by maximizing the similarity between the images, to derive the set of warped atlases,

$$T_a = \arg \max_{T_a} \text{ImageSimilarity}(I, \mathcal{I}_a), \text{ and } \begin{cases} A_a = T_a(\mathcal{I}_a) \\ L_a = T_a(\mathcal{L}_a) \end{cases}, \quad (6.1)$$

in which  $T_a$  is the resulting transformation of the registration, and  $\{(A_a, L_a) \mid a = 1, \dots, N\}$  are respectively the warped atlas intensity image and corresponding segmentation result. Here, we employ the hierarchical registration for segmentation propagation, which was specifically designed for the whole heart MRI images and consists of three steps, namely the global affine registration for localization of the whole heart, the local affine registration for the initialization of the substructures, and the fully deformable registration for local detail refinement [58]. Image similarity metrics evaluate how similar the atlas and target image are. In this work we propose to use the spatially encoded mutual information (SEMI) method, which has been shown to be robust against intensity non-uniformity and different intensity contrast [59], that is

$$\text{ImageSimilarity}(I, \mathcal{I}_a) = \{S_1, \dots, S_{n_s}\} \quad (6.2)$$

where  $\{S_1, \dots, S_{n_s}\}$  are the SEMI and computed based on the spatially encoded joint histogram,

$$H_s(I, \mathcal{I}_a) = \sum_{x \in \Omega} w_1(I(x)) w_2(\mathcal{I}_a(x)) W_s(x). \quad (6.3)$$

Here,  $w_1(I(x))$  and  $w_2(\mathcal{I}_a(x))$  are Parzen window estimation and  $W_s(x)$  is a weighting function to encode the spatial information [59].

After the multi-atlas propagation, a label fusion algorithm is required to generate one final segmentation of the LA from the 30 propagated results,

$$L_I = \text{LabelFusion}(\{(A_1, L_1) \dots (A_N, L_N)\}). \quad (6.4)$$

The label fusion decides how to combine the multiple classification results into one labelling result. Since the atlases can produce segmentations with dramatically different accuracy at different locations, it should evaluate the performance of each atlas locally and assign different weights for the atlases at each pixel of the target image in decision fusion.

The recent literatures have many new methods [60]–[67] on improving multi-atlas segmentation using sophisticatedly designed algorithms, which generally need to evaluate local similarity between patches from the atlases and the target image for local weighted label fusion,

$$L_I(x) = \arg \max_{l \in \{l_{bk}, l_{la}\}} \sum_a w_a(S(I, A_a, x)) \delta(L_a(x), l), \quad (6.5)$$

in which  $l_{bk}$  and  $l_{la}$  indicate the labels of the background and left atrium, respectively, and the local weight  $w_a(\cdot) \propto S(\cdot)$  is determined by the local similarity  $S(\cdot)$  between the target image and the atlas.  $\delta(a, b)$  is the Kronecker delta function which returns 1 when  $a = b$  and returns 0 otherwise.

For the LA segmentation, we propose to use the multi-scale patch based label fusion (MSP-LF). This is because the intensity distribution of the blood pool in the LA is almost identical to that of the blood pool in the other chambers and great vessels. The multi-scale space theory can handle different level information within a small patch and has been applied to feature extraction/detection and image matching [54], [67]–[73]. The patches we compute from different scale spaces can represent the different levels of structural information, with low scale capturing local fine structure and high scale suppressing fine structure but providing global structural information of the image. This is different from the conventional patch-based methods, which only compute the local structural information within the patch. To avoid increasing the computational complexity, we adopt the multi-resolution implementation and couple it with the MSP where the high-scale patch can be efficiently computed using a low-resolution image space. The local similarity between two images using the MSP measure is computed, as follows,

$$S_{\text{msp}}(I, A_a, x) = \sum_s S(I^{(s)}, A_a^{(s)}, x) \quad (6.6)$$

where  $I^{(s)} = I * \text{Gaussian}(0, \sigma_s)$  is the target image from  $s$  scale-space, which is computed from the convolution of the target image with Gaussian kernel function with scale  $s$ . Here, we compute the local similarity in multi-scale image using the conditional probability of the images,

$$S(I^{(s)}, A_a^{(s)}, x) = p(i_x | j_x) = \frac{p(i_x, j_x)}{p(j_x)} \quad (6.7)$$

where  $i_x = I^{(s)}(x)$  and  $j_x = A_a^{(s)}(x)$  and the conditional image probability is obtained from the joint and marginal image probability, which can be calculated using the Parzen window estimation [74].

For each patient, the Roadmap dataset was then registered to the LGE MRI dataset using the DICOM header data, and then refined by affine and nonrigid registration steps [59]. The resulting transformation was applied to the MA-WHS derived cardiac anatomy to define the endocardial LA boundary and PV on the LGE MRI dataset for each patient.

#### 6.2.4 IMAGE OVER-SEGMENTATION USING SLIC SUPER-PIXELS

We used a Simple Linear Iterative Clustering (SLIC) based super-pixel method [46] to over-segment LGE MRI images in order to separate potential enhanced atrial scarring regions from other tissues. The SLIC method has been used successfully in many medical image analysis problems [75]–[77]. The SLIC super-pixel method, which is an unsupervised learning [78] based method, groups pixels into perceptually meaningful patches with similar size, which can be used to replace the regular pixel grid. Consequently, the derived super-pixel patches can capture and mitigate image redundancy, and therefore provide a significant primitive from which image features can be calculated effectively and efficiently.

In summary, super-pixel methods have been proven to have the following benefits: (1) super-pixels can adhere well to perceptually meaningful object boundaries in images; (2) super-pixels can reduce computational complexity of extracting image features; (3) for segmentation applications, super-pixels can improve performance while reducing the computation time [79]. In this study, we proposed to use a SLIC based super-pixel method, which has been successfully applied to solve various medical image analysis problems (e.g., [75], [80]). It has also demonstrated better segmentation accuracy and superior adherence to object boundaries, and it is faster and more memory efficient



compared to other state-of-the-art super-pixels methods [46]. Based on local k-means clustering, the SLIC method iteratively groups pixels into super-pixels. The clustering proximity is estimated in both intensity and spatial domains that is

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{S}\right)^2} m^2, \quad (6.8)$$

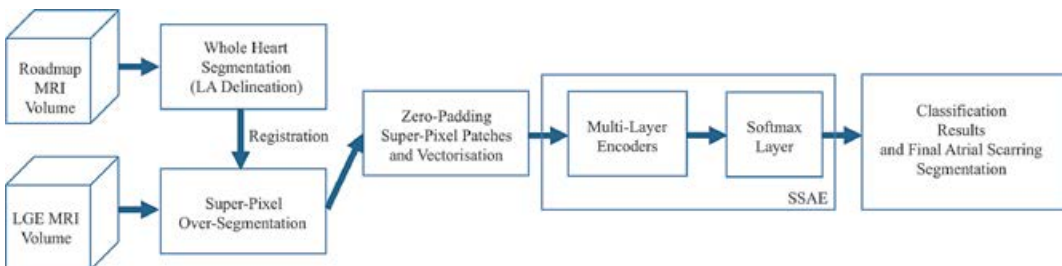
in which  $d_c = \sqrt{(I_j - I_i)^2}$  measures the pixel intensity difference of a gray scale image and  $d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$  describes the spatial distance between each pixel and the geometric center of the super-pixel. SLIC is initialized by sampling the target slice of the LGE MRI image into a regular grid space with grid interval of  $S$  pixels. To speed up the iteration, SLIC limits the size of search region of similar pixels to  $2S \times 2S$  around the super-pixel center (namely, local k-means clustering). In addition, parameter  $m$  balances the weighting between intensity similarity  $d_c$  and spatial proximity  $d_s$ . In this study we initialized  $S$  to 4 pixels that is  $2.8 \times 2.8 \text{ mm}^2$  considering the LA wall thickness is approximately 3mm [23], [81], and also take into account that the super-pixel size is still large enough to extract statistics of the grouped pixel intensities. In addition,  $m$  was chosen by visual inspection of the over-segmented results, and it was fixed when the super-pixel results adhered well with the LA wall boundary.

### 6.2.5 ATRIAL SCARRING SEGMENTATION BASED ON SUPER-PIXELS CLASSIFICATION

The LA regions, including LA wall, blood pool and atrial scarring, have been over-segmented into super-pixel patches. Next the atrial scarring segmentation will be based on classification of these super-pixel patches. This can be categorized as a supervised learning based segmentation method (Figure 6.1). First, we need to construct a training dataset with ground truth labelling, that is, each super-pixel patch will be labelled as scar or non-scar. Second, we can train our classifier based on the paired super-pixel patches and their labels. Finally, the trained model will be used to predict the atrial scarring in new input LGE MRI images.

#### 6.2.5.1 Training Data Construction and Ground Truth Definition

In order to train the following classifier, we built a training dataset containing enhanced and non-enhanced super-pixel patches. This has been done by (1) experienced expert-cardiologists specialized in cardiac MRI performing manual mouse clicks to select the enhanced scar regions; (2) combining the mouse clicks and SLIC segmentation to label the enhanced super-pixels; (3) applying morphological dilation (3mm) to the segmented endocardial LA boundary and PV from MA-WHS to extract the LA wall and PV; (4) finding the overlapped regions of the LA wall and PV and



**FIGURE 6.1** Whole pipeline of the atrial scarring segmentation workflow.



the labelled enhanced super-pixels; and (5) labelling the other super-pixels overlapped with LA wall and PV as non-enhancement. Details of each step are given as following:

1. Manual mouse clicks: Instead of manually drawing the boundaries of the enhanced atrial scarring regions, we asked experienced cardiologists specialized in cardiac MRI to perform manual mouse clicks on the LGE MRI images to label the regions that they believed to be enhanced (i.e., atrial scarring tissue). This is because manual boundary drawing of enhancement on the thin LA wall is a very challenging task and subject to large inter- and intra-observer variances. Mouse clicks on the enhancement regions are much easier and much more efficient. The manual mouse clicks were done on the original LGE MRI images without the super-pixel grid overlaid. This is because: (a) the mouse clicks will not be biased by super-pixel patches and (b) the super-pixel grid may reduce the visibility of the enhancement on LGE MRI images.
2. The coordinates of the mouse clicks were used to select the enhanced super-pixels. Because our cardiologists performed the mouse clicks on the original LGE MRI images without having prior knowledge about the super-pixels, we asked them to have relatively dense mouse clicks. These mouse clicks will ensure all the enhanced regions can be included, but only one mouse click will be taken into account if multiple clicks dwell in the same super-pixel.
3. The endocardial LA boundary and PV were extracted using our MA-WHS method. We then applied a morphological dilation to extract the LA wall and PV assuming that the thickness of LA wall is 3mm. The blood pool regions were extracted by a morphological erosion (5mm) from the endocardial LA boundary. And the pixel intensities were normalized according to the mean and standard deviation of the blood pool intensities [39].
4. We masked the selected enhanced super-pixels [derived from step (2)] using the LA wall and PV segmentation. Only the super-pixels having a defined overlap with the LA wall and PV segmentation were selected as enhancement for building the training data (overlapping ratio was set to  $\geq 20\%$ ). Other super-pixels (overlapping ratio  $< 20\%$ ) were discarded as they were considered as enhancement from other substructures of the heart (such as the mitral valve and aorta) but not enhancement of the LA wall and PV. Although we assumed that the LA wall thickness is 3mm, our enhanced super-pixels are not restricted to this wall thickness.
5. The other super-pixels overlapped with the LA wall and PV but not selected as enhancement were considered as non-enhancement (overlapping ratio was set to  $\geq 20\%$ ).

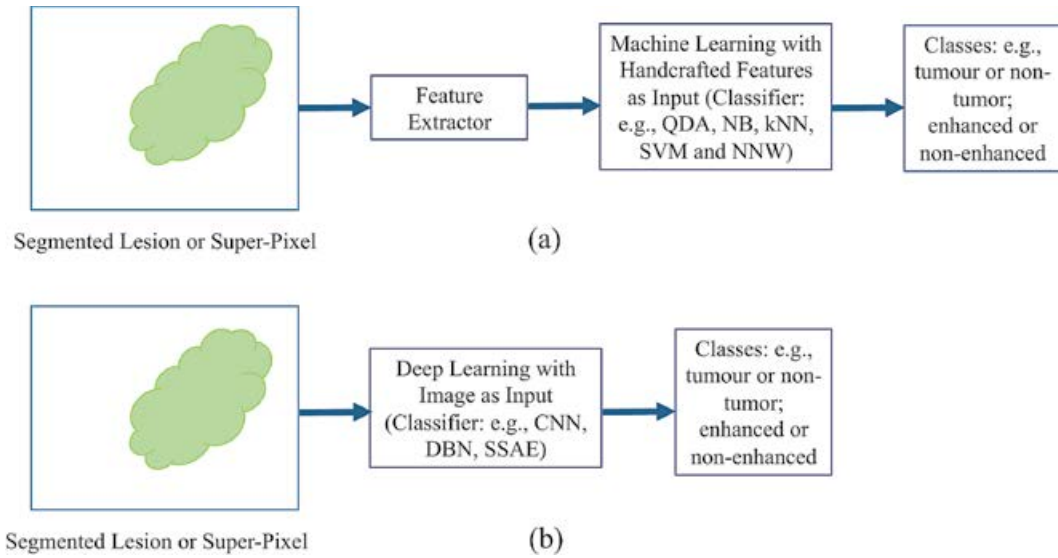
By performing these five steps, we constructed a training dataset that contains super-pixel patches labelled either enhancement or non-enhancement within the LA wall and PV.

In order to form the ground truth of the enhanced atrial scarring on the LGE MRI images, we performed the following two further steps:

6. Once we extracted the enhanced super-pixels, they were combined to create a binary image for each slice (i.e., 1 for enhanced super-pixels and 0 for unenhanced).
7. The binary image was overlaid on the original LGE MRI images, and our cardiologists performed manual corrections to create the final boundaries (ground truth) of the enhanced atrial scarring. In so doing, we minimized the bias toward a better performance of the segmentation using classified super-pixels.

#### 6.2.5.2 Deep Learning Using Stacked Sparse Auto-Encoders

The segmentation of the atrial scarring is performed using the classification of the over-segmented super-pixels. Conventional classification tasks are normally solved using particular machine



**FIGURE 6.2** Machine learning and deep learning for classification. (a) Conventional machine learning for a general classification task (i.e., machine learning with feature input or feature-based machine learning) in the field of medical image analysis. Handcrafted features (e.g., contrast, circularity, and effective diameter) are extracted from a segmented lesion or super-pixels in an image. (b) Deep learning based classification: deep learning with image input. Thus, one of the major and essential differences between machine learning and deep learning is the direct training of pixels in images.

learning algorithms. In this use, a machine learning algorithm is often called a classifier. Widely used classification methods include quadratic discriminant analysis (QDA) [82], Naïve Bayes (NB) [83], k-Nearest Neighbor algorithm (kNN) [84], Support Vector Machines (SVM) [85] [86], and shallow Neural Networks (NNW) [87] [88]. A crucial step in the design of such a classification framework is the extraction of discriminant features from the images (or over-segmented super-pixel patches in our case). This process is normally done by human researchers or a particular automated filtering pre-processing that is denoted as handcrafted features (Figure 6.2a). However, meaningful or task-related handcrafted features are designed mostly by human experts on the basis of their domain knowledge, making it hard for non-experts to exploit machine learning techniques for their own studies. Moreover, filtering based feature engineering methods may be biased to the particular pre-defined basis function that is only sensitive to specific features of the images.

Recently, deep learning based methods (e.g., Convolutional Neural Networks [CNN] and Deep Belief Network [DBN]) are rapidly becoming the state of the art, leading to superior performance in different medical image reconstruction, segmentation, and analysis applications [89]–[91]. Compared to conventional machine learning based classification methods, deep learning has overcome the obstacles of handcrafted features by incorporating the feature engineering step into a learning step. That is, instead of extracting features manually or using pre-defined filtering, deep learning requires only a set of data with minimal preprocessing and discovers the informative representations in a self-taught manner [92] (Figure 6.2b). In so doing, the burden of feature engineering has shifted from humans (subjective) to computers (objective), allowing non-experts in machine learning to effectively use deep learning for their own applications, especially in medical image analysis [92].

In this study, after we obtained the over-segmented super-pixels, the Stacked Sparse Auto-Encoders (SSAE) [93] were used to perform the classification. The SSAE were initially pre-trained

in an unsupervised manner without using the labels of the super-pixels. An auto-encoder neural network tries to learn an approximation to the identity function to replicate its input at its output using a back-propagation algorithm, that is  $\hat{X} = h_{w,b}(X) \approx X$ , in which  $X = \{x_1, x_2, \dots, x_m\}$ ,  $X \in \mathcal{R}^{n \times m}$  is a matrix storing all the input training vectors  $x_i \in \mathcal{R}^n$ . Each input vector  $x_i$  was formed by: (1) zero-padding all the super-pixels into a  $20 \times 20$  matrix, which is the smallest bounding box for the largest super-pixel dimensions, and (2) vectorizing the  $20 \times 20$  matrix into a  $400 \times 1$  vector. The cost function of this pre-training can be written as

$$\arg \min_{W^l} J_a(W^l) = \frac{1}{2m} \sum_{i=1}^m \|\hat{x}_i - x_i\|_2^2 + \frac{\lambda}{2} \|W^l\|_2^2 + \beta \sum_{j=1}^k \text{KL}(\rho \parallel \hat{\rho}_j), \quad (6.9)$$

where  $m$  is the number of input training vectors,  $k$  is the number of hidden nodes,  $\lambda$  is the coefficient for the  $L_2$  regularization term,  $\beta$  is the weight of sparsity penalty, KL is the Kullback-Leibler divergence function  $\text{KL}(\rho \parallel \hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j}$ ,  $\rho$  is sparsity parameter that specifies the desired level of sparsity,  $\hat{\rho}_j$  is probability of firing activity that is  $\hat{\rho}_j = \frac{1}{m} \sum_{i=1}^m h_j(x_i)$ . The unsupervised pre-training is performed one layer at a time by minimizing the error in reconstructing its input and learning an encoder and a decoder, which yields an optimal set of weights  $W$  and biases  $b$  stored in  $W^l$ . If the number of hidden nodes  $k$  is less than the number of visible input nodes  $n$ , then the network is forced to learn a compressed and sparse representation of the input [93].

Second, a Softmax layer was added as the activity classification model  $h_\theta(x_i)$  to accomplish the super-pixels classification task [93]. In addition, it can be jointly trained with the SSAE during fine-tuning of the parameters with labeled instances in a *supervised fashion*. The weight matrix  $\theta$  is obtained by solving the convex optimization problem as follows.

$$\arg \min_{\theta} J_s(\theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{c=1}^C \mathbf{1}\{y_i = c\} \times \log P(y_i = c | x_i; \theta) + \frac{\lambda}{2} \|\theta\|_2^2, \quad (6.10)$$

where  $c \in \{1, C = 2\}$  is the class label,  $\tilde{X} = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$  represents a set of labeled training instances, and the last term for the  $L_2$  regularization.

Finally, fine-tuning was applied to boost the classification performance, and it treats all layers of the SSAE and the Softmax layer as a single model and improves all the weights of all layers in the network by using the backpropagation technique [93].

### 6.2.5.3 Hyper-Parameters Settings

Figure 6.3 shows the detailed network architecture of the implemented SSAE. Hyper-parameters were not optimized explicitly but were determined via trial and error. Here are the final defined

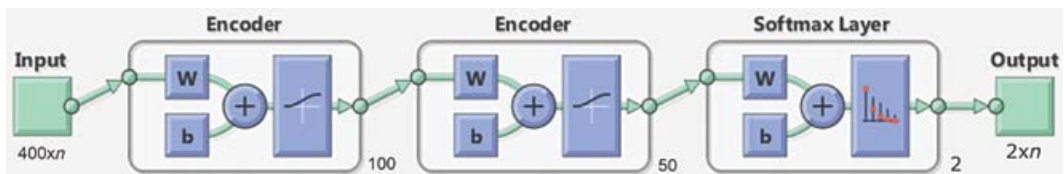


FIGURE 6.3 Detailed network architecture of the implemented SSAE.

hyper-parameters (values) used in this study: maximum epochs of the SSAE (200), maximum epochs of the Softmax and fine-tuning (500), hidden layers size of the SSAE (100 and 50), sparsity parameter  $\rho$  (0.1), sparsity penalty  $\beta$  (5),  $L_2$  regularization term  $\lambda$  for the SSAE and the Softmax (0.0001).

### 6.2.6 VALIDATION APPROACHES

#### 6.2.6.1 Validation for the Whole Heart Segmentation

One experienced cardiologist (>5 years' experience and specialized in cardiac MRI) manually segmented the endocardial LA boundary and labelled the PV slice-by-slice in the LGE MRI images for all the patients. A second senior cardiologist (>25 years' experience and specialized in cardiac MRI) confirmed the manual segmentation. The evaluation and validation of our MA-WHS has been done against this manual segmentation, which is assumed to be the ground truth. We used three metrics: Dice Score [94], [95], Hausdorff Distance (HD) [96] and Average Surface Distance (ASD). Dice Score, which is defined as  $\text{Dice Score} = \frac{2 \times |F_{\text{Manual}} \cap F_{\text{Auto}}|}{|F_{\text{Manual}}| + |F_{\text{Auto}}|}$  ( $F_{\text{Manual}}$ : ground truth segmentation;  $F_{\text{Auto}}$ : automatic segmentation;  $|\bullet|$ : the number of pixels assigned to the segmentation), measures the overlap between two segmentations. The higher the values of Dice Score, the better the overall performance of the segmentation will be. HD and ASD measure the boundary distance (in mm) between two contours of segmentation, which are defined as

$$\text{HD}(P_{\text{Manual}}, P_{\text{Auto}}) = \max(d(P_{\text{Manual}}, P_{\text{Auto}}), d(P_{\text{Auto}}, P_{\text{Manual}}))$$

$$\text{where } d(P_{\text{Manual}}, P_{\text{Auto}}) = \sup_{p_m \in P_{\text{Manual}}} \inf_{p_a \in P_{\text{Auto}}} \|p_m - p_a\| \quad (6.11)$$

$$d(P_{\text{Auto}}, P_{\text{Manual}}) = \sup_{p_a \in P_{\text{Auto}}} \inf_{p_m \in P_{\text{Manual}}} \|p_m - p_a\|,$$

$$\text{ASD} = \frac{1}{2} \left( \frac{\sum_{p_m \in P_{\text{Manual}}} \min_{p_a \in P_{\text{Auto}}} \|p_m - p_a\|}{\sum_{p_m \in P_{\text{Manual}}} 1} + \frac{\sum_{p_a \in P_{\text{Auto}}} \min_{p_m \in P_{\text{Manual}}} \|p_m - p_a\|}{\sum_{p_a \in P_{\text{Auto}}} 1} \right), \quad (6.12)$$

in which  $P_{\text{Manual}} = \{p_{m1}, \dots, p_{mn}\}$  and  $P_{\text{Auto}} = \{p_{a1}, \dots, p_{an}\}$  are two finite point sets of the two segmented contours (using the ground truth segmentation and automatic segmentation),  $\|\bullet\|$  denotes  $L_2$  norm, sup denotes supremum and inf denotes infimum. The lower the values of HD and ASD, the better agreement between manual delineation and fully automatic segmentation.

#### 6.2.6.2 Validation for the Atrial Scarring Validation

We evaluated our SSAE based classification by: (i) leave-one-patient-out cross-validation (LOO CV) [78], [97], which provides an unbiased predictor and is capable of creating sufficient training data for studies with small sample size [98] and (ii) the cross-validated classification accuracy, sensitivity, specificity [76], [77], [99], and average area under the receiver operating characteristic (ROC) curve (AUC) [88], [100]. For evaluation of the atrial scarring segmentation, we used the Dice score.

**TABLE 6.3****Quantitative Results of the Whole Heart Segmentation and Atrial Scarring Segmentation.**

Tasks	Acc.	Sens.	Spec.	AUC	Dice	HD (mm)	ASD (mm)
MM-WHS	-	-	-	-	0.90±0.12	9.53±6.01	1.47±0.89
Atrial Scarring Segmentation	0.91	0.95	0.75	0.95	0.82±0.05	-	-

## 6.3 RESULTS

### 6.3.1 WHOLE HEART SEGMENTATION

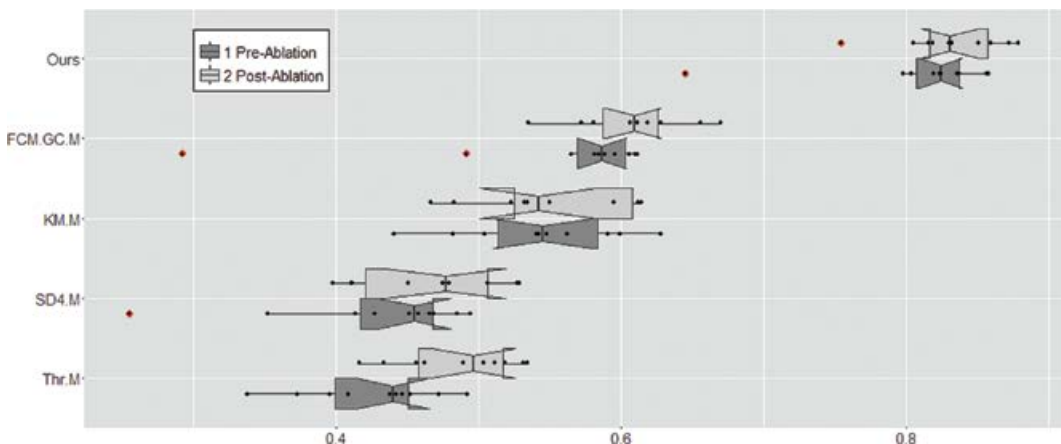
The quantitative evaluations show that the MA-WHS based method achieved  $0.90 \pm 0.12$  Dice score,  $9.53 \pm 6.01$  mm HD, and  $1.47 \pm 0.89$  mm ASD (Table 6.3).

### 6.3.2 ATRIAL SCARRING SEGMENTATION

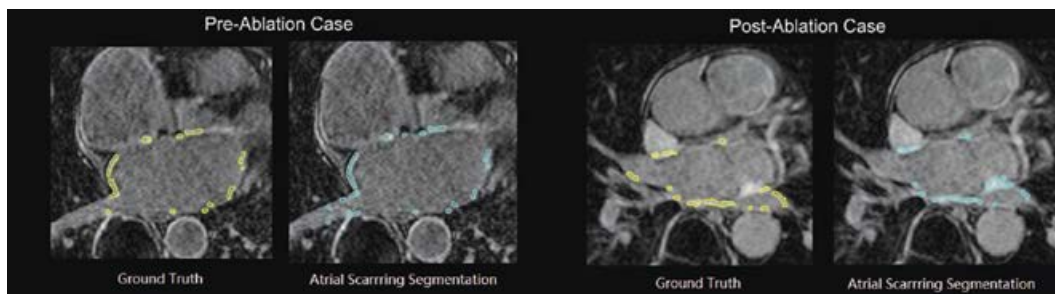
For the SSAE based super-pixels classification, we obtained LOO CV accuracy of 0.91, sensitivity of 0.95, specificity of 0.75, AUC of 0.95, and the Dice score for the final atrial scarring segmentation was found to be  $0.82 \pm 0.05$  (Table 6.3).

In addition, for the atrial scarring segmentation, we compared our fully automatic framework with existing semi-automatic methods with manually delineated anatomical structure of the LA and PVs. The four methods we compared in this study were described in the benchmarking work [39], namely simple thresholding (Thr), conventional standard deviation (4 SDs were tested, i.e., SD4), k-means clustering (KM), and fuzzy c-means clustering with graph-cuts (FCM+GC). Figure 6.4 shows that our fully automatic framework obtained more accurate and more consistent results across 20 AF patient cases (Figure 6.4, red dots represent outliers).

Figure 6.5 demonstrates that qualitatively our fully automatic atrial scarring segmentation is in accordance with the manual segmented results. However, if there are enhancements from the nearby mitral valve or blood pool regions, our method may misclassify them as enhanced atrial fibrosis that is the major contribution for the false positives.



**FIGURE 6.4** Boxplot for the comparison results of the Dice scores obtained by our fully automatic framework and other four methods (Thr, SD4, KM, and FCM+GC) with manual delineated LA wall and PVs (+M).



**FIGURE 6.5** Final atrial scarring segmentation results (cyan regions) for an example pre-ablation (left) and an example post-ablation (right) case compared to the ground truth (yellow regions).

## 6.4 DISCUSSION

In this study, we developed a novel fully automatic segmentation pipeline to detect enhanced atrial scarring in LGE MRI images. The achievements of this work are:

- a. MA-WHS of Roadmap MRI for cardiac anatomy segmentation: a MA-WHS segmentation method has been proposed with a new MSP-LF scheme to define the LA and PV geometry while minimizing the mis-segmentation from confounding tissues of other substructures of the heart;
- b. Super-pixel classification based method for atrial scar segmentation from LGE MRI: the super-pixel classification has two uses. Firstly, it has been used as a by-product tool, which can help cardiologists to easily construct a manual ground truth segmentation by indicating enhanced scarring regions with a number of mouse clicks. This is subject to limited manual corrections and is more efficient and reliable than direct manual drawing around enhanced region borders, especially when the boundaries are highly irregular and ill-defined. Secondly, based on the trained classifier, the super-pixel classification allows a fully automatic atrial scarring segmentation to be achieved by running super-pixel based over-segmentation and classification for the new input LGE MRI data;
- c. SSAE based deep learning classifier: the SSAE based classifier has been proposed and implemented that achieved high classification accuracy without any handcrafted features needed;
- d. Validation: our developed fully automatic pipeline was tested and validated directly on real clinical datasets.

Compared to manual ground truth construction that took 25mins to 50mins per patient case, the SSAE based prediction only took  $5.8 \pm 0.9$ secs to segment one patient case, while for a single loop of the LOO CV, the training on 20 patients took about 20 hours. All the experiments were performed using a Windows 7 workstation with 6-cores 1.9GHz Intel® Xeon® E5-2609v3/64 GB RAM and NVIDIA GeForce® GTX Titan X with an in-house Matlab implementation. With the proven efficiency and efficacy, the application of our method to real clinical problems is straightforward. Overall, results of this study offer compelling evidence that our fully automatic pipeline is capable of detecting enhanced atrial scarring from LGE MRI images acquired from a longstanding persistent AF cohort.

Segmentation of the atrial scarring from LGE MRI images is very challenging. This is not only because the atrial scarring is difficult to distinguish in the thin LA wall but also because the image quality can be poor due to motion artefacts, noise contamination, and contrast agent wash-out during the long acquisition. Moreover, the enhancement from the surrounding tissues (i.e., other substructures of the heart or non-heart tissues) and enhanced blood flow are confounding issues for atrial scarring segmentation and result in increased false positives. For example, the aortic wall is generally enhanced and when close to the LA, can be mistaken for atrial scarring and contribute a



false positive result. Moreover, false positives can be increased from the misclassification of other enhanced regions, for example, other non-heart tissues and other substructures of the heart or fat tissues surrounding the LA. However, most of these confounding enhancement regions can be distinguished subject to accurate heart anatomy delineation using our MA-WHS. Another source of artefact originates from the respiratory navigator and results in enhanced signal from blood flow in the right pulmonary veins. Recent advances in sequence design have reduced navigator artefacts considerably [52], [101]. Due to the subjective understanding of the LGE MRI images, our cardiologists may also miss labelling some enhanced regions.

In this study, we showed good accuracy for segmenting the LA and proximal PVs. The segmentation accuracy in the more distal veins was less good, but this is not an issue for clinical ablation for AF patients as the ablation points or clamps are never placed far away from the LA chamber. In addition, the segmentation of the PVs at the more distal regions might just introduce more artefacts from the enhanced navigator beam regions without improving the accuracy of the actual atrial scarring segmentation. There are some previous studies that endeavoured to segment the detailed sub-branches of PVs and especially at the more distal regions (e.g., [42] using MRA). However, this can only improve the accuracy of LA+PV delineation against manual delineated ground truth, but may not be a benefit for the final atrial scarring segmentation. For the post-ablation cases, the enhancement at or near the PVs can indicate the efficacy of the treatment, for example, identify the gaps in the ablation line. Therefore, based on our experiments and observations, a robust segmentation method to delineate PV anatomy variations is still in demand, and it is a more important task than accurate delineation of the detailed sub-branches of the PVs.

In our study, 2D SLIC was applied instead of using its 3D version. This is because: (1) our 3D LGE MRI data are anisotropic with fine in-plane resolution and relatively coarse resolution in the third dimension (i.e.,  $0.75 \times 0.75 \times 2 \text{ mm}^3$ ); therefore, 2D SLIC can provide better adherence to the more detailed anatomical edges shown in the 2D in-plane slice, which has a higher spatial resolution; (2) the clinical image viewing is still a slice-by-slice procedure in 2D, and the 2D SLIC results are more intuitive for manual super-pixel labelling by our expert-cardiologists, which is a necessary step for atrial scarring ground truth construction; (3) the 2D SLIC processing is efficient ( $8.6 \pm 2.6$  seconds per 2D slice and  $\sim 5$  mins per patient case) although a slice-by-slice computing was performed. In addition, for our application, we did not perform further optimization for the two parameters of the SLIC algorithm (i.e., the size of the super-pixels  $S$  and the compactness  $m$ ). In our study, the size of the super-pixels was restricted by the LA wall thickness and  $S$  was initialized to 4 pixels. The compactness term  $m$  controls the regularity of the super-pixel shapes (i.e., the resulting super-pixels can adhere more tightly to object boundaries in the image when  $m$  is small), but these super-pixels have less regular shape. We chose  $m = 4$  based on visual inspections of the over-segmentation results.

As is well known, LGE studies are specifically designed to highlight fibrosis and scarring, and to null all signal from normal tissue. Although it is possible to see the cardiac anatomy in LGE images, the SNR is limited and accurate and automatic segmentation of the LA and PV walls would be very difficult. To the best of our knowledge, all previous studies have relied on manual drawing on LGE images or have used MRA images or b-SSFP based Roadmap scans for the delineation of the LA and PV anatomy. In this current study, we use the 3D Roadmap images to resolve the LA and PV anatomy. For each AF patient in our study, the Roadmap data took  $\sim 6.5$  mins to acquire (241 cardiac cycles at 800ms per cardiac cycle with a typical respiratory efficiency of 50%). In subsequent work, we have reduced the spatial resolution of the 3D Roadmap to  $1.8 \times 1.8 \times 1.8 \text{ mm}^3$ , which reduces the acquisition duration to  $\sim 4$  mins while preserving good results. In addition, although the requirement for a Roadmap acquisition might be a limitation, in practice, many patients have these as part of their standard clinical cardiac MRI examination.

The comparison work reported in this study is limited. While we validated our techniques against manual segmentation, we also wished to compare our results with other previously published techniques. A number of advanced techniques have been proposed, such as unsupervised learning based clustering and graph-cuts based methods [39]. However, implementation of these is difficult as the



fine-tuned hyper-parameters used are not always clearly described and the methodologies cannot be reproduced exactly. Moreover, our patient cohort is different from that in which these algorithms were optimized and tested. In this study, we have therefore only compared our technique against the simple thresholding and conventional standard deviation based methods, as these have fully standard implementations. When compared to manual segmentation (ground truth) in post-ablation scans, these standard techniques gave median DICE of 38–48%, while our fully automatic technique achieved a median DICE of 82%. The results that we obtained here with the standard techniques are similar to those reported with these same techniques in the benchmarking study described in [39], while the latter score is similar to the best-performing methods reported in that same study.

In general, the segmentation algorithms performed better on post-ablation LGE MRI scans compared to the performance on pre-ablation ones. This is likely due to better image quality post-ablation (when the heart has reverted to sinus rhythm) and to higher levels of fibrosis. For the LA+PV segmentation, many of the algorithms previously published rely on manual segmentation. In our study, we have compared the efficacy of the atrial scarring segmentation algorithms using automatically segmented LA+PV geometry against the same algorithms with manually delineated geometry showing very similar results, thus confirming that our fully automatic MA-WHS method is capable of accurately defining the relevant cardiac anatomy.

In addition, compared to a method of using multi-atlas segmentation for the four chambers [42], our MA-WHS method obtained superior DICE (90% vs. 86% reported by [42]) and ASD (1.5 mm vs. 1.8 mm reported by [42]). Even compared with their results obtained after refinement by MRA [42], our MA-WHS still has comparable results. For the segmentation of the atrial scarring, in our study, the results of the standard segmentation techniques in pre-ablation cases were better than those reported by the benchmarking study [39]. This is likely to be due to our patients all having longstanding persistent AF and therefore having higher levels of pre-ablation scarring. For the post-ablation cases, we have obtained comparable median DICE compared to the best performing algorithm [39]. Of note is that in the benchmarking study the variances of all of the techniques tested are large, while in our manuscript, the results are more consistent with a relatively small variance. This may be due to our patient cohort being more tightly defined, whereas in the previous study, datasets were analyzed from patients at multiple institutions using a variety of imaging protocols. It may also be that our automatic technique is based on supervised learning while previous methods are unsupervised, and the derived model parameters may not be optimized for all the patient cases.

In our study, although techniques like SLIC and super-pixel based classification may be very well known and widely used methods [61], [80], [102]–[105], they are reliable algorithms that, when uniquely combined in the proposed pipeline, enable fully automatic segmentation and assessment of atrial scarring. This is an important advance, as LGE MRI is becoming a preferred method for noninvasive imaging of atrial scarring.

One possible limitation of our study is that the SSAE based classifier has many hyper-parameters, which need to be carefully tuned (e.g., maximum epochs of the SSAE, maximum epochs of the Softmax and fine-tuning, hidden layers size of the SSAE, sparsity parameter  $\rho$ , sparsity penalty  $\beta$ , L2 regularization term  $\lambda$  for the SSAE and the Softmax). Currently these hyper-parameters were tuned via trial and error, which may limit the final classification accuracy.

## 6.5 CONCLUSION

To the best of our knowledge, this is the first study that developed a deep learning based fully automatic segmentation pipeline for atrial scarring segmentation with quantitative validation on LGE MRI scans. The proposed pipeline has demonstrated an effective and efficient way to objectively segment and assess the atrial scarring. The evaluation has been done on 20 LGE MRI scans in longstanding persistent AF patients that contain both pre-ablation and post-ablation cases. The validation results have shown that both our MA-WHS and super-pixel classification based atrial

scarring segmentation have obtained satisfactory accuracy. Based on the results, we can envisage a straightforward deployment of our framework for clinical usage.

## Acknowledgments

The studies described in this chapter were funded by the British Heart Foundation Project Grant (Project Number: PG/16/78/32402), the NIHR Cardiovascular Biomedical Research Unit, Royal Brompton Hospital & Harefield NHS Foundation Trust and Imperial College London, and the Chinese NSFC research fund (81301283) and the NSFC-RS fund (81511130090). Data were obtained during the NIHR Efficacy and Mechanism Evaluation Programme (Project Number: 12/127/127).

## References

1. Y. Miyasaka, "Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence," *Circulation*, vol. 114, no. 2, pp. 119–125, Jul. 2006.
2. S. S. Chugh, R. Havmoeller, K. Narayanan, D. Singh, M. Rienstra, E. J. Benjamin, R. F. Gillum, Y.-H. Kim, J. H. McAnulty, Z.-J. Zheng, M. H. Forouzanfar, M. Naghavi, G. A. Mensah, M. Ezzati, and C. J. L. Murray, "Worldwide epidemiology of atrial fibrillation: A global burden of disease 2010 study," *Circulation*, vol. 129, no. 8, pp. 837–847, 2014.
3. M. Allesie, "Electrical, contractile and structural remodeling during atrial fibrillation," *Cardiovasc. Res.*, vol. 54, no. 2, pp. 230–246, 2002.
4. S. Stewart, C. L. Hart, D. J. Hole, and J. J. V. McMurray, "A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study," *Am. J. Med.*, vol. 113, no. 5, pp. 359–364, 2002.
5. C. T. January, L. S. Wann, J. S. Alpert, H. Calkins, J. E. Cigarroa, J. C. Cleveland, J. B. Conti, P. T. Ellinor, M. D. Ezekowitz, M. E. Field, K. T. Murray, R. L. Sacco, W. G. Stevenson, P. J. Tchou, C. M. Tracy, and C. W. Yancy, "2014 AHA/ACC/HRS Guideline for the Management of Patients with Atrial Fibrillation: Executive Summary," *J. Am. Coll. Cardiol.*, vol. 64, no. 21, pp. 2246–2280, Dec. 2014.
6. S. H. Hohnloser, K.-H. Kuck, and J. Lilienthal, "Rhythm or rate control in atrial fibrillation—Pharmacological Intervention in Atrial Fibrillation (PIAF): a randomised trial," *Lancet*, vol. 356, no. 9244, pp. 1789–1794, Nov. 2000.
7. I. C. Van Gelder, V. E. Hagens, H. A. Bosker, J. H. Kingma, O. Kamp, T. Kingma, S. A. Said, J. I. Darmanata, A. J. M. Timmermans, J. G. P. Tijssen, and H. J. G. M. Crijns, "A comparison of rate control and rhythm control in patients with recurrent persistent atrial fibrillation," *N. Engl. J. Med.*, vol. 347, no. 23, pp. 1834–1840, Dec. 2002.
8. M. Haïssaguerre, P. Jaïs, D. C. Shah, A. Takahashi, M. Hocini, G. Quiniou, S. Garrigue, A. Le Mouroux, P. Le Métayer, and J. Clémenty, "Spontaneous initiation of atrial fibrillation by ectopic beats originating in the pulmonary veins," *N. Engl. J. Med.*, vol. 339, no. 10, pp. 659–666, Sep. 1998.
9. D. J. Wilber, C. Pappone, P. Neuzil, A. De Paola, F. Marchlinski, A. Natale, L. Macle, E. G. Daoud, H. Calkins, B. Hall, V. Reddy, G. Augello, M. R. Reynolds, C. Vinekar, C. Y. Liu, S. M. Berry, D. A. Berry, and ThermoCool AF Trial Investigators, "Comparison of antiarrhythmic drug therapy and radiofrequency catheter ablation in patients with paroxysmal atrial fibrillation: a randomized controlled trial," *JAMA*, vol. 303, no. 4, pp. 333–340, Jan. 2010.
10. F. Ouyang, R. Tilz, J. Chun, B. Schmidt, E. Wissner, T. Zerm, K. Neven, B. Köktürk, M. Konstantinidou, A. Metzner, A. Fuernkranz, and K. H. Kuck, "Long-term results of catheter ablation in paroxysmal atrial fibrillation: lessons from a 5-year follow-up," *Circulation*, vol. 122, no. 23, pp. 2368–2377, 2010.
11. R. Weerasooriya, P. Khairy, J. Litalien, L. MacLe, M. Hocini, F. Sacher, N. Lellouche, S. Knecht, M. Wright, I. Nault, S. Miyazaki, C. Scavee, J. Clémenty, M. Haïssaguerre, and P. Jais, "Catheter ablation for atrial fibrillation: are results maintained at 5 years of follow-up?" *J. Am. Coll. Cardiol.*, vol. 57, no. 2, pp. 160–166, 2011.
12. J. R. Edgerton, J. H. McClelland, D. Duke, M. W. Gerdisch, B. M. Steinberg, S. H. Bronleewe, S. L. Prince, M. A. Herbert, S. Hoffman, and M. J. Mack, "Minimally invasive surgical ablation of atrial fibrillation: six-month results," *J. Thorac. Cardiovasc. Surg.*, vol. 138, no. 1, pp. 109–113; discussion 114, Jul. 2009.
13. J. R. Edgerton, W. M. Jackman, C. Mahoney, and M. J. Mack, "Totally thorascopic surgical ablation of persistent AF and long-standing persistent atrial fibrillation using the 'Dallas' lesion set," *Heart Rhythm*, vol. 6, no. 12 Suppl, pp. S64–S70, Dec. 2009.

14. S. P. J. Krul, A. H. G. Driessen, W. J. van Boven, A. C. Linnenbank, G. S. C. Geuzebroek, W. M. Jackman, A. A. M. Wilde, J. M. T. de Bakker, and J. R. de Groot, "Thoracoscopic video-assisted pulmonary vein antrum isolation, ganglionated plexus ablation, and periprocedural confirmation of ablation lesions: first results of a hybrid surgical-electrophysiological approach for atrial fibrillation," *Circ. Arrhythm. Electrophysiol.*, vol. 4, no. 3, pp. 262–70, Jun. 2011.
15. L. V. A. Boersma, M. Castella, W. van Boven, A. Berruezo, A. Yilmaz, M. Nadal, E. Sandoval, N. Calvo, J. Brugada, J. Kelder, M. Wijffels, and L. Mont, "Atrial fibrillation catheter ablation versus surgical ablation treatment (FAST): a 2-center randomized clinical trial," *Circulation*, vol. 125, no. 1, pp. 23–30, Jan. 2012.
16. T. J. Badger, R. S. Oakes, M. Daccarett, N. S. Burgon, N. Akoum, E. N. Fish, J. J. E. Blauer, S. N. Rao, Y. Adjei-Poku, E. G. Kholmovski, S. Vijayakumar, E. V. R. Di Bella, R. S. MacLeod, and N. F. Marrouche, "Temporal left atrial lesion formation after ablation of atrial fibrillation," *Heart Rhythm*, vol. 6, no. 2, pp. 161–168, 2009.
17. H. Calkins, K. H. Kuck, R. Cappato, J. Brugada, A. J. Camm, S.-A. Chen, H. J. G. Crijns, R. J. Damiano, D. W. Davies, J. DiMarco, J. Edgerton, K. Ellenbogen, M. D. Ezekowitz, D. E. Haines, M. Haissaguerre, G. Hindricks, Y. Iesaka, W. Jackman, J. Jalife, P. Jais, J. Kalman, D. Keane, Y.-H. Kim, P. Kirchhof, G. Klein, H. Kottkamp, K. Kumagai, B. D. Lindsay, M. Mansour, F. E. Marchlinski, P. M. McCarthy, J. L. Mont, F. Morady, K. Nademanee, H. Nakagawa, A. Natale, S. Nattel, D. L. Packer, C. Pappone, E. Prystowsky, A. Raviele, V. Reddy, J. N. Ruskin, R. J. Shemin, H.-M. Tsao, D. Wilber, and Heart Rhythm Society Task Force on Catheter and Surgical Ablation of Atrial Fibrillation, "2012 HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design," *Heart Rhythm*, vol. 9, no. 4, pp. 632–696.e21, Apr. 2012.
18. H. Zhong, J. M. Lacomis, and D. Schwartzman, "On the accuracy of CartoMerge for guiding posterior left atrial ablation in man," *Heart Rhythm*, vol. 4, no. 5, pp. 595–602, May 2007.
19. E. J. Schmidt, R. P. Mallozzi, A. Thiagalingam, G. Holmvang, A. D'Avila, R. Guhde, R. Darrow, G. S. Slavin, M. M. Fung, J. Dando, L. Foley, C. L. Dumoulin, and V. Y. Reddy, "Electroanatomic mapping and radiofrequency ablation of porcine left atria and atrioventricular nodes using magnetic resonance catheter tracking," *Circ. Arrhythm. Electrophysiol.*, vol. 2, no. 6, pp. 695–704, Dec. 2009.
20. P. Kellman and A. E. Arai, "Cardiac imaging techniques for physicians: late enhancement," *J. Magn. Reson. Imaging*, vol. 36, no. 3, pp. 529–542, Sep. 2012.
21. D. C. Peters, J. V. Wylie, T. H. Hauser, K. V. Kissinger, R. M. Botnar, V. Essebag, M. E. Josephson, and W. J. Manning, "Detection of pulmonary vein and left atrial scar after catheter ablation with three-dimensional navigator-gated delayed enhancement MR imaging: initial experience," *Radiology*, vol. 243, no. 3, pp. 690–695, 2007.
22. R. S. Oakes, T. J. Badger, E. G. Kholmovski, N. Akoum, N. S. Burgon, E. N. Fish, J. J. E. Blauer, S. N. Rao, E. V. R. DiBella, N. M. Segerson, M. Daccarett, J. Windfelder, C. J. McGann, D. Parker, R. S. MacLeod, and N. F. Marrouche, "Detection and quantification of left atrial structural remodeling with delayed-enhancement magnetic resonance imaging in patients with atrial fibrillation," *Circulation*, vol. 119, no. 13, pp. 1758–1767, 2009.
23. M. Akçakaya, H. Rayatzadeh, T. Basha, S. N. Hong, R. H. Chan, K. V. Kissinger, T. H. Hauser, M. E. Josephson, and W. J. Manning, "Accelerated late gadolinium enhancement cardiac MR imaging with isotropic spatial resolution using compressed sensing: initial experience," *Radiology*, vol. 264, no. 3, pp. 691–699, 2012.
24. G. R. Vergara, S. Vijayakumar, E. G. Kholmovski, J. J. E. Blauer, M. a. Guttman, C. Gloschat, G. Payne, K. Vij, N. W. Akoum, M. Daccarett, C. J. McGann, R. S. MacLeod, and N. F. Marrouche, "Real-time magnetic resonance imaging-guided radiofrequency atrial ablation and visualization of lesion formation at 3 Tesla," *Heart Rhythm*, vol. 8, no. 2, pp. 295–303, 2011.
25. N. F. Marrouche, D. Wilber, G. Hindricks, P. Jais, N. Akoum, F. Marchlinski, E. Kholmovski, N. Burgon, N. Hu, L. Mont, T. Deneke, M. Duytschaever, T. Neumann, M. Mansour, C. Mahnkopf, B. Herweg, E. Daoud, E. Wissner, P. Bansmann, and J. Brachmann, "Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: the DECAAF study," *JAMA*, vol. 311, no. 5, pp. 498–506, 2014.
26. N. Akoum, M. Daccarett, C. McGann, N. Segerson, G. Vergara, S. Kuppahally, T. Badger, N. Burgon, T. Haslam, E. Kholmovski, R. MacLeod, and N. Marrouche, "Atrial fibrosis helps select the appropriate patient and strategy in catheter ablation of atrial fibrillation: a DE-MRI guided approach," *J. Cardiovasc. Electrophysiol.*, vol. 22, no. 1, pp. 16–22, 2011.

27. G. R. Vergara and N. F. Marrouche, "Tailored management of atrial fibrillation using a LGE-MRI based model: from the clinic to the electrophysiology laboratory," *J. Cardiovasc. Electrophysiol.*, vol. 22, no. 4, pp. 481–487, 2011.
28. L. C. Malcolm-Lawes, C. Juli, R. Karim, W. Bai, R. Quest, P. B. Lim, S. Jamil-Copley, P. Kojodjojo, B. Ariff, D. W. Davies, D. Rueckert, D. P. Francis, R. Hunter, D. Jones, R. Boubertakh, S. E. Petersen, R. Schilling, P. Kanagaratnam, and N. S. Peters, "Automated analysis of atrial late gadolinium enhancement imaging that correlates with endocardial voltage and clinical outcomes: a 2-center study," *Heart Rhythm*, vol. 10, no. 8, pp. 1184–1189, 2013.
29. D. Ravanelli, E. C. Dal Piaz, M. Centonze, G. Casagrande, M. Marini, M. Del Greco, R. Karim, K. Rhode, and A. Valentini, "A novel skeleton based quantification and 3-D volumetric visualization of left atrium fibrosis using late gadolinium enhancement magnetic resonance imaging," *IEEE Trans. Med. Imaging*, vol. 33, no. 2, pp. 566–576, 2014.
30. C. McGann, N. Akoum, A. Patel, E. Kholmovski, P. Revelo, K. Damal, B. Wilson, J. Cates, A. Harrison, R. Ranjan, N. S. Burgon, T. Greene, D. Kim, E. V. R. DiBella, D. Parker, R. S. MacLeod, and N. F. Marrouche, "Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI," *Circ. Arrhythmia Electrophysiol.*, vol. 7, no. 1, pp. 23–30, 2014.
31. A. S. Jadidi, H. Cochet, A. J. Shah, S. J. Kim, E. Duncan, S. Miyazaki, M. Sermesant, H. Lehrmann, M. Lederlin, N. Linton, A. Forclaz, I. Nault, L. Rivard, M. Wright, X. Liu, D. Scherr, S. B. Wilton, L. Roten, P. Pascale, N. Derval, F. Sacher, S. Knecht, C. Keyl, M. Hocini, M. Montaudon, F. Laurent, M. Haïssaguerre, and P. Jaïs, "Inverse relationship between fractionated electrograms and atrial fibrosis in persistent atrial fibrillation: combined magnetic resonance imaging and high-density mapping," *J. Am. Coll. Cardiol.*, vol. 62, no. 9, pp. 802–812, 2013.
32. V. Y. Reddy, E. J. Schmidt, G. Holmvang, and M. Fung, "Arrhythmia recurrence after atrial fibrillation ablation: can magnetic resonance imaging identify gaps in atrial ablation lines?" *J. Cardiovasc. Electrophysiol.*, vol. 19, no. 4, pp. 434–437, 2008.
33. T. J. Badger, M. Daccarett, N. W. Akoum, Y. a. Adjei-Poku, N. S. Burgon, T. S. Haslam, S. Kalvaitis, S. Kuppahally, G. Vergara, L. McMullen, P. A. Anderson, E. Kholmovski, R. S. MacLeod, and N. F. Marrouche, "Evaluation of left atrial lesions after initial and repeat atrial fibrillation ablation; Lessons learned from delayed-enhancement MRI in repeat ablation procedures," *Circ. Arrhythmia Electrophysiol.*, vol. 3, no. 3, pp. 249–259, 2010.
34. S. Kapa, B. Desjardins, D. J. Callans, F. E. Marchlinski, and S. Dixit, "Contact electroanatomic mapping derived voltage criteria for characterizing left atrial scar in patients undergoing ablation for atrial fibrillation," *J. Cardiovasc. Electrophysiol.*, vol. 25, no. 10, pp. 1044–1052, Oct. 2014.
35. F. Bisbal, E. Guiu, P. Cabanas-Grandío, A. Berruezo, S. Prat-Gonzalez, B. Vidal, C. Garrido, D. Andreu, J. Fernandez-Armenta, J. M. Tolosana, E. Arbelo, T. M. De Caralt, R. J. Perea, J. Brugada, and L. Mont, "CMR-guided approach to localize and ablate gaps in repeat AF ablation procedure," *JACC Cardiovasc. Imaging*, vol. 7, no. 7, pp. 653–663, 2014.
36. K. Higuchi, M. Akkaya, N. Akoum, and N. F. Marrouche, "Cardiac MRI assessment of atrial fibrosis in atrial fibrillation: implications for diagnosis and therapy," *Heart*, vol. 100, no. 7, pp. 590–596, 2014.
37. J. L. Harrison, H. K. Jensen, S. A. Peel, A. Chiribiri, A. K. Grondal, L. O. Bloch, S. F. Pedersen, J. F. Bentzon, C. Kolbitsch, R. Karim, S. E. Williams, N. W. Linton, K. S. Rhode, J. Gill, M. Cooklin, C. A. Rinaldi, M. Wright, W. Y. Kim, T. Schaeffter, R. S. Razavi, and M. D. O'Neill, "Cardiac magnetic resonance and electroanatomical mapping of acute and chronic atrial ablation injury: a histological validation study," *Eur. Heart J.*, vol. 35, no. 22, pp. 1486–1495, 2014.
38. D. Perry, A. Morris, N. Burgon, C. McGann, R. MacLeod, and J. Cates, "Automatic classification of scar tissue in late gadolinium enhancement cardiac MRI for the assessment of left-atrial wall injury after radiofrequency ablation," in *In SPIE Medical Imaging*, 2012, p. 83151D.
39. R. Karim, R. J. Housden, M. Balasubramaniam, Z. Chen, D. Perry, A. Uddin, Y. Al-Beyatti, E. Palkhi, P. Acheampong, S. Obom, A. Hennemuth, Y. Lu, W. Bai, W. Shi, Y. Gao, H.-O. Peitgen, P. Radau, R. Razavi, A. Tannenbaum, D. Rueckert, J. Cates, T. Schaeffter, D. Peters, R. MacLeod, and K. Rhode, "Evaluation of current algorithms for segmentation of scar tissue from late gadolinium enhancement cardiovascular magnetic resonance of the left atrium: an open-access grand challenge," *J. Cardiovasc. Magn. Reson.*, vol. 15, pp. 105–122, 2013.
40. J. L. Harrison, C. Sohns, N. W. Linton, R. Karim, S. E. Williams, K. S. Rhode, J. Gill, M. Cooklin, C. A. Rinaldi, M. Wright, T. Schaeffter, R. S. Razavi, and M. D. O'Neill, "Repeat left atrial catheter ablation: cardiac magnetic resonance prediction of endocardial voltage and gaps in ablation lesion sets," *Circ. Arrhythmia Electrophysiol.*, vol. 8, no. 2, pp. 270–278, Apr. 2015.

41. B. R. Knowles, D. Caulfield, M. Cooklin, C. A. Rinaldi, J. Gill, J. Bostock, R. Razavi, T. Schaeffter, and K. S. Rhode, "3-D visualization of acute RF ablation lesions using MRI for the simultaneous determination of the patterns of necrosis and edema," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1467–1475, 2010.
42. Q. Tao, E. G. Ipek, R. Shahzad, F. F. Berendsen, S. Nazarian, and R. J. van der Geest, "Fully automatic segmentation of left atrium and pulmonary veins in late gadolinium-enhanced MRI: towards objective atrial scar assessment," *J. Magn. Reson. Imaging*, vol. 44, no. 2, pp. 346–354, Jan. 2016.
43. R. Karim, A. Arujuna, R. J. Housden, J. Gill, H. Cliffe, K. Matharu, C. A. Rinaldi, M. O'Neill, D. Rueckert, R. Razavi, T. Schaeffter, and K. Rhode, "A method to standardize quantification of left atrial scar from delayed-enhancement MR Images," *Transl. Eng. Heal. Med. IEEE J.*, vol. 2, pp. 1–15, 2014.
44. D. Perry, A. Morris, N. Burgon, C. McGann, R. MacLeod, and J. Cates, "Automatic classification of scar tissue in late gadolinium enhancement cardiac MRI for the assessment of left-atrial wall injury after radiofrequency ablation," 2012, p. 83151D.
45. G. Veni, S. Y. Elhabian, and R. T. Whitaker, "ShapeCut: Bayesian surface estimation using shape-driven graph," *Med. Image Anal.*, vol. 40, pp. 11–29, Aug. 2017.
46. R. Achanta, A. Shaji, K. Smith, and A. Lucchi, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2281, 2012.
47. G. Yang, X. Zhuang, H. Khan, S. Haldar, E. Nyktari, X. Ye, G. Slabaugh, T. Wong, R. Mohiaddin, J. Keegan, and D. Firmin, "Segmenting atrial fibrosis from late gadolinium-enhanced cardiac MRI by deep-learned features with stacked sparse auto-encoders," in *Annual Conference on Medical Image Understanding and Analysis, Communications in Computer and Information Science*, 2017, pp. 195–206.
48. G. Yang, X. Zhuang, H. Khan, S. Haldar, E. Nyktari, L. Li, R. Wage, X. Ye, G. Slabaugh, R. Mohiaddin, T. Wong, J. Keegan, and D. Firmin, "Fully automatic segmentation and objective assessment of atrial scars for longstanding persistent atrial fibrillation patients using late gadolinium-enhanced MRI," *Med. Phys.*, vol. 45, no. 4, pp. 1562–1576, 2018.
49. G. Yang, X. Zhuang, H. Khan, S. Haldar, E. Nyktari, X. Ye, G. Slabaugh, T. Wong, R. Mohiaddin, J. Keegan, and D. Firmin, "A fully automatic deep learning method for atrial scarring segmentation from late gadolinium-enhanced MRI images," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 844–848.
50. D. C. Peters, J. V. Wylie, T. H. Hauser, R. Nezafat, Y. Han, J. J. Woo, J. Taclas, K. V. Kissinger, B. Goddu, M. E. Josephson, and W. J. Manning, "Recurrence of atrial fibrillation correlates with the extent of post-procedural late gadolinium enhancement. A pilot study," *JACC Cardiovasc. Imaging*, vol. 2, no. 3, pp. 308–316, 2009.
51. J. Keegan, P. Jhooti, S. V. Babu-Narayan, P. Drivas, S. Ernst, and D. N. Firmin, "Improved respiratory efficiency of 3D late gadolinium enhancement imaging using the continuously adaptive windowing strategy (CLAWS)," *Magn. Reson. Med.*, vol. 71, no. 3, pp. 1064–1074, 2014.
52. J. Keegan, P. Drivas, and D. N. Firmin, "Navigator artifact reduction in three-dimensional late gadolinium enhancement imaging of the atria," *Magn. Reson. Med.*, vol. 785, pp. 779–785, 2013.
53. X. Zhuang, W. Bai, J. Song, S. Zhan, X. Qian, W. Shi, Y. Lian, and D. Rueckert, "Multiatlas whole heart segmentation of CT data using conditional entropy for atlas ranking and selection," *Med. Phys.*, vol. 42, no. 7, pp. 3822–3833, 2015.
54. X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI," *Med. Image Anal.*, vol. 31, pp. 77–87, Jul. 2016.
55. G. Yang, X. Zhuang, H. Khan, S. Haldar, E. Nyktari, L. Li, X. Ye, G. Slabaugh, T. Wong, R. Mohiaddin, J. Keegan, and D. Firmin, "Multi-atlas propagation based left atrium segmentation coupled with super-voxel based pulmonary veins delineation in late gadolinium-enhanced cardiac MRI," in *SPIE Medical Imaging*, 2017, pp. 1013313–1.
56. G. Yang, X. Ye, G. Slabaugh, J. Keegan, R. Mohiaddin, and D. Firmin, "Differentiation of pre-ablation and post-ablation late gadolinium-enhanced cardiac MRI scans of longstanding persistent atrial fibrillation patients," in *SPIE Medical Imaging*, 2017, p. 1013400–1.
57. C. Tobon-Gomez, A. Geers, J. Peters, J. Weese, K. Pinto, R. Karim, T. Schaeffter, R. Razavi, and K. Rhode, "Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets," *IEEE Trans. Med. Imaging*, vol. 34, no. 7, pp. 1460–1473, Feb. 2015.
58. X. Zhuang, K. S. Rhode, R. S. Razavi, D. J. Hawkes, and S. Ourselin, "A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI," *IEEE Trans. Med. Imaging*, vol. 29, no. 9, pp. 1612–1625, 2010.
59. X. Zhuang, S. Arridge, D. J. Hawkes, and S. Ourselin, "A nonrigid registration framework using spatially encoded mutual information and free-form deformations," *IEEE Trans. Med. Imaging*, vol. 30, no. 10, pp. 1819–28, Oct. 2011.

60. X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de-Solorzano, "Combination strategies in multi-atlas image segmentation: application to brain MR data," *IEEE Trans. Med. Imaging*, vol. 28, no. 8, pp. 1266–1277, Aug. 2009.
61. W. Bai, W. Shi, C. Ledig, and D. Rueckert, "Multi-atlas segmentation with augmented features for cardiac MR images," *Med. Image Anal.*, vol. 19, no. 1, pp. 98–109, 2015.
62. P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, "Nonlocal patch-based label fusion for hippocampus segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, 2010, pp. 129–136.
63. S. F. Eskildsen, P. Coupé, V. Fonov, J. V. Manjón, K. K. Leung, N. Guizard, S. N. Wassef, L. R. Østergaard, and D. L. Collins, "BEaST: Brain extraction based on nonlocal segmentation technique," *Neuroimage*, vol. 59, no. 3, pp. 2362–2373, Feb. 2012.
64. S. Hu, P. Coupé, J. C. Pruessner, and D. L. Collins, "Nonlocal regularization for active appearance model: application to medial temporal lobe segmentation," *Hum. Brain Mapp.*, vol. 35, no. 2, pp. 377–395, Feb. 2014.
65. M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," *IEEE Trans. Med. Imaging*, vol. 29, no. 10, pp. 1714–29, Oct. 2010.
66. H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 611–623, Mar. 2013.
67. G. Wu, M. Kim, G. Sanroma, Q. Wang, B. C. Munsell, D. Shen, and Alzheimer's Disease Neuroimaging Initiative, "Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition," *Neuroimage*, vol. 106, pp. 34–46, Feb. 2015.
68. T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.
69. T. Lindeberg, "Generalized Gaussian Scale-Space Axiomatics Comprising Linear Scale-Space, Affine Scale-Space and Spatio-Temporal Scale-Space," *J. Math. Imaging Vis.*, vol. 40, no. 1, pp. 36–81, May 2011.
70. M. Holden, L. D. Griffin, N. Saeed, and D. L. G. Hill, "Multi-channel mutual information using scale space," in *Medical Image Computing and Computer-Assisted Intervention*, 2004, pp. 797–804.
71. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
72. S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *2011 International Conference on Computer Vision*, 2011, pp. 2548–2555.
73. Y. Wu, G. Liu, M. Huang, J. Guo, J. Jiang, W. Yang, W. Chen, and Q. Feng, "Prostate segmentation based on variant scale patch and local independent projection," *IEEE Trans. Med. Imaging*, vol. 33, no. 6, pp. 1290–1303, Jun. 2014.
74. P. Thévenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2083–99, 2000.
75. M. Soltaninejad, G. Yang, T. Lambrou, N. Allinson, T. L. Jones, T. R. Barrick, F. A. Howe, and X. Ye, "Automated brain tumour detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI," *Int. J. Comput. Assist. Radiol. Surg.*, Sep. 2016.
76. M. Soltaninejad, X. Ye, G. Yang, N. Allinson, and T. Lambrou, "An image analysis approach to MRI brain tumour grading," *Oncol. News*, vol. 9, no. 6, pp. 204–207, Jan. 2015.
77. M. Soltaninejad, X. Ye, G. Yang, N. Allinson, and T. Lambrou, "Brain tumour grading in different MRI protocols using SVM on statistical features," in *Medical Image Understanding and Analysis*, 2014.
78. G. Yang, F. Raschke, T. R. Barrick, and F. A. Howe, "Manifold learning in MR spectroscopy using nonlinear dimensionality reduction and unsupervised clustering," *Magn. Reson. Med.*, vol. 74, no. 3, pp. 868–878, 2015.
79. X. Ye, G. Beddoe, and G. Slabaugh, "Automatic graph cut segmentation of lesions in CT using mean shift superpixels," *Int. J. Biomed. Imaging*, vol. 2010, p. 983963, Jan. 2010.
80. Y. Zhao, Y. Zheng, Y. Liu, J. Yang, Y. Zhao, D. Chen, and Y. Wang, "Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy," *IEEE Trans. Med. Imaging*, vol. 36, no. 1, pp. 51–63, Jan. 2017.
81. K. Lemola, M. Sneider, B. Desjardins, I. Case, J. Han, E. Good, K. Tamirisa, A. Tsemo, A. Chugh, F. Bogun, F. Pelosi, E. Kazerooni, F. Morady, and H. Oral, "Computed tomographic analysis of the anatomy of the left atrium and the esophagus: implications for left atrial catheter ablation," *Circulation*, vol. 110, no. 24, pp. 3655–60, Dec. 2004.

82. P. A. Lachenbruch and M. Goldstein, "Discriminant analysis," *Biometrics*, vol. 35, no. 1, pp. 69–85, 1979.
83. G. John and P. Langley, "Estimating continuous distributions in Bayesian classifiers," *Proc. Elev. Conf. Uncertain. Artif. Intell.*, pp. 338–345, 1995.
84. J. Friedman, J. Bentley, and R. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. Math. Softw.*, vol. 1549, pp. 209–226, 1977.
85. N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, First. Cambridge University Press, 2000.
86. G. Yang, T. L. Jones, F. A. Howe, and T. R. Barrick, "Morphometric model for discrimination between glioblastoma multiforme and solitary metastasis using three-dimensional shape analysis," *Magn. Reson. Med.*, Jul. 2015.
87. M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*. PWS Pub. Co., 1996.
88. G. Yang, T. L. Jones, T. R. Barrick, and F. A. Howe, "Discrimination between glioblastoma multiforme and solitary metastasis using morphological features derived from the p:q tensor decomposition of diffusion tensor imaging," *NMR Biomed.*, vol. 27, no. 9, pp. 1103–1111, Sep. 2014.
89. Y. Mo, F. Liu, J. Zhang, G. Yang, T. He, and Y. Guo, "Deep poicare map for robust medical image segmentation," Mar. 2017.
90. H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks," in *Annual Conference on Medical Image Understanding and Analysis, Communications in Computer and Information Science*, 2017, pp. 506–517.
91. S. Yu, H. Dong, G. Yang, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, D. Firmin, and Y. Guo, "Deep de-aliasing for fast compressive sensing MRI," May 2017. arXiv:1705.07137
92. D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, Jun. 2017.
93. M. Hasan and A. K. Roy-Chowdhury, "A continuous learning framework for activity recognition using deep hybrid feature models," *IEEE Trans. Multimed.*, vol. 17, no. 11, pp. 1909–1922, Nov. 2015.
94. X. Zhuang, "Challenges and methodologies of fully automatic whole heart segmentation: a review," *J. Healthc. Eng.*, vol. 4, no. 3, pp. 371–408, 2013.
95. G. Yang, J. H. Hipwell, M. J. Clarkson, C. Tanner, T. Mertzaniidou, S. Gunn, S. Ourselin, D. J. Hawkes, and S. R. Arridge, "Combined reconstruction and registration of digital breast tomosynthesis," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6136 LNCS, pp. 760–768, 2010.
96. D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, 1993.
97. G. Yang, F. Raschke, T. R. Barrick, and F. A. Howe, "Classification of brain tumour 1H MR spectra: extracting features by metabolite quantification or nonlinear manifold learning?" in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2014, pp. 1039–1042.
98. B. Efron, "Estimating the error rate of a prediction rule: improvement on cross-validation," *J. Am. Stat. Assoc.*, vol. 78, no. 382, p. 316, 1983.
99. T. L. Jones, T. J. Byrnes, G. Yang, F. A. Howe, B. A. Bell, and T. R. Barrick, "Brain tumor classification using the diffusion tensor image segmentation (D-SEG) technique," *Neuro. Oncol.*, vol. 17, no. 3, pp. 466–476, Aug. 2014.
100. K. Murphy, *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
101. M. H. Moghari, D. C. Peters, J. Smink, L. Goepfert, K. V. Kissinger, B. Goddu, T. H. Hauser, M. E. Josephson, W. J. Manning, and R. Nezafat, "Pulmonary vein inflow artifact reduction for free-breathing left atrium late gadolinium enhancement," *Magn. Reson. Med.*, vol. 66, no. 1, pp. 180–6, Jul. 2011.
102. B. Irving, J. M. Franklin, B. W. Papież, E. M. Anderson, R. A. Sharma, F. V. Gleeson, S. M. Brady, and J. A. Schnabel, "Pieces-of-parts for supervoxel segmentation with global context: application to DCE-MRI tumour delineation," *Med. Image Anal.*, vol. 32, pp. 69–83, Aug. 2016.
103. F. Kanavati, T. Tong, K. Misawa, M. Fujiwara, K. Mori, D. Rueckert, and B. Glocker, "Supervoxel classification forests for estimating pairwise image correspondences," *Pattern Recognit.*, vol. 63, pp. 561–569, Mar. 2017.
104. M. P. Heinrich, I. J. A. Simpson, B. W. Papież, S. M. Brady, and J. A. Schnabel, "Deformable image registration by combining uncertainty estimates from supervoxel belief propagation," *Med. Image Anal.*, vol. 27, pp. 57–71, Jan. 2016.
105. Z. Tian, L. Liu, Z. Zhang, and B. Fei, "Superpixel-based segmentation for 3D prostate MR images," *IEEE Trans. Med. Imaging*, vol. 35, no. 3, pp. 791–801, 2016.