Medical Image Analysis on Left Atrial LGE MRI for Atrial Fibrillation Studies: A Review

Lei Li[a,b,c], Veronika A. Zimmer[c,d], Julia A. Schnabel[c,d,e], Xiahai Zhuang*[a]

[a]*School of Data Science, Fudan University, Shanghai, China*
[b]*School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China*
[c]*School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK*
[d]*Department of Informatics, Technical University of Munich, Germany*
[e]*Helmholtz Center Munich, Germany*

ARTICLE INFO

ABSTRACT

Late gadolinium enhancement magnetic resonance imaging (LGE MRI) is commonly used to visualize and quantify left atrial (LA) scars. The position and extent of LA scars provide important information on the pathophysiology and progression of atrial fibrillation (AF). Hence, LA LGE MRI computing and analysis are essential for computer-assisted diagnosis and treatment stratification of AF patients. Since manual delineations can be time-consuming and subject to intra- and inter-expert variability, automating this computing is highly desired, which nevertheless is still challenging and under-researched.

This paper aims to provide a systematic review on computing methods for LA cavity, wall, scar, and ablation gap segmentation and quantification from LGE MRI, and the related literature for AF studies. Specifically, we first summarize AF-related imaging techniques, particularly LGE MRI. Then, we review the methodologies of the four computing tasks in detail and summarize the validation strategies applied in each task as well as state-of-the-art results on public datasets. Finally, the possible future developments are outlined, with a brief survey on the potential clinical applications of the aforementioned methods. The review indicates that the research into this topic is still in the early stages. Although several methods have been proposed, especially for the LA cavity segmentation, there is still a large scope for further algorithmic developments due to performance issues related to the high variability of enhancement appearance and differences in image acquisition.

1. Introduction

1.1. Clinical goals

Atrial fibrillation (AF) is the most common cardiac arrhythmia encountered in the clinic, occurring in up to 2% of the population and rising in prevalence along with advancing age (Chugh et al., 2014). Fig. 1 presents a comparison of sinus rhythm and AF. One can see that there are chaotic electrical signals in the atrium of AF patients compared to sinus rhythm, resulting in a rapid and irregular heart rhythm. Radiofrequency catheter ablation via pulmonary vein isolation (PVI) is a promising procedure for treating AF, especially for paroxysmal AF patients (Calkins et al., 2007). The left atrium (LA) is a crucial structure in the pathophysiology of AF, and the observation of LA remodeling can be important for the initial evaluation of AF (Tops et al., 2010). Besides, structural changes in the LA wall (especially changes in the wall thickness) are known to

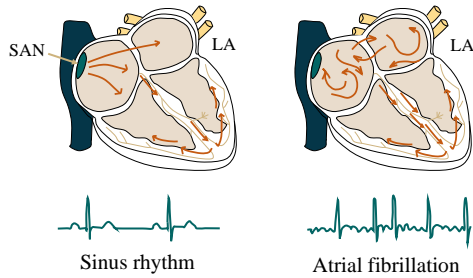*URL:* zxh@fudan.edu.cn (Xiahai Zhuang*)

Fig. 1. The electrical activities of the left atrium (LA) in sinus rhythm and atrial fibrillation (AF), respectively. The sinoatrial node (SAN) produces an electrical impulse, which is regular in the sinus rhythm and can be overwhelmed by disorganized electrical waves, usually originating from the pulmonary veins.

Table 1. AF patient classification that depends on fibrosis extent (Akoum et al., 2011).

Utah grade	Percentage	Success rate	AF recurrence
Utah 1 (minimal)	$\leq 5\%$	100%	0
Utah 2 (mild)	5~20%	81.8%	28%
Utah 3 (moderate)	20~35%	62.5%	35%
Utah 4 (extensive)	$\geq 35\%$	0	56%

occur in AF patients (Karim et al., 2018). The wall thickness can be used to predict the response to invasive treatment of AF and has the potential for improving the safety of AF ablation (Whitaker et al., 2016). The wall thickness is also important to measure the transmurality of scars which is related to the AF recurrence (Ranjan et al., 2011). The success of AF treatments is highly related to the formation of a contiguous scar completely encircling the veins (Ranjan et al., 2011). Unfortunately, the encircling lesion is often incomplete with a combination of ablation scars and gaps of healthy tissue (Miller et al., 2012). Therefore, the extent and distribution of both scars and gaps are important information for AF patient selection (Akoum et al., 2011), diagnosis prediction (Arujuna et al., 2012), and treatment stratification (Njoku et al., 2018). For example, patients were divided into four grades according to their degrees of fibrosis (refers to preexisting scars) in Akoum et al. (2011), shown in Table 1. Based on the scoring, various therapeutic strategies were suggested by electrophysiologists.

Recently, late gadolinium enhancement magnetic resonance imaging (LGE MRI) has evolved as a tool for defining the extent of fibrosis/ scars and visualizing the ablation gaps (Siebermair et al., 2017; Li et al., 2020b; Nuñez-Garcia et al., 2019). *Therefore, it is crucial to develop techniques for the four progressive tasks, i.e., (1) LA cavity segmentation, (2) LA wall segmentation together with wall thickness measurement, (3) scar segmentation and quantification, and (4) ablation gap localization from LGE MRI.* Fig. 2 provides the clinical pipeline for AF ablation procedures, where the role of LGE MRI is highlighted and the four closely related tasks of clinical interests are presented, followed by several related clinical applications.

1.2. Challenges of LA LGE MRI computing

Manual delineations of the LA, LA wall, scars, and ablation gaps are all labor-intensive and prone to be subjective, so their

Table 2. Search engines and expressions used to identify potential papers for review.

Engine	Google scholar, PubMed, IEEE-Xplore, and Citeseer
Term	"Atrial fibrillation" or AF **and**
	"Late gadolinium/ delayed enhancement/ contrast enhanced (cardiac) magnetic resonance" or "LGE/ DE/ CE MR(I)" or "LGE-/ DE-/ CE-MR(I)" or "LGE/ DE/ CE CMR" or "LGE-/ DE-/ CE-CMR" **and**
	Classif*/ segment*/ quantif*/ localiz*/ detect* **and**
	"Left atrium/ atrial" or LA **or**
	"Atrial wall/ myocardi*" or "wall thickness" **or**
	"Atrial scars/ fibrosis/ lesion" or "ablation pattern" **or**
	"Ablat*"/ lesion gaps" or "gaps in ablation lesion" or "incomplete ablation pattern"

automation is highly desired, which nevertheless remains challenging. The challenges for automatic LA cavity segmentation are mainly from the large variations in terms of LA shape, intensity range as well as poor image quality. For the LA wall analysis, two additional difficulties are presented, i.e., the intrinsic thin wall thickness and the complex structure of the LA wall. Here, the complex structure refers to the multiple openings in its 3D structure such as the pulmonary veins (PV) and mitral valve (MV) of the LA. For the scar analysis, its unique challenge lies in the enhanced noise from surrounding tissues. For the gap quantification, the large variability in PV morphology (position, orientation, size, thickness) and the robustness to scar segmentation changes are the two major concerns. Fig. 3 illustrates and explains part of these challenges in an intuitive way.

1.3. Study inclusion and literature search

In this work, we aim to provide the reader with a survey of the state-of-the-art image computing techniques, important results as well as the related literature for AF studies. To ensure comprehensive coverage, we have screened publications from the last 10 years related to this topic. Our main sources of references were Internet searches using engines such as Google Scholar, PubMed, IEEE-Xplore, and Citeseer. To cover as many related works as possible, flexible search terms have been employed when using these search engines, as summarized in Table 2. Both peer-reviewed journal papers and conference papers were included here. We have also followed the references found in papers from these sites, and finally collected a comprehensive library of more than 130 papers. Fig. 4 presents the distributions of papers in segmentation and quantification from LGE MRI for AF patients per year/task. Note that we generally picked the most detailed and representative ones for this review when we encountered several papers from the same authors about the same subject.

1.4. Related review literature

Table 3 lists existing review papers related to AF. One can see that most current AF-related review papers focused on a clinical survey instead of the methodology of image computing, such as segmentation or quantification algorithms. Only two reviews, Pontecorboli et al. (2017) and Jamart et al. (2020),
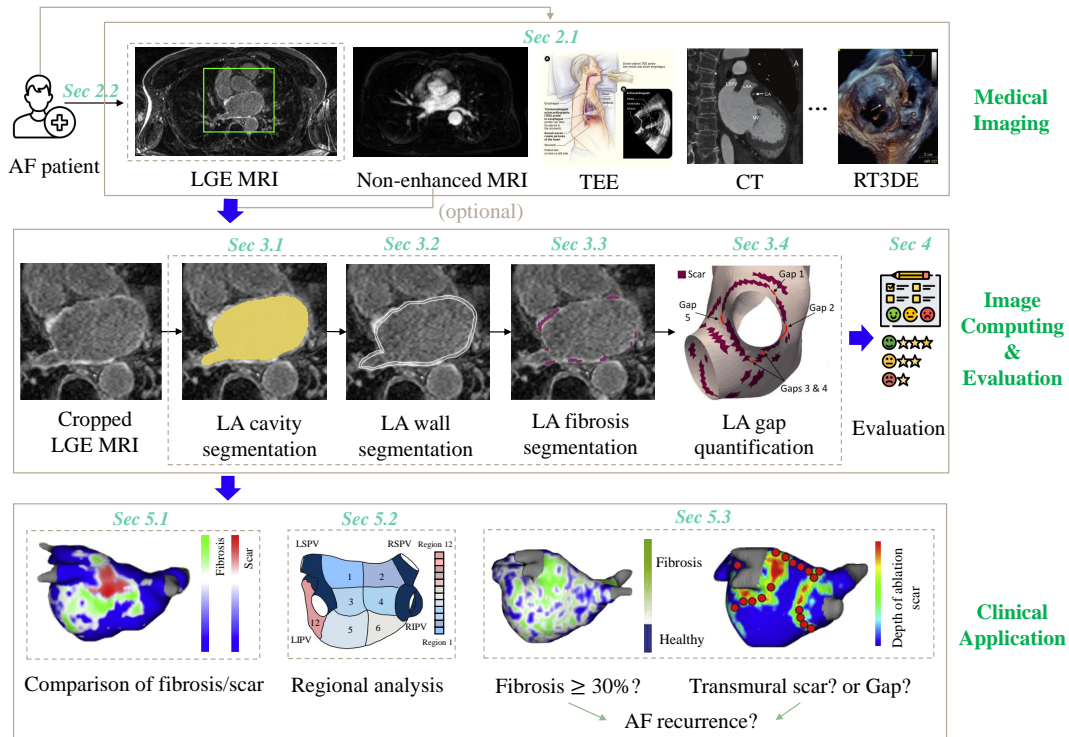
Fig. 2. The pipeline of LA image computing for AF studies and the structure of this review. Top row: common image modalities for AF treatments, such as late gadolinium enhancement magnetic resonance imaging (LGE MRI), non-enhanced MRI, transesophageal echocardiography (TEE) (image adapted from *Stanford HEALTH CARE*), CT, and real-time 3D echocardiography (RT3DE) (images adapted from Regazzoli et al. (2015) with permission); Middle row: computation and evaluation steps for LA analysis reviewed in this study (images adapted from Li et al. (2021c); Nuñez-Garcia et al. (2019) with permission); Bottom row: possible clinical applications (images adapted from Siebermair et al. (2017) with permission).

are similar to ours in terms of the topic (LGE MRI) and style (technical). However, only conventional thresholding methods or only deep learning (DL)-based methods were reviewed in each work. Fig. 5 visualizes the scopes of current reviews as well as this review, and one can see that the scopes are different although partial overlaps can be found. Besides, our review organizes the related works according to the clinical pipeline (see Fig. 2), resulting in an intuitive structure of the paper.

1.5. Structure of this review

The remainder of the paper is organized as follows (compare Fig. 2): Section 2 presents the current common imaging tools used in AF ablation and the importance of LGE MRI in the management of AF. Section 3 systematically reviews the state-of-the-art image computing techniques and results of LA cavity, wall, scar, and ablation gap segmentation and quantification. Section 4 presents the public data, evaluation measures, and state-of-the-art evaluation results on the public data for each task. Potential clinical applications are provided in Section 5. Discussion of current LA LGE MRI computing challenges and future perspectives are given in Section 6, along with a conclusion in Section 7.

2. Imaging of atrial fibrillation

Medical images can offer crucial information for the evaluation and treatment of AF patients, and have been widely used

in the ablation process (Tops et al., 2010; Obeng-Gyimah and Nazarian, 2020). Table 4 summarizes the common imaging modalities used in three ablation stages (before, during, and after catheter ablation), mainly referring to Tops et al. (2010) and Obeng-Gyimah and Nazarian (2020). One can see that diverse imaging modalities have been introduced in the ablation process, each of which assists in various aspects of the procedure.

2.1. Imaging for ablation procedures

Before catheter ablation (CA), the first step is to exclude contraindication, such as the LA appendage (LAA) thrombi which are normally detected using transesophageal echocardiography (TEE) (Ellis et al., 2006; Calkins et al., 2007; Pathan et al., 2018). MRI and computed tomography (CT) can be used to detect LA thrombi, but both tend to have a low inter-observer agreement (Mohrs et al., 2006; Gottlieb et al., 2008). In addition, the images are statically acquired a few seconds after the arrival of contrast to the LAA. Hence, it could be difficult to differentiate LAA thrombi from sluggish flow (Romero et al., 2013). To select patients expected for successful CA, the assessment of LA, PVs, and fibrosis are the key steps (Berruezo et al., 2007; Akoum et al., 2011). Three-dimensional (3D) imaging techniques, such as CT and MRI, are generally used for PV anatomy assessment. PV anatomy can also be measured by TEE, achieving up to 95% concordance with MRI (Toffanin et al., 2006). Moreover, cardiac MRI remains the gold standard for fibrosis assessment (Obeng-Gyimah and Nazarian, 2020). Especially, LGE MRI appears to be a promising alternative for
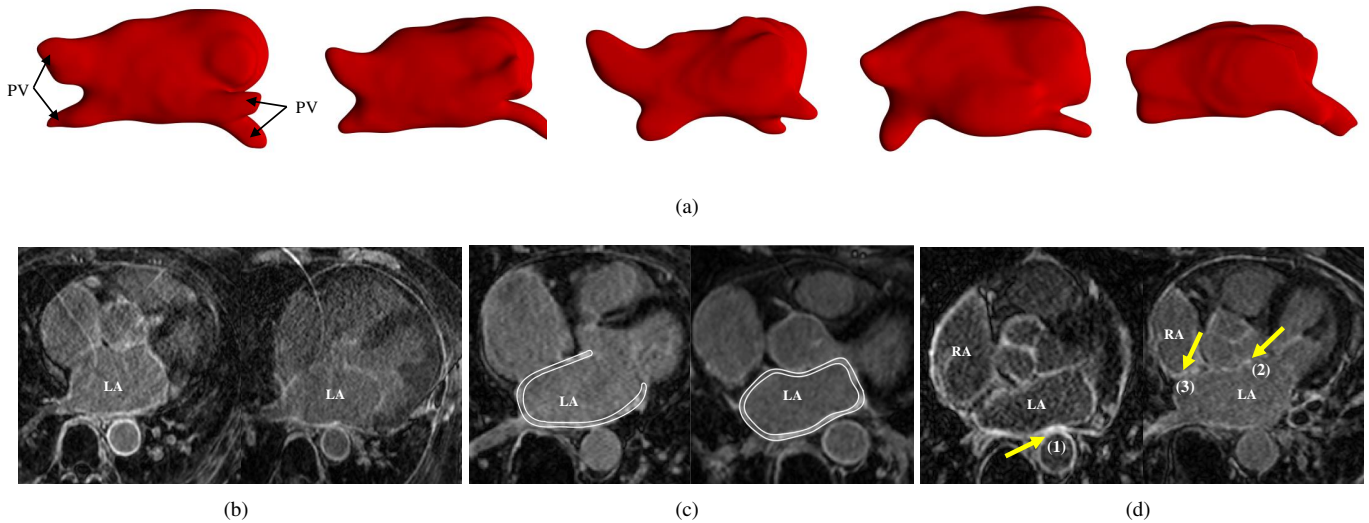
Fig. 3. The challenges of automatic segmentation and quantification of LGE MRI for AF: (a) various LA and pulmonary vein (PV) shapes; (b) two typical LGE MRIs with poor quality; (c) thin atrial walls highlighted using bright white color in the figure; (d) surrounding enhanced regions pointed out by the arrows, where (1) and (2) indicate the enhanced walls of descending and ascending aorta, respectively; and (3) denotes the enhanced walls of the right atrium (RA). Images (b)-(d) adapted from Li et al. (2020b) with permission.

Table 3. Summary of the review papers related to AF. EAM: electroanatomical mapping; JAF: Journal of Atrial Fibrillation; JACC: Journal of the American College of Cardiology; RMPBM: Magnetic Resonance Materials in Physics, Biology and Medicine; JICRM: The Journal of Innovations in Cardiac Rhythm Management; FCM: Frontiers in Cardiovascular Medicine; CET: Cardiovascular Engineering and Technology; DL: deep learning; CT: computed tomography.

Source	Venue	Scope	Limitation
Cox (2003)	Europace	Surgical treatment of AF	Clinical review
Rolf et al. (2014)	JAF	EAM of AF	Clinical review
Dzeshka et al. (2015)	JACC	Mechanisms and clinical implications of AF	Clinical review
Whitaker et al. (2016)	Europace	Wall thickness measurement for CT	Target image is not LGE MRI
Peng et al. (2016)	RMPBM	Cardiac chamber segmentation	Target partially includes LA cavity
Pontecorboli et al. (2017)	Europace	Fibrosis segmentation from LGE MRI	Only thresholding methods are included
Siebermair et al. (2017)	JACC	LGE fibrosis imaging	Clinical review
Obeng-Gyimah and Nazarian (2020)	JICRM	Imaging for AF ablation	Clinical review
Jamart et al. (2020)	FCM	LA cavity segmentation from LGE MRI	Only DL-based methods are included
Chen et al. (2020)	FCM	DL-based cardiac segmentation	Target partially includes LA and its scars
Habijan et al. (2020)	CET	Whole heart and chamber segmentation	Target partially includes LA cavity

pre-ablation scar visualization and quantification (Siebermair et al., 2017).

During CA, fluoroscopy is the most commonly employed imaging technique in the electrophysiology laboratory. Intra-ardiac echocardiography (ICE) offers real-time imaging of the PVs and adjacent structures and enhances the safety of transseptal puncture by visualizing inter-atrial septum and puncture needle (Jongbloed et al., 2005a). Both ICE and fluoroscopy can visualize the LA and PVs (Saad et al., 2002). Note that the integration of different imaging modalities during CA is promising (Tops et al., 2010), but is out of the scope of this review.

After CA and during the follow-up study, the main target of post-procedural imaging is to monitor complications and help predict recurrence. The most frequently occurring complications of AF ablation include PV stenosis, pericardial effusion, and atrio-oesophageal fistul. Multi-slice CT and MRI are usually used for accurate assessment of PV stenosis and esophageal injury (Holmes et al., 2009). Transthoracic echocardiography (TTE) is a recommended imaging tool for screening

to detect pericardial effusion (Calkins et al., 2007). To predict recurrence, LA size and functions are important indices, as LA ablation can lead to the formation of scars and subsequent changes in LA anatomy (Casaclang-Verzosa et al., 2008). *For the follow-up analysis of LA volumes*, TTE is typically used, but 3D techniques, such as real-time 3D echocardiography (Zhang et al., 2017), multi-slice CT (Polaczek et al., 2019), and MRI (Tsao et al., 2005), especially LGE MRI (McGann et al., 2014), may provide more accurate information. *For the measurement of LA wall thickness*, TEE has the advantages of high temporal resolution and short acquisition time, but it is difficult to obtain descriptive information on the LA wall due to its low spatial resolution (Nakamura et al., 2011). CT is an ideal modality, thanks to its high resolution, and MRI is widely considered to be the gold standard for the viability assessment of wall pathology (Karim et al., 2018).

LGE MRI has been recently widely explored for scar and ablation gap quantification (Nuñez-Garcia et al., 2019; Mishima et al., 2019). Note that T1 mapping MRI could be used to obtain

Table 4. The role of different imaging modalities in AF ablation procedures. CA: catheter ablation; LAA: LA appendage; ICE: intracardiac echocardiography; TTE: transthoracic echocardiography; STE: speckle tracking echocardiography.

Stage	Target	Imaging modality	Important summary
Before CA	Assessment of LAA thrombus	TEE	Clinical reference for LAA thrombi identification (Calkins et al., 2007)
		CT/ MRI	Low inter-observer agreement (Mohrs et al., 2006; Gottlieb et al., 2008)
	Assessment of LA size and anatomy	TTE	The most commonly used imaging technique in daily clinical practice (Tops et al., 2007)
		RT3DE/STE	New techniques for the assessment of LA volumes (Cameli et al., 2012)
		MRI	Gold standard for the assessment of LA volumes (Kuchynka et al., 2015)
	Assessment of PV anatomy	CT/ MRI	Provides detailed 3D information on PV anatomy as a "road-map" for ablation (Bhagirath et al., 2014)
	Assessment of fibrosis	LGE MRI	The most widely used MRI protocol for LA fibrosis imaging (Siebermair et al., 2017)
During CA	Positioning catheters	Fluoroscopy	Standard imaging modality in the electrophysiology laboratory; used to visualize catheters and devices (Bourier et al., 2016)
	Transseptal puncture	ICE	Used to enhance the safety of transseptal puncture and catheter tissue contact; used to visualize inter-atrial septum and puncture needle (Jongbloed et al., 2005a)
	Visualization of LA and PVs	Fluoroscopy	New rotational angiography technique to accurately identify PV anatomy and diameters (Thiagalingam et al., 2008)
		ICE	Real-time assessment of PV ostium with a limitation on the detection of small proximal branches from PVs (Saad et al., 2002; Wood et al., 2004; Jongbloed et al., 2005b)
After CA	Assessment of PV stenosis	CT/ MRI	Preferably, these 3D techniques are correlated with pre-procedural images for detection of PV stenosis (Holmes et al., 2009)
	Detection of pericardial	TTE	Routine echocardiography should be performed before discharge and during the follow-up study (Calkins et al., 2007)
	Esophageal injury	CT/ MRI	Performed when atrio-oesophageal fistula is suspected (Calkins et al., 2007)
	Assessment of LA size and function	TTE	Conventional method for the detection of LA volumes and function (Blondheim et al., 2005)
		RT3DE/CT/(LGE) MRI	3D assessment of LA volumes allows the detection of LA reverse remodelling (Zhang et al., 2017; Polaczek et al., 2019; Tsao et al., 2005; McGann et al., 2014)
	Assessment of wall thickness	TEE/CT/(LGE) MRI	Increased atrial wall thickening was seen in the post-ablation scans (Nakamura et al., 2011; Karim et al., 2018; Habibi et al., 2015)
	Assessment of scars and gaps	LGE MRI	Promising in the ablation lesion visualization (McGann et al., 2008)
		T1 mapping MRI	New technique without contrast agent for the assessment of scars (Beinart et al., 2013)

valuable imaging-based biomarkers for diffused cardiac fibrosis, which has been validated against histological studies (Sibley et al., 2012). For example, it is possible with T1 mapping to non-invasively quantify myocardium extracellular volume fraction, which is a biomarker of diffuse reactive fibrosis (Taylor et al., 2016). Nevertheless, it can be difficult to localize fibrosis using T1 mapping MRI, and it is therefore not appropriate for ablation procedure guidance or ablation gap identification. LGE MRI remains a promising method to detect focal and cohesive fibrosis (Pontecorboli et al., 2017).

2.2. LGE MRI for AF studies

LGE MRI is mainly used to evaluate fibrosis and scars of AF patients before and after ablation. This is because LGE MRI can discriminate scarring and healthy tissues by their altered wash-in and wash-out contrast agent kinetics (Marrouche et al., 2014). Scars are thus visualized as the regions of being enhanced or high signal intensity compared to healthy tissues (Yang et al., 2018a). There is still no consensus on the

option and dosage of the contrast agent, nor on the timing of image acquisition after contrast administration, as Table 5 shows. Among the listed protocols, the DECAAF (Delayed-Enhancement MRI Determinant of Successful Radiofrequency Catheter Ablation of Atrial Fibrillation) protocol can be considered the most widely used one for LA fibrosis imaging (Siebermair et al., 2017). Considering the importance and advances of LGE MRI in AF studies, in this review we mainly focus on the computing works on LGE MRI.

3. Image computing

We structure the review of image computing methodology according to the segmentation and quantification tasks in question, as presented in Fig. 2. To understand the key elements of methodologies, we further classify the methods applied in each task (see Fig. 6). In the following sections, we will elaborate and discuss these methods and the corresponding results of different tasks in detail.

Table 5. Imaging parameters for the LGE scar assessment utilized in several leading centers worldwide. SA: Siemens Avanto; PA: Philips Achieva; TR: repetition time; TE: echo time; Acq. T: acquired time after contrast agent injection; CARMA: Comprehensive Arrhythmia Research and Management; DECAAF: Delayed-Enhancement MRI Determinant of Successful Radiofrequency Catheter Ablation of Atrial Fibrillation. Here, † refers to multiple centers.

| Source	Center	Scanner	TR/TE (ms)	Acq. T (min)	Gadolinium dose	Spacing (mm^3)
Badger et al. (2010)	Utah, USA	1.5 T, SA	5.5/2.3	15	0.1 mmol/kg	$1.25 \times 1.25 \times 2.5$
Taclas et al. (2010)	Boston, USA	1.5 T, PA	3.8/1.52	15~25	0.2 mmol/kg	$1.3 \times 1.3 \times 4.0/5.0$
Hunter et al. (2013)	Imperial/Barts, UK	1.5 T, PA	N/A	20	0.4 mmol/kg	$1.5 \times 1.5 \times 4.0$
Bisbal et al. (2014)	Barcelona, Spain	3 T	2.3/1.4	25~30	0.2 mmol/kg	$1.25 \times 1.25 \times 2.5$
McGann et al. (2014)	CARMA†	1.5 T; 3 T, SA	5.2/2.4; 3.1/1.4	5~9; 6~12	0.1 mmol/kg	$1.25 \times 1.25 \times 2.5$
Fukumoto et al. (2015)	John Hopkins, USA	1.5 T, SA	3.8/1.52	10~32	0.2 mmol/kg	$1.3 \times 1.3 \times 2.0$
Harrison et al. (2015a)	KCL, UK	1.5 T, PA	6.2/3.0	20	0.2 mmol/kg	$1.3 \times 1.3 \times 4.0$
Akoum et al. (2015)	DECAAF†	1.5 T; 3 T	5.2/2.4; 3.1/1.4	15	0.1~0.2 mmol/kg	$1.25 \times 1.25 \times 2.5$
Cochet et al. (2015)	Bordeaux, France	1.5 T, SA	6.1/2.4	15~30	0.2 mmol/kg	$1.25 \times 1.25 \times 2.5$
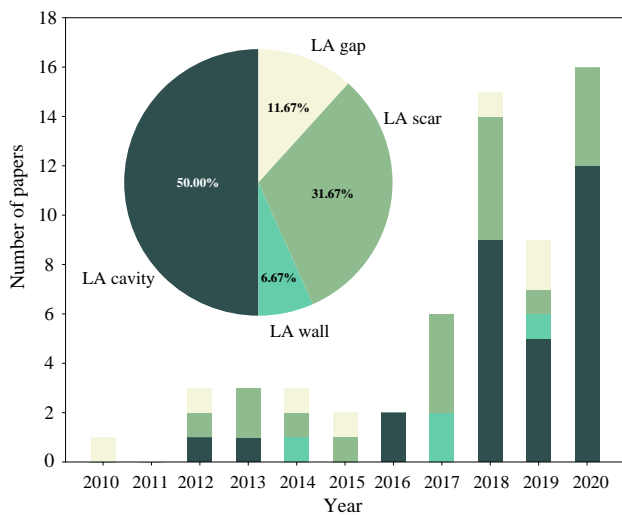
Fig. 4. The distributions of papers of LGE MRI segmentation and quantification for AF patients per year and task.

Fig. 5. Comparison of the scopes of related review studies via Venn diagram.

3.1. LA cavity segmentation

In recent years, many algorithms have been proposed to perform automatic LA cavity segmentation from medical images, but mostly for non-enhanced imaging modalities. Conversely, a limited number of works for the LA cavity segmentation from LGE MRI were reported in the literature before 2018. Most of the current studies on the LA cavity segmentation from LGE MRI are still based on time-consuming and error-prone manual segmentation methods (Higuchi et al., 2018; Njoku et al., 2018). This is mainly because LA cavity segmentation methods in non-enhanced imaging modalities are difficult to directly apply to LGE MRI, due to the existence of contrast agents and low-contrasted boundaries. Existing conventional automatic LA LGE MRI segmentation approaches generally require additional information, such as shape priors (Zhu et al., 2013) or other images, such as non-enhanced 3D MRI (Li et al., 2020b) and contrast enhanced magnetic resonance angiogram (MRA) (Ravanelli et al., 2014; Tao et al., 2016a; Roney et al., 2020). Recently, with the development of DL in medical image processing, numerous DL-based algorithms are proposed for the automatic LA cavity segmentation directly from LGE
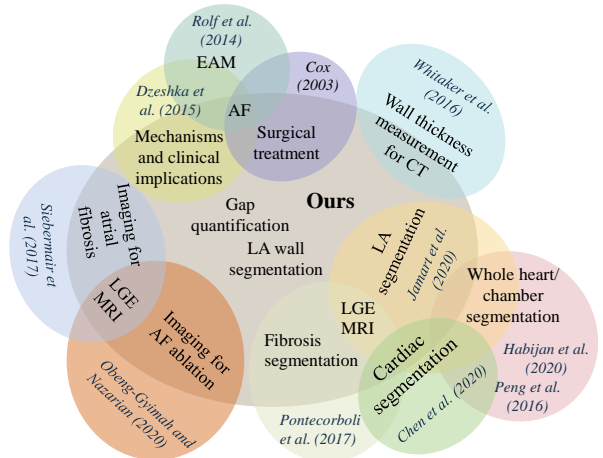
MRI (Xiong et al., 2020). Table 6 summarizes the representative methods and their results in chronological order. The upper and lower parts of the table summarize conventional (non-DL-based) and DL-based methods, respectively.

3.1.1. Conventional methods for LA cavity segmentation

Conventional methods for LA cavity segmentation can be classified into four kinds, i.e., shape models, clustering algorithms, deformable models (region growing, activate contour, and level-set), and atlas-based methods.

Shape models/ clustering algorithms. Many works incorporated anatomical or shape priors to improve the robustness against the large variability of LA shapes and intensity distributions. For example, Gao et al. (2010) used shape learning and region-based active contour evolution for the LA cavity segmentation. The shape learning aimed to utilize prior shape knowledge, to solve the unclear boundary problem in LGE MRI when using the active contour method. Zhu et al. (2013) achieved the LA cavity segmentation using a variational region growing with a moments-based shape prior. They adjusted the weights between the data-driven term and shape prior
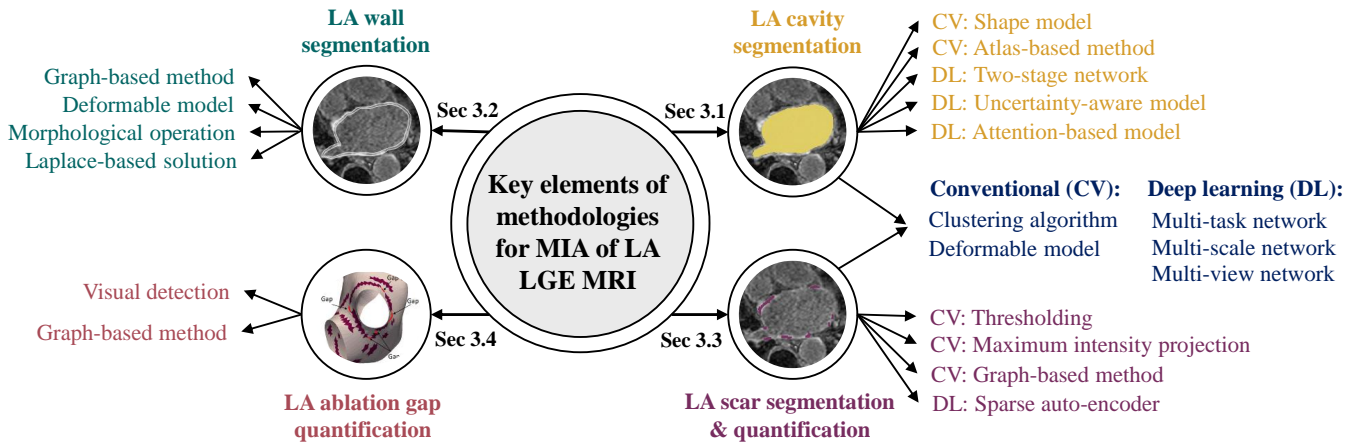
Fig. 6. Key elements of LA LGE MRI computing methodologies on the four tasks. MIA: medical image analysis.

constraint to adapt for the changes in the volume of the target region. Nuñez-Garcia et al. (2018) constructed LGE MRI atlases via multi-atlas segmentation (MAS) and then clustered the LA shapes using principal component analysis to perform a second MAS for the LA cavity segmentation, as presented in Fig. 7. It remains too complicated so far to cover the large shape variation between LA cavities of different subjects by simply imposing a shape prior.

Deformable models. The major challenge of deformable models on the LA cavity segmentation arises from the wide variability of the intensity distribution in LGE MRI. To solve this, Zhu et al. (2013) designed a variational region growing method to reduce its sensitivity to the change of intensity distribution. The seed search in their work was performed by incorporating certain geometric information of PVs relative to the LA. Instead of performing global optimization, Tao et al. (2016a) and Qiao et al. (2018) employed level-set for local refinement on the global segmentation obtained by MAS. The advantage of deformable models is that they do not have a prior assumption about the object geometry and are therefore skillful at capturing local shape variations, such as the PV regions of the LA. Therefore, it is effective to combine deformable models for local attention with other models considering the global shape information of LA. Examples include Gao et al. (2010) and Zhu et al. (2013) where a shape prior was employed as a global constraint.

Atlas-based methods. An alternative way is to use atlas-based methods that can be robust to the LA cavity with high anatomical variations. For instance, Tao et al. (2016a) and Li et al. (2020b) utilized atlas-based methods employing the label of another image (from the same patient) with better anatomical information to assist the LA cavity segmentation of LGE MRI. Tao et al. (2016a) employed MAS to segment the LA cavity from the MRA, and then mapped the generated label to LGE MRI followed by a level-set based refinement. They compared the results with that of solely using LGE MRI (directly employing MAS on LGE MRI) and found that the former achieved better results. They also tested their method on the public dataset

from the *Atrial Segmentation Challenge* where only LGE MRI was provided (Qiao et al., 2018), and achieved better performance in terms of Dice compared to that in Tao et al. (2016a) (0.88 ± 0.03 vs. 0.86 ± 0.05). This may be due to the difference in the dataset, as the public data includes both pre- and post-ablation images. Similarly, Li et al. (2020b) employed an auxiliary MRI sequence to assist the LA cavity segmentation of LGE MRI using MAS methods and obtained a better Dice score (0.898 ± 0.044) than other conventional methods. Particularly, Li et al. (2020b) and Nuñez-Garcia et al. (2018) adopted a multi-atlas based whole heart segmentation (MA-WHS) and then extracted the LA sub-structure. This is because the LGE MRIs employed in their studies cover the whole heart, and MA-WHS could be helpful to exclude surrounding sub-structures of LA. Although in clinical routine LGE MRI may have limited field-of-view, all current public LA LGE MRI datasets were specifically acquired to cover the whole heart with the development of novel whole-heart high-resolution LGE techniques (Toupin et al., 2021). Although auxiliary images can provide better anatomical information, the anatomy extracted from them may be highly deformed compared to that acquired from LGE MRI. It may cause difficulties in the co-registration step and lead to subsequent incorrect segmentation of the LA cavity. Moreover, conventional atlas-based methods are generally time-consuming due to multiple image registration steps.

3.1.2. Deep learning-based methods for LA cavity segmentation

For the LA cavity segmentation, many basic neural network architectures have been employed. To boost the feature learning ability of networks, a series of works have focused on optimizing network structures, investigating different loss functions, and applying anatomical constraints. Here, we mainly classify these DL-based methods according to the network architectures, and will also discuss the loss functions and anatomical constraints used to train the networks.

Architecture of network. Recently, many methods based on different network structures were developed with the launch of the *Atrial Segmentation Challenge* in MICCAI 2018, where U-Net

Table 6. Summary of previously published works on the automatic *LA cavity segmentation* from LGE MRI. CNN: convolutional neural network; MAS: multi-atlas segmentation; MA-WHS: multi-atlas whole heart segmentation; SVM: support vector machine; KNN: K nearest neighbor; FCN: fully convolutional network; HAANet: hierarchical aggregation network; ASD: average surface distance; 95HD: 95% Hausdorff distance; VO: volume overlap; Jc: Jaccard; Acc: Accuracy; Sen: Sensitivity; Spe: Specificity; Cf: Conform; APD: average perpendicular distance; Diam. Err: antero-posterior diameter error; Volume Err: volume error. Note that the results in studies labeled via ‡ are from the benchmark paper (Xiong et al., 2020) for a fair comparison. The reported values in their manuscript may be inconsistent with the results reported in Xiong et al. (2020) as they may employ parts of training data from the challenge as test data for evaluation.

Study	Num	Pre/ Post	Algorithm	Evaluation	$Dice_{LA}$
Gao et al. (2010)	20	Post + Pre	Active contours + statistical shape learning	N/A	N/A
Kutra et al. (2012)	59	Pre	Multi-model based fitting + SVM	Acc	N/A
Zhu et al. (2013)	64	Post + Pre	Variational region growing + shape prior	Dice, VO, 95HD, ASD	0.79 ± 0.05
Deng and Zhang (2016)	64	Post + Pre	KNN + super pixel voting	Dice, VO	0.81 ± 0.08
Tao et al. (2016a)	56	Pre	MAS + 3D level-set	Dice, ASD	0.86 ± 0.05
Nuñez-Garcia et al. (2018)‡	154	Post + Pre	MA-WHS + shape clustering	Dice, HD, ASD	0.859 ± 0.061
Qiao et al. (2018)‡	154	Post + Pre	MAS + level-set	Dice, APD	0.861 ± 0.036
Li et al. (2020b)	58	Post	MA-WHS	Dice	0.898 ± 0.044
Chen et al. (2018b)	100	Post + Pre	Multi-view two-task network	Dice, Acc, Spe, Sen	0.908 ± 0.031
Xiong et al. (2018)	154	Post + Pre	Dual CNNs	Dice, HD, Spe, Sen	0.942 ± 0.014
Chen et al. (2018a)‡	154	Post + Pre	Multi-task 2D U-Net	Dice, Jc, HD, ASD	0.921 ± 0.026
Vesal et al. (2018)‡	154	Post + Pre	3D U-Net+ dilated + residual	Dice, Jc, Acc	0.925 ± 0.027
Savioli et al. (2018)‡	154	Post + Pre	3D volumetric FCN	Dice, HD	0.851 ± 0.051
Li et al. (2018a)‡	154	Post + Pre	Attention based 3D HAANet	Dice	0.923 ± 0.029
Bian et al. (2018)‡	154	Post + Pre	ResNet101 + 2D pyramid Network	Dice, Cf, Jc, HD, ASD	0.926 ± 0.022
Puybareau et al. (2018)‡	154	Post + Pre	VGG-16 + transfer learning + "pseudo-3D"	Dice	0.923 ± 0.023
Liu et al. (2018)‡	154	Post + Pre	2D U-Net + FCN	Dice	0.903 ± 0.032
Preetha et al. (2018)‡	154	Post + Pre	2D U-Net	Dice	0.887 ± 0.031
de Vente et al. (2018)‡	154	Post + Pre	2D U-Net	Dice	0.897 ± 0.035
Jia et al. (2018)‡	154	Post + Pre	Two-stage 3D U-Net + contour loss	Dice, HD, Spe, Sen	0.907 ± 0.031
Xia et al. (2018)‡	154	Post + Pre	Two-stage 3D V-net	Dice	0.932 ± 0.022
Yang et al. (2018b)‡	154	Post + Pre	Two-stage 3D U-Net + transfer learning	Dice, Cf, Jc, HD, ASD	0.925 ± 0.023
Borra et al. (2018)	154	Post + Pre	Otsu's algorithm + 3D U-Net	Dice	0.898
Jamart et al. (2019)	154	Post + Pre	Two-stage 2D V-net	Dice, Jc, HD, ASD, Diam. Err, Volume Err	0.937
Yu et al. (2019)	100	Post + Pre	Uncertainty-aware model	Dice, Jc, 95HD, ASD	0.889
Wang et al. (2019a)	100	Post + Pre	Ensembled U-Net	Dice	0.921 ± 0.020
Du et al. (2020)	100	Post + Pre	Multi-scale dual-path network	Dice, Cf, Jc, HD	0.936 ± 0.005
Borra et al. (2020)	100	Post + Pre	2D/ 3D U-Net	Dice, HD, Spe, Sen	0.895 ± 0.025/ 0.914 ± 0.015
Xiao et al. (2020)	100	Post + Pre	Multi-view network	Dice, HD, ASD	0.912
Zhao et al. (2021)	100	Post + Pre	ResNet101 + hybrid loss	Dice, 95HD, ASD	0.918 ± 0.011
Li et al. (2021c)	60	Post	Multi-task 3D U-Net + spatial encoding	Dice, HD, ASD	0.913 ± 0.032

was commonly employed as the backbone. For example, Vesal et al. (2018) employed a 3D U-Net with dilated convolutions at the bottom of the network and residual connections between encoder blocks, to incorporate both local and global knowledge. Li et al. (2018a) proposed an attention-based hierarchical aggregation network for the LA cavity segmentation, and the basic network is a 3D U-Net. Borra et al. (2020) tested both 2D and 3D U-Net for the LA cavity segmentation and found that 3D pipelines showed significantly better performance compared to the 2D pipelines. Wang et al. (2019a) utilized ensemble attention U-Net, dense U-Net, and residual U-Net models to segment LA. Liu et al. (2018), Preetha et al. (2018), and de Vente et al. (2018) all employed 2D U-Net for the LA cavity segmentation, and Liu et al. (2018) also tested the performance of fully convolutional networks (FCNs). Instead of using U-Net as the backbone, Bian et al. (2018) used ResNet101 for the LA cavity segmentation and adopted a pyramid module to learn multi-

scale semantic information in the feature map. Puybareau et al. (2018) achieved the LA cavity segmentation by transfer learning from VGG-16, a pre-trained network used to classify natural images. Savioli et al. (2018) presented a 3D volumetric FCN for the LA cavity segmentation. Besides the architecture, Jamart et al. (2020) emphasized the importance of relevant loss function selection for the LA cavity segmentation. Jia et al. (2018) proposed a novel contour loss function to include distance information for good shape consistency. Zhao et al. (2021) employed a hybrid loss to focus on the boundaries as much as on regions, and therefore reduced the impact of noisy neighboring tissues. Li et al. (2021c) introduced a spatial encoding (SE) loss to incorporate continuous spatial information of the LA. Their experiments showed that the SE loss could be effective to remove noisy patches in the final predicted segmentation, and therefore evidently reduced the Hausdorff distance (HD) value. For the loss function selection, one could refer to the review paper (Ma
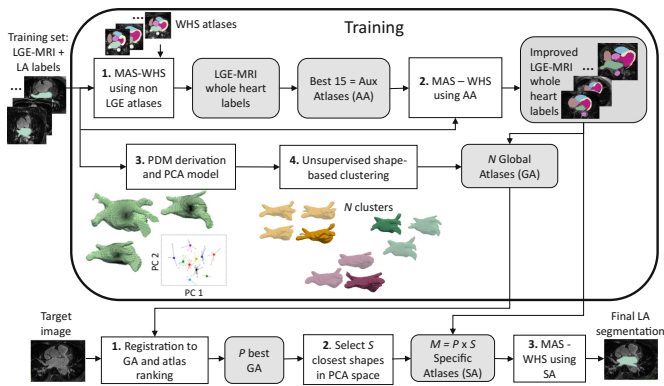
Fig. 7. A framework example for the LA cavity segmentation from Nuñez-Garcia et al. (2018). The LA cavity segmentation was achieved via multi-atlas whole heart segmentation (MAS-WHS) and shape modeling of the LA. Image adapted from Nuñez-Garcia et al. (2018) with permission.

et al., 2021), where Dice-related compound loss functions were recommended for medical image segmentation tasks.

Multi-task networks. Multi-task learning has been adopted for the LA cavity segmentation to utilize its possible relationship with other auxiliary tasks. For example, Chen et al. (2018b) and Li et al. (2021c) performed simultaneous LA cavity and scar segmentation via multi-task learning. The simultaneous optimization scheme showed better performance than solving the two tasks independently which ignored the intrinsic spatial relationship between the LA cavity and scars. Chen et al. (2018a) designed a two-task network for both LA cavity segmentation and pre/ post ablation image classification to learn additional anatomical information. The results indicated that multi-task learning obtained better segmentation performance compared to baseline U-Net method training with a single segmentation task.

Two-stage networks. A two-stage training strategy has been gradually employed to replace conventional pre-processing (such as the Otsu's algorithm employed in Borra et al. (2018)) for the region of interest (ROI) extraction. For instance, Jia et al. (2018), Xia et al. (2018), Yang et al. (2018b), and Jamart et al. (2019) all utilized two-stage U-Net/ V-Net and achieved top performances in the LA cavity segmentation. The first stage was to roughly locate the LA cavity center for ROI extraction, while the second stage was to perform the LA cavity segmentation from the cropped ROI. In this way, a memory-efficient and accurate framework was developed, and the class imbalance problem was also mitigated. It is worth mentioning that Xia et al. (2018) obtained the first-ranked results (mean Dice score of 0.932 ± 0.022) in *Left Atrium Segmentation Challenge* by using the two-stage network.

Multi-view networks. The major drawback of 2D networks is that they ignore the inter-slice correlation in the 3D LGE MRI. To solve this, a number of works have employed multi-view images as the input of networks to learn additional contextual information, namely multi-view learning. Examples include Chen et al. (2018b), Yang et al. (2020), and Xiao et al. (2020)

where the features learned from axial, sagittal, and coronal views were combined for the LA cavity segmentation. Specifically, Chen et al. (2018b) and Yang et al. (2020) regarded axial view as the main view due to its finer spatial resolution and extracted information by sequential learning; and then employed dilated residual learning to extract complementary information from sagittal and coronal views (with lower spatial resolution). Instead of employing 2D networks, Xiao et al. (2020) constructed three 3D deep convolutional streams to extract features from the patches of three views, and then fused the features for the LA cavity segmentation.

Multi-scale networks. There exists inconsistency in the sizes of LA anatomical structures such as the PVs among different patients in LGE MRI. Multi-scale networks are therefore commonly used to learn both local and global features from LGE MRI. For instance, Du et al. (2020) adopted a dual-path structure network with a multi-scale strategy for the LA cavity segmentation from LGE MRI. Xiong et al. (2018) proposed an AtriaNet consisting of a multi-scale and dual pathway architecture, to capture both local LA tissue geometries and global positional information. They evaluated their algorithm on 154 LGE MRIs and obtained average Dice scores of 0.940 ± 0.014 and 0.942 ± 0.014 for the LA epicardium and endocardium, respectively.

Uncertainty-aware models. LA structures such as the mitral valve are difficult to segment due to the lack of a clear anatomical border between the LA and the LV. The ambiguity of the boundary gives rise to uncertainty for the LA cavity segmentation. Yang et al. (2018b) designed a composite loss to combat uncertainty, and the main idea was to enlarge the gap between background and foreground predictions. Yu et al. (2019) proposed an uncertainty-aware self-ensembling model for semi-supervised LA cavity segmentation. This is achieved by encouraging the segmentation to be consistent for the same input under different perturbations of the unlabeled data. Therefore, they could use abundant unlabeled data for training and obtained similar performance compared to the fully supervised methods using abundant labeled data.

3.1.3. Summary of LA cavity segmentation methods

In summary, conventional methods generally rely on the information from shape priors or additional paired MRI/ MRA for accurate LA cavity segmentation from LGE MRI. However, acquiring the auxiliary images requires extra work, and may introduce further errors, i.e., misalignment between LGE MRI and the auxiliary images. Recently, with the development of DL and the release of public data, many methods could directly segment the LA cavity from LGE MRI, and achieved promising results. However, there still exist large errors in the PV and MV regions. This is mainly due to the small size, the large variability of PVs, including the number, position and orientation of the PVs, and the unclear boundary of MV. Note that PVs are crucial structures for AF analysis, as scars and ablation gaps are mainly located around PVs after PVI procedures. To improve the performance of DL-based methods, multi-task learning is effective,

and a two-stage network is also a recommended training strategy. It is also important to include shape prior or spatial information into the DL-based framework for robust LA cavity segmentation, especially when the size of training dataset is small. Besides, the accuracy of segmentation was found to be correlated to the image quality of LGE MRI (Pearson's correlation = 0.38, *p*-value = 0.005) (Xiong et al., 2020). It is interesting that the reviewed methods show that 2D and 3D convolutional neural networks (CNNs) had comparable performance, though the target LGE MRI belongs to a 3D image.

3.2. LA wall segmentation

To the best of our knowledge, there are limited works reported for automatic LA wall segmentation in the literature, especially from LGE MRI. Many groups estimated the LA wall from LGE MRI just as an initialization step for the LA scar segmentation (Karim et al., 2013; Yang et al., 2018a; Wu et al., 2018). These works are not included in this section, as most of them simply dilated the generated LA endocardium by assuming a fixed wall thickness for approximated LA wall segmentation (Karim et al., 2013). However, LA wall thickness varies with positions of the same patient and patients with different gender, age, and disease status (Pan et al., 2008). With an accurate segmentation result, the wall thickness, which is useful in clinic studies, could be calculated. For the review of existing techniques of wall thickness measurement, one can refer to Table 1 of the benchmark paper (Karim et al., 2018). *Considering the limited number of works reported on LA wall segmentation, in this section we further review the segmentation on other modalities, including non-enhanced MRI and CT.* Table 7 summarizes the representative works and results from (LGE) MRI and CT.

3.2.1. Conventional methods for LA wall segmentation

Morphological operations. The most straightforward method is to perform morphological operations on the LA endocardium by assuming a fixed wall thickness. For example, Bishop et al. (2016) adopted morphological operations on the segmented blood pool for wall segmentation from CT. This method ignores the thickness variation among different LA positions.

Deformable models. In contrast, deformable models can dynamically adapt to the changes of wall thickness, and hence obtain more plausible LA wall segmentation results. For example, Tao et al. (2016b) used the level-set approach to extract the inner and outer LA surface for the final wall segmentation. Jia et al. (2016) adopted the region growing method for endocardial segmentation and then utilized Marker-controlled geodesic active contour for the epicardial segmentation. Karim et al. (2018) presented the LA wall segmentation and thickness measurement results using three conventional methods, i.e., level-set, region growing, and watershed. The results showed that level-set performed evidently better than the other two methods; region growing generally over-estimated thickness and performed poorly in the wall segmentation task. They also found that algorithms performed worse in MRI than in CT, which may be due to the fact that the image quality of MRI was generally worse

than CT. However, CT has limited soft tissue contrast, so Tao et al. (2016b) employed nonlinear intensity transformation to enhance the LA wall region in CT.

Laplace-based solutions. Laplace-based solutions generate a series of smooth non-intersecting field lines between two boundaries in space and are ideal for simulating the highly variable LA epicardial and endocardial surfaces. Wang et al. (2019b) employed the multi-planar convex hull approach to extract the epicardial and endocardial surfaces, and then used the coupled partial differential equations (PDE) for the wall thickness measurement. They evaluated their method on both LGE MRI and ex vivo data, and observed that wall thickness values in LGE MRI were more difficult to measure and validate. Besides, there was a discrepancy in wall thickness measured by ex vivo data and LGE MRI. Specifically, the wall thickness values measured from ex vivo data were consistently higher than those measured in LGE MRI. Zhao et al. (2017) calculated the wall thickness by solving the Laplace equations on both epicardial and endocardial surfaces. Despite its prominence, the Laplace-based method still requires explicitly calculating gradient as well as distance trajectories, which are time-consuming and error-prone (Wang et al., 2019c).

Graph-based methods. Graph-based methods are promising alternatives. Veni et al. (2017) proposed a shape-based generative model namely ShapeCut, to extract epicardial and endocardial surfaces for the LA wall segmentation from LGE MRI, as presented in Fig. 8. The model could incorporate both local and global shape priors within a maximum-a-posterior estimation framework, and the shape parameters could be optimized via graph-cuts algorithm. The optimization could be executed in two phases in an iterative manner, i.e., one for multi-surface updates based on multi-column graphs and the other for global shape refinement based on closed forms. For evaluation, besides directly assessing the LA wall segmentation performance, they also adopted the LA scar segmentation based on their LA wall segmentation for further evaluation. Specifically, they extracted the scars using thresholding based on both manual and automatic wall segmentations. Then, they plotted the fibrosis percentage from manual annotations versus that from automatic ones for each scan. They obtained a linear relation with a small error, demonstrating a high overlap between the manual and automatic scarring regions. Here, the linear relation error was indicated using the MSE and R-square values.

3.2.2. Summary of LA wall segmentation methods

In summary, currently reported works were all based on conventional methods, and no DL-based method has been reported, to the best of our knowledge. This could be due to the limited number of relevant public datasets and the large inter- and intra-observer variations of the manual segmentation. As Karim et al. (2018) reported, a common error of LA wall segmentation arises from the surrounding tissue such as the neighboring aortic wall. Improving the image quality may mitigate this problem, and the active contour-based methods with shape constraints and coupled level-set approaches could be helpful. One of the main applications of LA wall segmentation is to measure

Table 7. Overview of previously published works on the *LA wall segmentation* from (LGE) MRI and CT. MSE: mean square error; A: anterior; P: posterior; Tk: thickness; GAC: geodesic active contour; PDE: partial differential equations. Note that the evaluation measures and results in Inoue and Drangova (2016), Tao et al. (2016b) and Jia et al. (2016) are from the benchmark paper (Karim et al., 2018).

| Study	Data	Algorithm	Evaluation	Result
Hsing et al. (2014)	55 LGE MRI	Manual	Tk	Tk = 7.0 ± 1.8 mm (before ablation) Tk = 10.7 ± 4.1 mm (after ablation)
Veni et al. (2017)	72 LGE MRI + 170 Synthetic	ShapeCut	ASD, HD, clinical evaluation	Synthetic: ASD = 0.25 ± 0.04 mm; HD = 1.95 ± 0.38 mm LGE MRI: ASD = 0.66 ± 0.14 mm LGE MRI scar segmentation: MSE = 3.07; R-square = 0.83
Zhao et al. (2017)	LGE MRI	Laplace equation	Tk	Tk = 3.7 ± 1.7 mm
Wang et al. (2019b)	154 LGE MRI + ex vivo data	Convex hull method + coupled PDE	Tk	LGE MRI: Tk = 0.4–11.7 mm and median = 3.88 mm
Karim et al. (2018)	10 MRI	Level-set Region growing Watershed	Tk, Dice, tissue mass	Tk = 2.16 ± 0.58 mm, Dice = 0.72 Tk = 6.04 ± 3.63 mm, Dice = 0.39 Tk = 3.45 ± 3.57 mm, Dice = 0.67
Inoue et al. (2014)	5 enhanced CT	Multi-region segmentation software + manual correction	Tk, visualization	Tk = 0.5-3.5 mm
Bishop et al. (2016)	10 CT	Morphological operations + Laplace equation	Tk	Errors ≤ 0.2 mm for Tk of 0.5–5.0 mm
Inoue and Drangova (2016)	10 CT	Mesh vertex normal traversal	Tk, Dice, tissue mass	Tk = 1.13 ± 1.02 mm (A), 1.26 ± 0.83 mm (P) Dice = 0.33 (A), 0.39 (P)
Tao et al. (2016b)	10 CT	Nonlinear intensity transformation + level-set	Tk, Dice, tissue mass	Tk = 1.34 ± 0.89 mm (A), 0.78 ± 0.41 mm (P) Dice = 0.43 (A), 0.21 (P)
Jia et al. (2016)	10 CT	Region growing + Marker-controlled GAC	Tk, Dice, tissue mass	Tk = 0.75 ± 0.38 mm (A), 1.46 ± 1.57 mm (P) Dice = 0.30 (A), 0.50 (P)
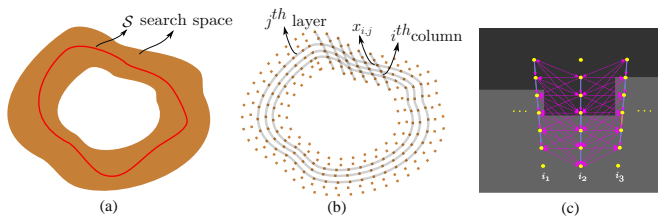
Fig. 8. A framework example for the LA wall segmentation from Veni et al. (2017). The LA wall segmentation was achieved by predicting the LA epicardium and endocardium respectively, so the task was converted into a surface estimation problem. ShapeCut constructed a geometric graph to discretize continuous parameterization of the set of possible surfaces for global optimization. (a) continuous parameterization of the surface estimate S; (b) discrete approximation of the continuous parameterization where the surface estimation performed. Here, each layer maintained a topology similar to the desired surface, and each column ensured that the estimated surface traverses it; (c) shape complying properly ordered graph construction. Picture modified from Veni et al. (2017) with permission.

wall thickness. Most of the reported algorithms relied on ruler-based assessments via digital calipers instead of performing a prior segmentation of the LA wall (Karim et al., 2018). Several works employed the Laplace equation or PDE to measure wall thickness after achieving the LA wall segmentation. Karim et al. (2018) demonstrated that their proposed wall thickness atlas could be effective for thickness prediction in new cases via atlas propagation. They constructed a flat thickness map via a surface flattening and unfolding strategy, to compare the mean thickness in each sub-region of the LA wall. Finally, though CT is a good modality for imaging the thin wall owing to its

high resolution, MRI could be effective to assess the wall tissue viability. Therefore, more attention is expected to the LA wall segmentation from MRI, especially LGE MRI.

3.3. LA scar segmentation and quantification

In the literature, a limited number of works have been reported targeting the fully automatic segmentation or quantification of LA scars, probably due to the particular challenge of this task. Most of the methods require an accurate initial manual segmentation of the LA cavity or LA wall for the following scar classification on the LA wall. For example, *Left Atrium Fibrosis and Scar Segmentation Challenge* (Rhode and Karim, 2012) provided LA cavity labels for participants to develop scar segmentation algorithms. Eight research teams contributed their methods to this task, including histogram analysis, thresholding, k-means clustering, region-growing with EM-fitting, active contour, and graph-cuts (Karim et al., 2013). The benchmark study showed that semi-automatic methods initialized with manual LA wall segmentation were much more reliable, and performed better than fully automatic approaches (Karim et al., 2013). Currently, the most commonly used approach for the LA scar segmentation is based on thresholding, which is nevertheless sensitive to intensity changes of LGE MRI (Pontecorboli et al., 2017). Table 8 summarizes all the works, where conventional methods are listed in the upper part and DL-based algorithms are enumerated in the bottom part.

Table 8. Summary of previously published works on the (semi-)automatic *LA fibrosis/ scar segmentation and quantification* from LGE MRI. ~M Post: ~ months post-ablation; Pre: pre-ablation; XOR: XOR overlap; Percentage: scar percentage; RMSE: root MSE; Volume: total scar volume; ROC: receiver operating characteristic; BER: balanced error rate; FCC: fuzzy c-means clustering; MIP: maximum intensity projection; GMM: Gaussian mixture model; SSAE: stacked sparse auto-encoders; MS-CNN: multi-scale CNN; EAM-c: correlation with electroanatomical mapping; FWHM: full-width-at-half-maximum; ICC: intraclass correlation coefficients; PCC: Pearson correlation coefficient. Here, the asterisk (*) indicates the method employed manual LA wall segmentation.

Study	Num	Pre/ Post	Algorithm	auto?	Evaluation	Dice$_{scar}$
Oakes et al. (2009)	81	Pre	2–4 SD	semi-auto	Percentage, EAM-c	N/A
Badger et al. (2010)	144	3-20M Post	3 SD	semi-auto	Percentage, EAM-c	N/A
Knowles et al. (2010)	7	Post + Pre	MIP	semi-auto	Percentage, EAM-c	N/A
Karim et al. (2011)	9	6M Post	Probabilistic intensity model	auto	Percentage	N/A
Perry et al. (2012)	34	3M Post	K-means clustering	semi-auto	Dice, XOR, percentage	0.807 ± 0.106
Karim et al. (2013)	60	Post + Pre	Hysteresis threshold	semi-auto	Dice, RMSE, volume	0.76^{post}; 0.37^{pre}
	60	Post + Pre	Region growing + EM-fitting			0.85^{post}; 0.22^{pre}
	40	Post + Pre	Graph-cuts + FCC			0.73^{post}; 0.17^{pre}
	15	Post	Active contour + EM-fitting			0.76^{post}
	30	Post + Pre	Simple threshold*			0.84^{post}; 0.48^{pre}
	60	Post + Pre	Graph-cuts			0.78^{post}; 0.30^{pre}
	60	Post + Pre	Histogram analysis + threshold*			0.78^{post}; 0.42^{pre}
	60	Post + Pre	K-means clustering*			0.72^{post}; 0.45^{pre}
	60	Post + Pre	2 SD	semi-auto	Dice, RMSE, volume	0.58^{post}; 0.24^{pre}
	60	Post + Pre	3 SD			0.17^{post}; 0.16^{pre}
	60	Post + Pre	4 SD			0.14^{post}; 0.31^{pre}
	30	1-6M Post	6 SD			0.35^{post}
	60	Post + Pre	FWHM			0.59^{post}; 0.05^{pre}
Ravanelli et al. (2014)	10	Pre	4 SD	semi-auto	Dice, EAM-c	0.600 ± 0.210
Karim et al. (2014)	15	6M Post	GMM + graph-cuts	semi-auto	Dice, ROC, volume	> 0.8
Tao et al. (2016a)	46	Pre	MIP	auto	Qualitative visualization	N/A
Veni et al. (2017)	72	Post + Pre	K-means clustering	auto	Percentage	N/A
Yang et al. (2018a)	37	Post	Super-pixel + SVM	auto	Dice, Acc, Sen, Spe, ROC, BER	0.790 ± 0.050
Wu et al. (2018)	36	Post	Multivariate mixture model	auto	Dice, Acc, Sen, Spe	0.556 ± 0.187
Razeghi et al. (2020)	207	12M Post + Pre	MIP	auto	ICC, PCC, RMSE	N/A
Yang et al. (2017b)	20	Post + Pre	Super-pixel + SSAE	auto	Dice, Acc, Sen, Spe, ROC	0.776 ± 0.146
Li et al. (2018b)	100	Post + Pre	Graph-cuts + CNN	auto	Dice, Acc, Sen, Spe	0.566 ± 0.140
Chen et al. (2018b)	100	Post + Pre	Multi-view two-task network	auto	Dice, Acc, Sen, Spe, percentage	0.78 ± 0.08
Yang et al. (2020)	190	Post + Pre	Multi-view two-task network	auto	Dice, Acc, Sen, Spe	0.870
Li et al. (2020b)	58	6M Post	Graph-cuts + MS-CNN	auto	Dice, Acc, Sen, Spe, GDice	0.702 ± 0.071
Li et al. (2020a)	60	3-27M Post	Multi-task network	auto	Dice, Acc, GDice	0.543 ± 0.097

3.3.1. Conventional methods for LA scar segmentation and quantification

Thresholding. Thresholding is the most popular method for LA scar segmentation. The threshold value is normally defined by assuming a fixed standard deviation (SD) above the average intensity value of the normal wall region or blood pool (Oakes et al., 2009; Badger et al., 2010; Ravanelli et al., 2014). For details, one can refer to the survey from Pontecorboli et al. (2017), where different thresholding-based scar segmentation techniques were reviewed and compared. These methods are easy to implement and intuitive, but also have several disadvantages. Firstly, the selection of threshold values is subjective, and the values can differ significantly across various scans, due to the difference of timing from gadolinium administration (Karim et al., 2014; Chubb et al., 2018). Secondly, the performance of scar segmentation highly relies on the accuracy of LA or LA wall segmentation that is also challenging, and therefore thresholding based LA scar segmentation was typically achieved via semi-automatic or manual approaches (Oakes et al., 2009; Badger et al., 2010). The benchmark paper (Karim et al., 2013) compared eight methods with the full-width-at-half-maximum (FWHM) and n-SD methods, and all thresholding methods employed manual LA cavity segmentation as initialization and three of them further utilized manual LA wall segmentation. In general, all the evaluated eight methods in the benchmark paper outperformed the FWHM and n-SD methods.

Maximum intensity projection. Similar to thresholding, maximum intensity projection (MIP) is also a scar quantification scheme that employs scar intensity characteristics. However, unlike thresholding, MIP is more robust to the inaccurate LA cavity segmentation due to the projection step. Examples in-

clude Knowles et al. (2010) and Tao et al. (2016a), where projection was performed at ±3 mm and ±2 mm along each normal vector of the LA surface respectively, to consider the potential errors of LA cavity segmentation. Razeghi et al. (2020) also employed MIP for scar segmentation (3 mm externally and 1 mm internally). Nevertheless, the projection range of MIP must be selected carefully, as it needs to be large enough to extend into the LA myocardium, but not too far to include the intensity of other regions.

Clustering algorithms. Considering the complex intensity distribution of LGE MRI, clustering algorithms could be another solution for LA scar segmentation. This is because clustering can provide a mechanism to statistically separate voxels into groups that are analogous to various tissue types, such as blood pool, healthy wall tissue, and scars. Perry et al. (2012) employed k-means clustering to segment scars from manually segmented LA wall regions. Veni et al. (2017) used the same k-means clustering method as Perry et al. (2012), and the LA wall was automatically segmented by their proposed ShapeCut method. Yang et al. (2018a) employed super-pixel via a linear iterative clustering algorithm to over segment scars, and then utilized the support vector machine algorithm to classify the over-segmented super-pixels into scarring and normal wall regions. They scored the image quality into 0 (non-diagnostic), 1 (poor), 2 (fair), 3 (good), and 4 (very good) on a Likert-type scale, according to the level of signal to noise ratio (SNR), appropriate T1, and the existence of navigator beam and ghost artifacts. Only subjects with image quality $\geq$ 2 were selected into their study for evaluation. Wu et al. (2018) combined LGE MRI with anatomical MRI for the scar quantification based on the multivariate mixture model (MvMM) and maximum likelihood estimator (MLE). They formulated a joint distribution of images using the MvMM (Zhuang, 2019), where the registration of the two MRIs and scar segmentation of LGE MRI were performed simultaneously. Then, the transformation and model parameters were optimized by an iterated conditional model algorithm within the MLE framework.

Deformable models. Two deformable models were employed to segment LA scars from LGE MRI, i.e., region growing and active contour with EM-fitting, as reported in Karim et al. (2013). Among the eight methods mentioned in Karim et al. (2013), region growing with EM-fitting method obtained the best performance on a post-ablation dataset in terms of Dice, even better than those methods that directly employed manual LA wall segmentation for initialization. For pre-ablation data, the three methods with manual LA wall initialization achieved evidently better Dice compared to the other five methods only with manual LA initialization. Similar to Yang et al. (2018a), Karim et al. (2013) classified the LGE MRIs into three types, i.e., good, average, and poor, according to its SNR and contrast ratio (CR) for scars. They found that most methods had a marginally lower Dice on scans with worse quality, but without statistical significance. This could be attributed to the minor quality difference and accurate initialization of manual LA cavity segmentation.

Graph-based methods. Graph-based methods naturally consider inter-dependencies by introducing links (or edges) between related objects, thus effectively capturing their long-range relatedness. It may be an effective solution to capture these small and diffuse scars distributed on the LA wall. Karim et al. (2011) proposed a probabilistic tissue intensity model which was formulated as a Markov random field and solved using graph-cuts. In their following work (Karim et al., 2014), they presented a scar quantification method by combining the scar intensity model priors and Gaussian mixture model (GMM). Besides, they added constraints via the graph-cuts approach to ensure smoothness and avoided discontinuities in the final scar segmentation. The proposed method was evaluated on both numerical phantoms and clinical datasets, and demonstrated a good concordance between the automatic results and manual delineations. Here, numerical phantoms could offer a wide range of variation in scar contrast, which is usually unavailable in clinical datasets.

3.3.2. Deep learning-based methods for LA scar segmentation and quantification

Yang et al. (2017b) was the first work applying a DL-based classifier for the LA scar segmentation. Specifically, they used super-pixel over-segmentation for feature extraction, and then adopted a supervised classification step via stacked sparse auto-encoders. However, they only used handcrafted intensity features, which provided limited information. Similar to the DL-based LA cavity segmentation methods, multi-scale, multi-view, and multi-task networks were also employed for LA scar segmentation and quantification.

Multi-scale networks. As Fig. 3 (d) shows, the surrounding enhanced regions can seriously disrupt the segmentation of scars. Multi-scale learning could be an effective strategy to alleviate the interference, as it provides both local and global views when learning features of scars. Li et al. (2018b) proposed a hybrid approach utilizing a graph-cuts framework combined with CNNs to predict edge weights of the graph for the automatic scar segmentation. They extended their work by introducing multi-scale CNN (MS-CNN) to learn local and global features simultaneously (Li et al., 2020b), as presented in Fig. 9. The experimental results showed that the multi-scale learning scheme (number of scales = 3) improved the performance when compared with a single scale (Dice$_{scar}$: 0.702 ± 0.071 vs. 0.677 ± 0.070). Besides, the scheme is also less dependent on an accurate LA cavity segmentation, which makes it more robust. A major limitation of this study was the lack of an end-to-end training style, as the framework was split into three subtasks, i.e., LA cavity segmentation as an initialization, feature learning via the MS-CNN, and optimization based on graph-cuts. This indicated the limitation of multi-scale patch strategies, which resulted in an expensive time and space complexity and an infeasible end-to-end training on the whole graph.

Multi-task/ multi-view networks. To achieve end-to-end optimization, multi-task learning is desired. Li et al. (2020a) developed a new framework where LA cavity segmentation, scar
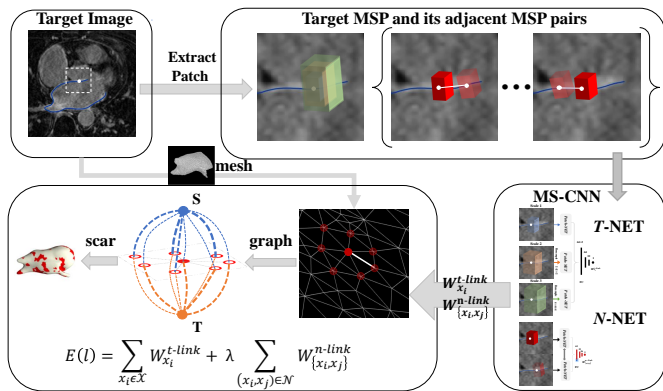
Fig. 9. A framework example for the LA scar quantification from Li et al. (2020b). The scar quantification was performed on the LA surface mesh where a graph was constructed, and the weights of a graph were explicitly learned via an MS-CNN. Image adapted from Li et al. (2020b) with permission.

projection onto the LA surface, and scar quantification are performed simultaneously in an end-to-end fashion based on a multi-task network. In this framework, they proposed a shape attention (SA) mechanism by an implicit surface projection, to utilize the inherent spatial relationship between the LA cavity and scars. The mechanism also alleviated the class-imbalance problem in the scar quantification and proved to be effective in the ablation study. Similarly, Chen et al. (2018b) and Yang et al. (2020) adopted multi-task learning for simultaneous LA and scar segmentation, but the spatial relationship between the two regions was not explicitly learned in their works. Moreover, as mentioned in Section 3.1.2, they employed multiple views as the input of multi-task networks.

3.3.3. Summary of LA scar segmentation and quantification methods

In summary, scar segmentation/ quantification from LGE MRI remains an open problem. Most methods relied on interactive correction/ manual initialization, or on accurate initial estimation of LA wall segmentation for following application of thresholding. These semi-automatic approaches generally obtained high accuracies in terms of Dice scores. Compared to the conventional automatic methods, DL-based algorithms could obtain better performance. However, DL-based models could have limited model generalization ability. In general, pre-ablation data with fibrosis is more challenging to segment than post-ablation data with scars. This may be attributed to the fact that fibrosis appears more diffusely compared to post-ablation scars (Karim et al., 2013). In addition, it is difficult to differentiate the native fibrosis and post-ablation scars for long-standing persistent AF patients (Yang et al., 2017a). One major challenge for scar segmentation/ quantification is the artifacts from the boundary regions, such as the right atrial (RA) wall and aorta wall. A good initialization, i.e., accurate LA or LA wall segmentation, could be helpful to counteract this problem. Li et al. (2020b) tried to reduce the dependence on accurate LA cavity segmentation via projection and MS-CNN, while Li et al. (2020a) introduced a distance-based spatial encoding loss

for training a deep neural network to learn the spatial information of scars around the LA boundary. Another challenge arises from the imaging, including poor image quality and data-mismatch issues in DL-based methods. Therefore, a more consistent and standard image acquisition protocol is highly required. Alternatively, domain generalization algorithms need to be considered to improve the model generalization ability across different sites or on unseen datasets (Li et al., 2021a; Campello et al., 2021).

3.4. LA ablation gap quantification

Gaps around PVs can be classified into electrical/ conduction gaps and anatomical ablation gaps. Conduction gaps refer to the electrical reconnection regions with high voltages in the electroanatomical mapping (EAM), and they can be detected using intra-cavitary catheters during a redo procedure. Ablation gaps indicate the gaps of healthy tissue in the (ideally continuous) scars, which are typically identified by LGE MRI. Therefore, in this section, we only focus on the developed methods to quantify ablation gaps from LGE MRI. Note that the ablation gaps do not belong to the inherent structure of the LA, but instead are "gaps" left during the LA ablation procedure. Table 9 summarizes representative (semi-)automatic LA ablation gap quantification methods, results, and main findings.

3.4.1. Conventional methods for LA gap quantification

Visual detection. To the best of our knowledge, most of the methods reported in the literature relied on visual inspection, which could result in biased estimations of gap characteristics, such as the number, length, and position of gaps. For instance, Badger et al. (2010) and Mishima et al. (2019) both employed thresholding for the scar segmentation and then detected ablation gaps visually. Moreover, as ablation gaps are highly correlated with scars, there is a certain overlap for quantification methods of scars and ablation gaps, such as MIP and thresholding. Bisbal et al. (2014) manually segmented the LA wall for an accurate initialization and then adopted MIP for the scar and gap classification. Linhart et al. (2018) used the image intensity ratio as a threshold for LA scar segmentation and defined the gaps as the discontinued ablation line $\leq$ 3 mm. Several software packages were also employed for ablation gap quantification, such as *Osirix* (Ranjan et al., 2012) and Custom-written software (Harrison et al., 2015a).

Graph-based methods. Recently, Nuñez-Garcia et al. (2019) proposed a reproducible framework for semi-automatic gap quantification using a graph-based method, as presented in Fig. 10. One can see that the gap quantification was performed via minimum path search in a graph where each node was a scarring patch, and the edges denoted the geodesic distances between patches. They proposed a quantitative measure to estimate the percentage of gaps around a vein, namely the relative gap measure. One major limitation of this work was that a fixed regional parcellation was assumed, i.e., four-PV configuration in the LA, but actually only around 70% of LA have four PVs (Prasanna et al., 2014).

Table 9. Summary of previously published works on the (semi-)automatic *ablation gap quantification* from LGE MRI. # gaps: mean number of gaps; GL: gap length; IIR: image intensity ratio; NAUC: normalized area under the curve; RSPV: right superior PV; LIPV: left inferior PV; LSPV: left superior PV.

Study	Num	Algorithm	Evaluation	Results & Main findings
Badger et al. (2010)	144	3 SD for scar segmentation + visually detect gap	Visual, EAM-c	Significant relationship between gaps and recurrence; Achieving complete circumferential lesions around the PV is difficult.
Ranjan et al. (2012)	12	Measurement tool in *Osirix*	GL, pathology correlation	The correlation coefficient (R^2) between the GL identified by LGE MRI and the gross pathology was 0.95; GL = 1.0 mm (via gross pathology) and GL = 1.4 mm (via LGE MRI); Real time MRI system can be used to identify gaps.
Bisbal et al. (2014)	50	Manual LA wall + MIP	GL, # gaps, EAM-c	# gaps = 4.4/patient; # gaps = 1.27 ± 0.41/PV Median GL = 13.33 ± 5.8 mm/gap; Position of highest # gaps: RSPV (=1.53); Position of fewest # gaps: LIPV (=0.67); The majority of patients (73.3%) had gaps in all PVs. LGE MRI may identify non-conducting gaps that could be related to later recurrences.
Harrison et al. (2015a)	20	Custom-written software	Visual, EAM-c	Weak point-by-point relationship (R^2=0.57) between scars and endocardial voltage in patients undergoing repeat LA ablation; The mean voltage within scar region is lower than that within normal wall region.
Linhart et al. (2018)	94	IIR for scar segmentation + gap is defined as discontinuation of the ablation line by ≤ 3 mm	GL, # gaps, EAM-c	# gaps = 5.4/patient; Mean GL = 7.3 mm/gap; 90 out of 94 patients (96%) had at least 1 anatomic gap; Anatomic gaps are frequently detected in LGE MRI at 3 months after first PVI; An increase of 10% relative GL increased the likelihood of AF recurrence by 16%.; The total relative GL was significantly associated with recurrence instead of # gaps.
Mishima et al. (2019)	10	2 SD for scar segmentation + visually detect gap	GL, # gaps, EAM-c	Mean GL = 11.6 ± 3.9 mm/gap; # gaps = 1.6/patient (1st ablation); # gaps = 1.4/patient (2nd ablation); Position of highest # gaps: RSPV (=2); Position of fewest # gaps: LIPV (=0); The location of electrical gaps are well matched to that on the LGE MRI.
Nuñez-Garcia et al. (2019)	50	Graph-based method	GL, # gaps, RGM	Position of highest # gaps: LSPV (=1.73); Position of fewest # gaps: LIPV (=1.16); No significant differences between left and right PVs; No significant relationship between gaps and recurrence.

3.4.2. Summary of LA ablation gap quantification methods

It is considered difficult to achieve complete circumferential lesions, so the majority of patients have gaps after ablation (Badger et al., 2010; Bisbal et al., 2014; Linhart et al., 2018). The most common locations appearing gaps are the area between the left superior PV (LSPV) and the LAA. This may be due to the presence of a thicker myocardium in this area, which leads to non-transmural lesions (Galand et al., 2016). In Bisbal et al. (2014) and Mishima et al. (2019), the largest number of gaps occurred in right superior PV (RSPV) was reported; while in Nuñez-Garcia et al. (2019) it appeared in LSPV. In contrast, the fewest of gaps occurred consistently in the left inferior PV (LIPV) (Bisbal et al., 2014; Mishima et al., 2019; Nuñez-Garcia et al., 2019). The different distributions of gaps in different PV positions could be attributed to the differences in imaging and limited accuracy of scar segmentation in these regions.

The relationship between electrical gaps of EAM and anatomical gaps of LGE MRI is still unclear. Mishima et al. (2019) found that the location of electrical gaps was well matched to that of the detected ablation gaps from LGE MRI. However, Harrison et al. (2015a) claimed a weak point-by-point relationship between scars and EAM in the patients with repeated LA ablation. Besides, the relationship between ablation gaps and AF recurrence is also controversial, with positive answers (Peters et al., 2009; Taclas et al., 2010; Badger et al., 2010; Bisbal et al., 2014; Linhart et al., 2018) but also negative conclusions (Spragg et al., 2012; Harrison et al., 2015b; Nuñez-Garcia et al., 2019). These are partially due to the lack of an objective and consistent method for ablation gap quantification, primarily depending on visual observation. The task has not been properly addressed in the literature, and research on this is still in an early stage.
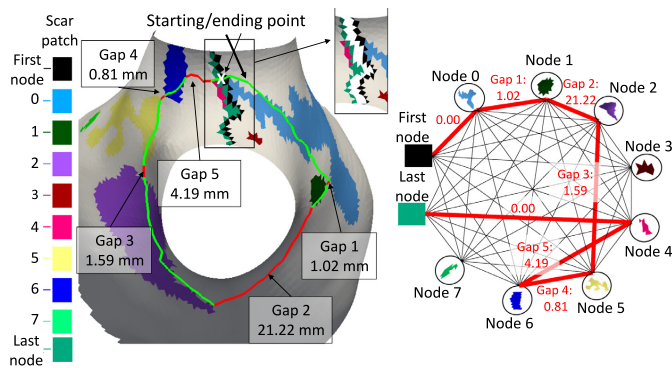
Fig. 10. A framework example for the LA gap quantification from Nuñez-Garcia et al. (2019). The gap quantification was performed through minimum path search in a graph where every node referred to a scarring patch and the edges denoted the geodesic distances between patches. Image adapted from Nuñez-Garcia et al. (2019) with permission.

3.5. Image computing and analysis on the LA LGE MRI

So far, we have presented and discussed the recent progress in LA LGE MRI computing. Table 10 summarizes the various properties of different targets with corresponding potential processing schemes. The LA cavity is a relatively large target but with variable shapes; the LA wall is equivalent to two surfaces with extremely small and inconsistent distance; and the LA scars/ ablation gaps belong to small, discrete, and space-constrained (scars and ablation gaps are localized at the LA wall) targets with distinct features. Most of the methods summarized here are customized to the corresponding attributes and challenges of each task. For example, due to the variable shapes of the LA cavity, many atlas-based methods were proposed to incorporate the shape priors. Auxiliary images, uncertainty-aware, and coarse-to-fine training schemes are also beneficial for LA cavity segmentation. Due to the properties of the LA wall, variants of deformable models were employed, such as coupled level-set, region growing, and watershed algorithms. With an accurate LA initialization, it is straightforward yet effective to adopt thresholding for the scar segmentation, as scarring regions are enhanced in intensity compared to the healthy wall. Moreover, due to the thin wall, some researchers proposed to project the scars onto the LA surface ignoring the wall thickness for scar quantification.

Nevertheless, there is a certain overlap in these reviewed approaches, mainly as the four tasks are coherent and share similar challenges (please refer to Sec 1.2 of the manuscript for the challenges of each task). Among the *conventional methods*, several classical algorithms were commonly employed, such as graph-based methods, deformable models, and clustering algorithms. For example, Fig. 8, Fig. 9, and Fig. 10 present the graph-based methods for LA wall segmentation, LA scar quantification, and LA gap quantification, respectively. It is evident that for different tasks the graphs were constructed in different styles. Specifically, for LA wall segmentation a graph was represented by a set of columns and a neighborhood structure among adjacent columns for a multi-surface update, namely a multi-column graph. For LA scar quantification, a graph was designed on the LA surface mesh, and the graph weights

were learned by MS-CNN. For LA gap quantification, the scar patches were regarded as the nodes of a graph, and the geodesic distances between patches were denoted as the edges. Among the *DL-based methods*, there are several commonalities for LA cavity and scar segmentation/ quantification, which can be categorized into three kinds, (1) alleviating the class imbalance problem via pre-processing, a two-stage pipeline, or weighted sampling; (2) improving the robustness of networks via multi-scale learning, multi-task learning, or multi-view feature fusion; (3) forcing the network to generate more plausible segmentation results by incorporating shape priors, applying anatomical constraints, or introducing uncertainty maps. It is worthwhile to highlight that for LA cavity and scar segmentation/ quantification, leveraging spatial relationship of LA cavity and scars via simultaneous optimization has been explored and shown to be beneficial for improving the accuracy.

There are apparent trade-offs between conventional and DL-based algorithms. Conventional approaches are transparent and well-established, while DL has potential of higher precision and versatility but with the cost of an enormous amount of data and computing resources (O'Mahony et al., 2019). Therefore, it is interesting to explore hybrid approaches combining the advantages of them. Several works have demonstrated their benefits for LA LGE MRI computing. For example, for LA cavity segmentation, Borra et al. (2018) utilized Otsu's algorithm to extract ROI and then performed segmentation on the ROI via U-Net. Li et al. (2021c) employed the conventional distance transform maps to incorporate continuous spatial information of the target label. The limited receptive view and spatial awareness in the standard CNN-based methods could lead to a noisy segmentation, especially for the target with highly variable shapes, such as LA. Their results showed the effectiveness of distance transform maps in the DL-based framework removing the noisy patches of the segmentation. Statistical shape models (SSMs) can be a promising alternative to combine CNN with prior knowledge of anatomical shapes for the LA cavity segmentation (Ambellan et al., 2019). Recently, DL-based cross-modality MAS frameworks are promising for the left ventricle (LV) myocardial segmentation (Ding et al., 2020), and could be extended for the LA cavity segmentation, especially when the additional paired modalities are available. For LA scar quantification, Li et al. (2020b) combined the conventional graph-cuts algorithm and MS-CNN (LearnGC) for hybrid representations of structural and local features. They employed MS-CNN to learn multi-scale features of patches corresponding to nodes on the graph and obtained better results than conventional graph-cuts algorithms which were based on hand-crafted features. For LA wall segmentation and gap quantification, no DL-based method has been reported, to the best of our knowledge. However, the conventional ShapeCut algorithm proposed by Veni et al. (2017) can be adapted for such application, by extracting features from the intensity profiles via CNN for more accurate LA wall segmentation. Similar schemes can be employed on the proposed graph-based method for LA gap quantification (Nuñez-Garcia et al., 2019). Moreover, the utility of level-set for LA wall segmentation has been proven (Karim et al., 2018), and the combination of DL and level-set for the

Table 10. Summary of current computing methods on the LA LGE MRI for AF analysis.

Target	Unique characteristic	Potential processing strategies & future perspective
LA cavity	large shape variability	introducing shape prior; combining complementary information from other paired modalities; uncertainty-aware schemes; coarse-to-fine training
LA wall	thin thickness; irregular opening; varying thickness across the wall	level-sets; ShapeCut and its possible future version combining DL-based feature extraction
LA scars	small; diffusion; spatial constrained; intensity-related; corresponding to the low voltage regions of EAM	multi-scale learning schemes; surface projection; multi-task learning schemes; thresholding based on an accurate initialization
LA gaps	without an unified definition; uncertain number of PVs	quantitative instead of visual qualitative quantification

LV segmentation obtained accurate results with small training sets (Ngo et al., 2017). Therefore, such combination and hybrid approaches are expected and should be further explored in the near future.

4. Data and evaluation measures

Validation work not only reveals the performance and limitations of a proposed method, but also clarifies the scope of its application (Jannin et al., 2006). Hence, it is essential to validate an algorithm before applying it to a clinical setting. This section examines and analyzes the validation methods used for each aforementioned task in the literature, including the data and performance measures. We also focus on the evaluation of clinically relevant measures, besides the evaluation of computing accuracy of the algorithms.

4.1. Public AF related datasets

Several challenge events have been organized in recent years at international conferences such as ISBI (International Symposium on Biomedical Imaging) and MICCAI (Medical Image Computing and Computer-Assisted Interventions), with corresponding public datasets released. For example, Zhuang et al. organized the *Multi-Modality Whole Heart Segmentation Challenge*, in conjunction with STACOM'17 and MICCAI'17. They provided 120 multi-modality images covering a wide range of cardiac diseases, such as AF, myocardial infarction, and congenital heart disease (Zhuang et al., 2019). Ten algorithms for CT data and eleven methods for MRI data have been evaluated, and most of the submitted algorithms were DL-based. The evaluated results showed that the LA cavity segmentation of AF patients was particularly more accurate compared to other categories of patients. Moreover, public datasets were released along with the challenge events focusing on a specific anatomical structure instead of the whole heart. Table 11 summarizes the public AF-related events and datasets with corresponding download links.

For LA cavity segmentation, Tobon-Gomez et al. organized the *Left Atrium Segmentation Challenge*, in conjunction with STACOM'13 and MICCAI'13. They offered a dataset including 30 CT and 30 MRIs with the manual LA cavity segmentation and presented the results of nine algorithms for CT and eight for MRI (Tobon-Gomez et al., 2015). Their results showed that the methodologies that combined statistical models with region-growing were the most suitable for the target

task. Zhao et al. organized the *Atrial Segmentation Challenge*, in conjunction with STACOM'18 and MICCAI'18. They provided 150 LGE MRIs with manual LA cavity segmentation generated from three experts, and the data covered both pre- and post-ablation images (Xiong et al., 2020). To explore the quality of the dataset, they calculated three measures, i.e., SNR, CR, and heterogeneity, which were in agreement. The quality measurements showed that less than 15% of the data had high quality (SNR>3), 70% had medium quality (SNR = 1~3), and over 15% was of low quality (SNR<1). In total, 27 teams contributed to the automatic LA cavity segmentation, and most of the methods were DL-based except for two MAS methods. The results showed that two-stage CNNs achieved superior results than other single CNN methods and conventional methods. This challenge event provided a significant step towards much-improved segmentation methods for the LA cavity segmentation of LGE MRI.

For LA wall segmentation, Karim et al. organized the *Left Atrial Wall Thickness Challenge*, in conjunction with STACOM'16 and MICCAI'16. The released images consisted of 10 CT and 10 MRIs of healthy and diseased subjects with manual LA wall segmentation. Only two of the three participants contributed to the automatic segmentation of the CT data, but no work on the MRI data was reported (Karim et al., 2018). The limited number of submitted algorithms generally performed poorly compared to the inter-observer variability, which revealed the difficulty of the wall segmentation task. Zhao and Xiong (2018) and Utah (2012) released a public LGE MRI dataset with LA wall segmentation. This segmentation was however generated using the morphological (dilation) operation from the LA cavity manual segmentation.

For LA scar segmentation, Karim et al. organized the *Left Atrium Fibrosis and Scar Segmentation Challenge* at ISBI 2012. They provided 60 multi-center and multi-vendor LGE MRIs with manual labels of both LA and scars, and summarized the submitted algorithms from seven institutions in Karim et al. (2013). To the best of our knowledge, no public dataset for gap quantification and evaluation has been reported.

4.2. Evaluation measures

The methods are evaluated in different ways for different tasks in the literature. However, all the measures are generally designed based on the idea of comparing automatic segmentation results with reference segmentations. In this section, we

Table 11. Summary of the public datasets whose research targets are AF patients or include AF patients. Here, the star ($\star$) indicates that the data is acquired from multiple centers and vendors. bSSFP: balanced steady-state free precession.

Source	Year	Data	Target	Name & Hyperlink
Utah (2012)	2012	155 LGE MRI + 3D MRI	LA cavity, LA scar	*CARMA, University of Utah*
Karim et al. (2013)	2012	60 LGE MRI*	LA scar	*Left Atrium Fibrosis and Scar Segmentation Challenge*
Tobon-Gomez et al. (2015)	2013	30 CT, 30 bSSFP MRI	LA cavity	*Left Atrium Segmentation Challenge*
Karim et al. (2018)	2016	10 CT, 10 black-blood MRI	LA wall	*Left Atrial Wall Thickness Challenge*
Zhuang et al. (2019)	2017	60 CT, 60 bSSFP MRI*	Whole heart including LA cavity	*Multi-Modality Whole Heart Segmentation Challenge*
Xiong et al. (2020)	2018	150 LGE MRI	LA cavity	*Atrial Segmentation Challenge*

summarize common measures employed in each LA computing task. The reader is referred to Fig. 11 for an illustration of each evaluation measure listed below.

4.2.1. LA cavity measures

For assessing the performance of LA cavity segmentation, a range of different measures have been explored, as shown in Table 6. The most widely used measures include the Dice coefficient/ score, Jaccard index, HD, and average surface distance (ASD). They are defined as follows,

$$\text{Dice}(V_{\text{auto}}, V_{\text{manual}}) = \frac{2\,|V_{\text{auto}} \cap V_{\text{manual}}|}{|V_{\text{auto}}| + |V_{\text{manual}}|}, \tag{1}$$

$$\text{Jaccard}(V_{\text{auto}}, V_{\text{manual}}) = \frac{|V_{\text{auto}} \cap V_{\text{manual}}|}{|V_{\text{auto}} \cup V_{\text{manual}}|}, \tag{2}$$

$$\text{HD}(X, Y) = \max\left[\sup_{x \in X}\inf_{y \in Y} d(x, y), \sup_{y \in Y}\inf_{x \in X} d(x, y)\right], \tag{3}$$

and

$$\text{ASD}(X, Y) = \frac{1}{2}\left(\frac{\sum_{x \in X} \min_{y \in Y} d(x, y)}{\sum_{x \in X} 1} + \frac{\sum_{y \in Y} \min_{x \in X} d(x, y)}{\sum_{y \in Y} 1}\right), \tag{4}$$

where V_{manual} and V_{auto} denote the set of pixels in the manual and automatic segmentation, respectively; X and Y represent two sets of contour points; $d(x, y)$ indicates the Euclidean distance between the two points x and y; and $|\cdot|$ refers to the number of pixels in set V. Dice and Jaccard are selected for volumetric overlap measurement, where Jacquard index can be more sensible and severe upon small variation compared to Dice (Jamart et al., 2019). ASD and HD are used to evaluate the shape and contour accuracy of the object of interest. ASD calculates the average of the distances between all pairs of pixels between two surfaces. HD calculates the largest error distance of the 3D segmentation defined for a prediction of the target. Therefore, HD can further measure the existence of outliers, and sometimes 95% HD will be used to eliminate the influence of a small subset of outliers.

In addition, three statistical measurements are employed, i.e., Accuracy (Acc), Specificity (Spe), and Sensitivity (Sen), defined as follows,

$$\text{Acc} = \frac{TP + TN}{TP + FP + FN + TN}, \tag{5}$$

$$\text{Spe} = \frac{TN}{TN + FP}, \tag{6}$$

and

$$\text{Sen} = \frac{TP}{TP + FN}, \tag{7}$$

where TP, TN, FN, and FP stand for the number of true positives, true negatives, false negatives, and false positives, respectively. Acc represents the proportion of true results (both TP and TN) among the total number of cases examined. Spe and Sen are used to reflect the success of the algorithm for the foreground and the background segmentation, respectively. Besides, the diameter and volume error calculations are used to assess the medical relevance of the automatic reconstructed LA volumes in the clinic.

4.2.2. LA wall measures

For the LA wall segmentation, wall thickness and Dice are currently the most commonly used measures. The thickness (Tk) of the LA wall can be calculated by averaging the thickness over each pixel $p_i \in S_{\text{epi}}$ from the epicardium S_{epi} to the endocardium S_{endo}, and therefore is defined as,

$$\text{Tk} = \frac{\sum_{p_i \in S_{\text{epi}}} d(p_i, S_{\text{endo}})}{|S_{\text{epi}}|}. \tag{8}$$

Actually, when the object size is much smaller than the background (as in the case of the LA wall), overlap-based metrics based on the four overlap cardinalities (TP, TN, FP, FN) are generally inappropriate (Taha and Hanbury, 2015). This is because they will provide the same metric value, regardless of the distance between two non-overlapping regions evaluated, ultimately affecting the objectivity in precision. Therefore, both Dice and Jaccard are not suitable since they can also be represented as,

$$\text{Dice} = \frac{2TP}{2TP + FP + FN}, \tag{9}$$

$$\text{Jaccard} = \frac{TP}{TP + FP + FN}. \tag{10}$$

In this case, distance-based metrics are recommended, as they consider the precision and accuracy of both the shape and local alignment of segmented regions. Apart from its small size, the LA wall is also accompanied by adjacent PV structures, which also exhibit large inter-observer variation and could be regarded
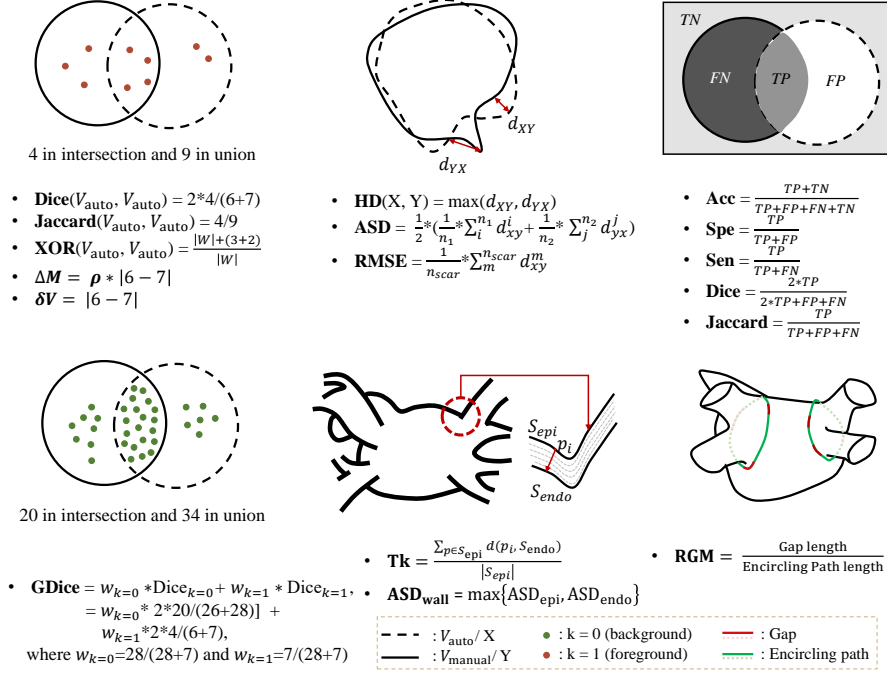
- **Dice**$(V_{auto}, V_{auto}) = 2*4/(6+7)$
- **Jaccard**$(V_{auto}, V_{auto}) = 4/9$
- **XOR**$(V_{auto}, V_{auto}) = \frac{|W|+(3+2)}{|W|}$
- $\Delta M = \rho * |6-7|$
- $\delta V = |6-7|$

- **HD**$(X, Y) = \max(d_{XY}, d_{YX})$
- **ASD** $= \frac{1}{2}*(\frac{1}{n_1}*\sum_i^{n_1} d_{xy}^i + \frac{1}{n_2}*\sum_j^{n_2} d_{yx}^j)$
- **RMSE** $= \frac{1}{n_{scar}}*\sum_m^{n_{scar}} d_{xy}^m$

- **Acc** $= \frac{TP+TN}{TP+FP+FN+TN}$
- **Spe** $= \frac{TP}{TP+FP}$
- **Sen** $= \frac{TP}{TP+FN}$
- **Dice** $= \frac{2*TP}{2*TP+FP+FN}$
- **Jaccard** $= \frac{TP}{TP+FP+FN}$

- **GDice** $= w_{k=0} *\text{Dice}_{k=0} + w_{k=1} * \text{Dice}_{k=1},$
  $= w_{k=0}* 2*20/(26+28)] +$
  $w_{k=1}*2*4/(6+7),$
  where $w_{k=0}=28/(28+7)$ and $w_{k=1}=7/(28+7)$

- **Tk** $= \frac{\sum_{p \in S_{epi}} d(p_i, S_{endo})}{|S_{epi}|}$
- **ASD$_{wall}$** $= \max\{ASD_{epi}, ASD_{endo}\}$

- **RGM** $= \frac{\text{Gap length}}{\text{Encircling Path length}}$

- - - : V_{auto}/ X   ● : k = 0 (background)   ——— : Gap
——— : V_{manual}/ Y   ● : k = 1 (foreground)   ••••••• : Encircling path

Fig. 11. Sketch map of measures employed in LA LGE MRI computing.

as outliers. Compared to HD which is sensitive to outliers, ASD is a better option for LA wall quantitative assessment. As the LA wall segmentation involves the two surfaces, i.e., the epicardium and endocardium, the ASD of the LA wall is defined as,

$$\text{ASD}_{wall} = \max\left\{ASD_{epi}, ASD_{endo}\right\}. \qquad (11)$$

Apart from these measurements, tissue mass and clinical evaluation are also employed for the evaluation of LA wall segmentation. The tissue mass M is designed to predict the volume error, and the difference in mass is defined as,

$$\Delta M = \rho \times |V - \hat{V}|, \qquad (12)$$

where ρ = 1.053 g/ml (Vinnakota and Bassingthwaighte, 2004) is the average wall tissue density, and V and $\hat{V}$ refer to the reference and predicted volume, respectively (Karim et al., 2018). Furthermore, Veni et al. (2017) proposed to compare the scar percentages within the manually and automatically segmented LA wall. The basic idea behind this is that the LA wall segmentation is usually regarded as an initial step for the scar segmentation as mentioned earlier.

4.2.3. LA scar measures

The optimal evaluation method to quantify scars from LGE MRI is still controversial due to the lack of ground truth. Currently, the EAM system is regarded as the clinical standard technique for the scar assessment, as presented in Fig. 12. The widely used bipolar voltage threshold defining the LA scars is ≤ 0.05 mV, which has been propagated through the literature and clinical practice (Harrison et al., 2014). However, the correlation between the LA scars identified by LGE MRI (enhanced regions) and EAM (low voltage regions) is still being questioned (Floria et al., 2020). The subjective and inaccurate scar segmentation might be one of the main reasons.

Alternatively, most algorithms employ manual segmented LA scars as the ground truth. For this evaluation, volume overlap measures and scar percentage are commonly used, as Table 8 shows. For example, Perry et al. (2012) proposed a novel overlap measure for the scar evaluation, namely XOR overlap,

$$\text{XOR}(V_{auto}, V_{manual}) = \frac{|W| + |V_{auto} \oplus V_{manual}|}{|W|}, \qquad (13)$$

where $|W|$ is the set of voxels that belong to the LA wall, and $\oplus$ refers to exclusive OR. The XOR overlap measure emphasizes the difference between overlapping scars, and will not be affected by the size of scars.

However, as mentioned in Section 4.2.2 volume overlap measures (such as Dice) could be highly sensitive to the mismatch of small structures (namely scars here), so in instances it will impose disproportionate penalties on the algorithm. To mitigate the effect of the small size of scars, Li et al. (2020b) proposed to project the appearance of scars onto the LA surface for both ground truth and automatic segmentation results, and then calculate the Dice scores of scars on the projected LA surface instead of on the 3D volume (Wu et al., 2018; Li et al., 2018b, 2020b,a). Furthermore, Li et al. (2020b,a) computed the generalized Dice (GDice) of scars from the projected LA surface for a better interpretation. GDice is defined as follows,

$$\text{GDice} = \frac{2 \sum_{k=0}^{N_k-1} \left| S_k^{auto} \cap S_k^{manual} \right|}{\sum_{k=0}^{N_k-1} \left(\left| S_k^{auto} \right| + \left| S_k^{manual} \right| \right)}, \qquad (14)$$

where S_k^{auto} and S_k^{manual} indicate the segmentation results of label k from the automatic method and manual delineation on the LA surface, respectively, N_k is the number of labels. Here, $N_k = 2$, where $k = 1$ represents normal wall and $k = 1$ refers to scarring regions.
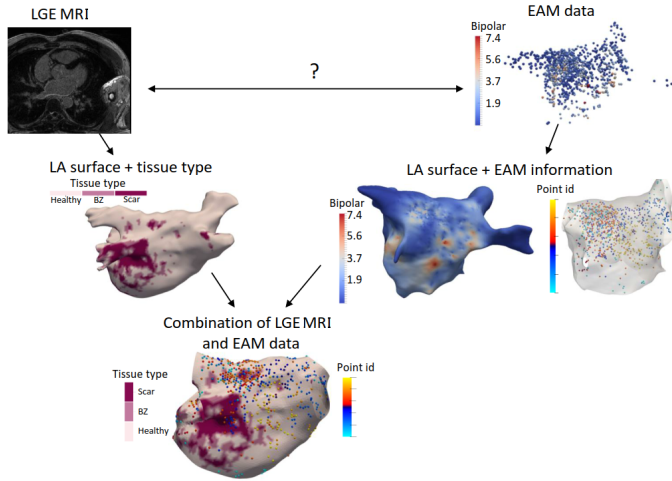
Fig. 12. The spatial correspondence of LGE MRI and EAM data. Image adapted from Núñez García et al. (2018) with permission.

Karim et al. (2013) proposed a surface-based metric, which employed MIP to calculate the distance error between the mesh vertex points on the LA surface. The distance error is defined as the root mean squared error (RMSE), i.e.,

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} d\left(v_i^{auto}, v_i^{manual}\right)^2}, \quad (15)$$

where v_i^{auto} and v_i^{manual} are the set of mesh vertices belonging to scars from the prediction and ground truth, respectively. The major limitation of the surface based metric is that targets with a significant amount of *FP* scars will have a low RMSE error. Nevertheless, it can be overcome by combining the surface measure with a volume-based index.

Scar percentage is directly related to clinical categorization of AF patients, as presented in Table 1, and thus should be appropriate as an assessment measure. Besides, one could analyze the relationship of scar percentages between manually and automatic scar segmentations, to evaluate the performance of automatic scar segmentation. For example, Veni et al. (2017) quantified the scar percentage correlation using the mean square error (MSE) and R-square value. Many works also calculate the volume error of scars for evaluation, which is defined as,

$$\delta V = |V_{auto} - V_{manual}|. \quad (16)$$

Statistical measurements related to scar classification could be employed for evaluation, including Acc, Sen, Spe, receiver operating characteristic (ROC) curve, and balanced error rate (BER).

4.2.4. LA ablation gap measures

As Table 9 shows, most gap quantification methods in the literature employed ablation gap characteristics (i.e., number, length, and position of gaps) for evaluation. Similar to the evaluation of scars, these works also analyzed the correlation with EAM, by comparing the ablation gaps in LGE MRI to the electrical gaps in EAM. However, the applicability of EAM for ablation gap quantification is limited. This is mainly because: 1)

the difficulty of the gap position registration between LGE MRI and EAM; 2) the voltage mapping does not entirely reflect scar/ gap formation; 3) the requirement of a voltage threshold for scar/ gap classification, with the same issues as for the LGE MRI threshold. Therefore, direct extrapolation of EAM data to verify LGE MRI should be performed carefully, in particular when they offer contradictory information (Nuñez-Garcia et al., 2019). Besides, Ranjan et al. (2012) calculated the correlation between gap length (GL) measured via LGE MRI for evaluation. Nuñez-Garcia et al. (2019) proposed a quantitative index, i.e., relative gap measure (RGM), to calculate the proportion of the ablation gaps on a defined standard LA parcellation.

$$RGM = \frac{\text{Gap length}}{\text{Encircling Path length}}, \quad (17)$$

where "Gap length" indicates the sum of all GLs along the "Encircling Path", and the "Encircling Path length" refers to the length of the complete closed-loop on the PVs. The RGM is between 0 and 1, which means that if RGM = 0, the vein is completely surrounded, and if RGM = 1, there are no scars around the veins. To alleviate the effect of the scar segmentation, one could adopt a multi-threshold scheme for the scar segmentation, and then integrate the results into the RGM calculation (Nuñez-Garcia et al., 2019).

4.3. Evaluation results on the AF-related public dataset

In general, the segmentation accuracy of different methods is not directly comparable, unless these methods are evaluated on the same dataset using the protocols. Therefore, we only summarize the state-of-the-art results of reviewed LA LGE MRI computing methods on the public dataset here, as presented in Table 12.

For LA cavity segmentation, three public datasets are available, and Dice, ASD and HD are commonly used for evaluation. On the dataset from Utah (2012), the state-of-the-art results of the LA cavity segmentation in all metrics were from Zhu et al. (2013). On the dataset from Zhuang et al. (2019), mean Dice scores from different methods have been reported for each pathology including AF. The methods evaluated on the public dataset from Xiong et al. (2020) have been separated into conventional methods and DL-based methods. For each metric, we list the state-of-the-art results from conventional and DL-based methods, and the best Dice score for the LA cavity segmentation was obtained by Xiong et al. (2018) (Dice = 0.942±0.014). The DL-based methods demonstrated great potential, as the best result in each metric was all obtained by DL-based methods on this dataset.

For LA wall segmentation, there is only one available public dataset for evaluation, which included 10 CT and 10 nonenhanced MRIs instead of the LGE MRI. We present both the state-of-the-art results and inter-observer variations for each metric. One can see that the results based on semi-automatic algorithms were generally comparable to the inter-observer variations for each metric. However, the size of this dataset is small, and current semi-automatic methods are labor-intensive and subjective.

Table 12. Summary of representative results for LA LGE MRI computing on public AF-related datasets. CoV: coefficient of variation; CV: conventional methods.

Public dataset source	Target	Representative result
Utah (2012)	LA cavity	Dice = 0.79 ± 0.05, OV = 0.65 ± 0.07, ASD = 2.79 ± 2.84 mm, 95% HD = 14.4 ± 3.65 mm (Zhu et al., 2013)
Xiong et al. (2020)	LA cavity	Dice = 0.861 ± 0.036 (CV) (Qiao et al., 2018), 0.942 ± 0.014 (DL) (Xiong et al., 2018) Jc = 0.758 (CV) (Qiao et al., 2018), 0.874 (DL) (Xia et al., 2018) HD = 11.8 mm (CV) (Qiao et al., 2018), 8.60 mm (DL) (Chen et al., 2018a) ASD = 1.473 mm (CV) (Nuñez-Garcia et al., 2018), 0.748 mm (DL) (Xia et al., 2018) Sen = 0.847 (CV) (Qiao et al., 2018), 0.949 (DL) (Preetha et al., 2018) Spe = 0.999 (CV) (Qiao et al., 2018), 1.000 (DL) (Vesal et al., 2018)
Zhuang et al. (2019)	Whole heart including LA cavity	Dice = 0.844 ± 0.097 (only on AF patients)
Karim et al. (2018)	LA wall	Dice = 0.43 (A) (Tao et al., 2016b), 0.39 (P) (Inoue and Drangova, 2016), 0.67 ± 0.22 (Inter-ob) on the CT; 0.72, 0.56 ± 0.14 (Inter-ob) on the MRI Tk error = 0 (P/A), 0.25 mm (P) and 0.20 mm (A) (Inter-ob) Tissue mass error = 3.84∼14.63 g, 10.03 ± 4.0 g (Inter-ob)
Karim et al. (2013)	LA scar	Dice = 0.85^{post}, 0.48^{pre} RMSE = $6.34 ± 8.2^{post}$ mm, $0.17 ± 0.1^{pre}$ mm (Lu et al., 2012) $\delta V = 6.34 ± 8.2^{post}$ ml, $1.25 ± 1.5^{pre}$ ml
Utah (2012)	LA scar	Dice = 0.807 ± 0.106 (semi-auto) (Perry et al., 2012) XOR = 0.916 ± 0.035 (semi-auto) (Perry et al., 2012) CoV = 0.62 (Andalò et al., 2018)

For LA scar segmentation, two public datasets are accessible, and typically Dice is used for evaluation. On both datasets, only semi-automatic algorithms were applied. There was performance variation among pre- and post-ablation images from Karim et al. (2013). Specifically, the best Dice scores were 0.48 and 0.85 on pre- and post-ablation LGE MRIs, respectively. However, in terms of RMSE and δV, the performance on the pre-ablation LGE MRIs was better than that on the post-ablation LGE MRI. The possible reason could be that the volume of post-ablation LGE MRI is generally larger than that of pre-ablation image. Nevertheless, pre-ablation LGE MRI is still generally more challenging for fibrosis segmentation due to its more diffuse distributions.

5. Potential clinical applications of the developed algorithms

It is essential to evaluate the clinical utility of the developed approaches for AF. Instead of blindly improving the accuracy of methods, researchers therefore can focus more on answering some clinical questions related to AF. The exploration and understanding of potential clinical applications of AF can guide the development of segmentation and quantification algorithms and answer important clinical questions. For example, we can employ the developed segmentation and quantification techniques to compare native and ablation-induced scars (Section 5.1), inspect the regional distribution of wall thickness (Section 5.2), fibrosis/ scars and ablation gaps from LGE MRI, and analyze the relationship between fibrosis/ scars/ gaps and AF recurrence (Section 5.3). Moreover, there are several other clinical applications, such as analyzing the relationship between the low-voltage regions in EAM and scars detected by LGE MRI, the relationship between ablation parameters (power of the radiofrequency signal, catheter contact force, etc.) and the created

chronic lesion detected by LGE MRI, as well as assessing the reproducibility of LGE MRI scar imaging with respect to imaging parameters. However, the latter three applications require additional EAM data or LGE MRIs with different ablation and imaging parameters, and therefore are out of the scope of this review.

To the best of our knowledge, there are a limited number of review papers targeting the clinical applications of LGE MRI. Zghaib and Nazarian (2018) summarized the new insights into the use of MRIs for the decision-making of AF management. They explored LGE, native T1-weighed, T2-weighted as well as cine MRI, and for LGE MRI they only reviewed studies on the relationship between the extent of scars on post-ablation LGE MRIs and the rate of AF recurrence. In this section, we will provide a comprehensive review from the perspective of the clinical applications for AF analysis.

5.1. Comparisons of native and ablation-induced scars

Recent studies demonstrated the differences in the extent and distribution of fibrosis/ scars of pre-/ post-ablation LGE MRI (Malcolme-Lawes et al., 2013; Fukumoto et al., 2015). For instance, Malcolme-Lawes et al. (2013) found that there was no difference of scars between ostial and LA cavity regions for pre-ablation data, but in post-ablation data the extent of scars in the ostia is larger than that in the LA cavity. They also reported a positive association between the extent of preexisting fibrosis and AF recurrence, which coincides with the finding in the literature (Verma et al., 2005b; Mahnkopf et al., 2010). However, they did not find any relationship between the amount of ablation-induced scars and AF recurrence, which should be negatively associated according to the studies of Peters et al. (2009); McGann et al. (2011). Fukumoto et al. (2015) demonstrated that ablation-induced scars are related to greater contrast affinity and thinner walls compared to preexisting fibro-

sis. Yang et al. (2017a) tried to distinguish native and ablation-induced scars via a texture based feature extraction. They stated the difficulty of the differentiation between native and ablation-induced scars, especially for longstanding persistent AF. Therefore, the understanding of the characteristics of pre- vs. post-ablation scars can be important and may inform future ablation strategies for AF.

5.2. Regional distribution analysis of wall thickness and fibrosis/ scars

To date, there are already several studies on LA wall thickness measurements, to analyze the relationships between wall thickness and patient age, AF stage/ type, scar formation, and AF recurrence (Karim et al., 2018). For example, Hall et al. (2006) studied 34 patients of different ages and found that the thinnest and thickest areas were the roof (1.06 ± 1.49 mm) and septum (2.2 ± 0.82 mm), respectively. They did not find any significant relationships between the wall thickness and age. In contrast, Pan et al. (2008) measured the wall thickness on 180 AF patients of various ages and concluded that the thickness increased with age. They also found that the anterior wall (2.0 ± 0.9 mm, 3.2 ± 0.2 mm and 3.7 ± 0.9 mm in 40~60, 60~80 and 80+ year olds) was thicker than the posterior wall (0.7 ± 0.2 mm, 1.8 ± 0.2 mm and 2.4 ± 0.4 mm in 40~60, 60~80 and 80+ year olds) among all the age groups. Beinart et al. (2011) and Hayashi et al. (2014) both observed that the middle superior posterior wall was the thinnest region with a thickness of 1.43 ± 0.44 mm and 1.44 ± 0.17 mm, respectively. Suenari et al. (2013) analyzed the thickness of 54 AF patients, and showed that the thickest wall area is in the left lateral ridge (4.42 ± 1.28 mm), while the thinnest is in the LIPV (1.68 ± 0.27 mm). Besides, they found that the thickness of the left lateral ridge was correlated to the AF recurrence (p=0.041). However, the superior right posterior wall was found to be significantly associated with both AF recurrence (p=0.048) and electrical reconnection (p=0.014) in Inoue et al. (2016). Despite this progress, most of these works were based on manual segmented the LA wall, and focused on CT images instead of LGE MRI. Note that transmural lesion formation is critical to the success of AF ablation and is dependent on the knowledge of regional LA wall thickness. Therefore, the distribution analysis of wall thickness from LGE MRI could be important and might provide insight into the progress of the AF.

As for the regional distribution of fibrosis/ scars in the LA LGE MRI, related information is limited and has not been comprehensively reported. Cochet et al. (2015) divided the LA into four segments and reported an irregular fibrosis anatomical distribution. However, they found that fibrosis generally occurred more often on the posterior LA wall than the anterior one, particularly in the area adjacent to and below LIPV. Benito et al. (2018) manually defined the LA parcellation with 12 sub-regions: 1~4, posterior wall; 5~6, floor; 7, septal wall; 8~11, anterior wall; 12, lateral wall (see Fig. 13 (a)). They selected 76 consecutive AF patients for analysis and also observed that the fibrosis was preferentially located at the posterior wall and floor around the antrum of the LIPV, i.e., segments 3 and 5 (40.42% and 25.82% fibrosis), as Fig. 13 (b) shows. In con-

trast, segments 8 and 10 (2.54% and 3.82% fibrosis) in the anterior wall contained the fewest fibrosis. Similar to the increased wall thickness in Pan et al. (2008), they found that age (>60 years old) was also significantly correlated to increased fibrosis (p=0.04). Recently, (Lee et al., 2019) separated the LA into nine segments, and also found that scars were most frequently seen at the posterior wall around the LIPV. Besides, they studied 195 paroxysmal and 121 persistent AF patients and observed that the presence of fibrosis assessed in LIPV from LGE MRI was associated with the chronicity of AF. This preliminary research suggests that the knowledge of preferential fibrosis/ scar position may open further perspectives in ablation strategies, patient selection, and AF recurrence prediction.

5.3. Relationship analysis between fibrosis/ scars/ gaps and AF recurrence

As mentioned in Section 5.1, both the extent of preexisting fibrosis and ablation-induced scars are correlated with AF recurrence, but with opposite effects. Specifically, AF recurrence is positively associated with the extent of preexisting scars, but negatively related to that of post-ablation scars. The characteristics of pre- vs. post-ablation scars may explain the seemingly paradox and inform future strategies for ablation (Fukumoto et al., 2015). With respect to the pre-ablation scars (also namely fibrosis), it has been regarded as a potential cause of the abnormalities in atrial activation, which may underlie the initiation and maintenance of AF. Note that AF belongs to a progressive disease, and several studies revealed that causality between AF and fibrosis may be bidirectional (Oakes et al., 2009). This might explain why patients with a greater extent of fibrosis normally suffer much higher recurrence rates after ablation. Apart from the extent of fibrosis, Oakes et al. (2009) investigated 81 AF patients with pre-ablation LGE MRI, and found that AF recurrence was also related to the locations of fibrosis. In their experiments, patients with recurrent AF presented fibrosis on the whole LA, whereas patients without recurrent AF had fibrosis only located primarily to the posterior wall and septum. As for post-ablation scars, robust evidence supports that complete circumferential and transmural lesion formation is critical to successful AF ablation (Cappato et al., 2003; Verma et al., 2005a; Ouyang et al., 2005). Here, the ablation lesion just refers to the post-ablation scars or can be named ablation-induced scars. Therefore, patients with a smaller degree of post-ablation scars on LGE MRI tend to recur AF after ablation. Similar to fibrosis, the location of post-ablation scars is also an important index for AF recurrence prediction. For example, several studies emphasized the importance of right inferior PV (RIPV) scars, which is the most highly correlated to clinical ablation success (Yamada et al., 2006; Peters et al., 2009). This could attribute to the reported technical difficulty in ablating the RIPV region due to poor catheter access, resulting in its greater variability of scars. For example, Peters et al. (2009) studied 35 AF patients undergoing the first ablation procedure, and compared the extent of scars on different sub-regions. They demonstrated that the PVs of patients without recurrence had more completely circumferential scars, especially on RIPV regions. In the case of ablation gaps, which are generally caused by incomplete PVI, the extent
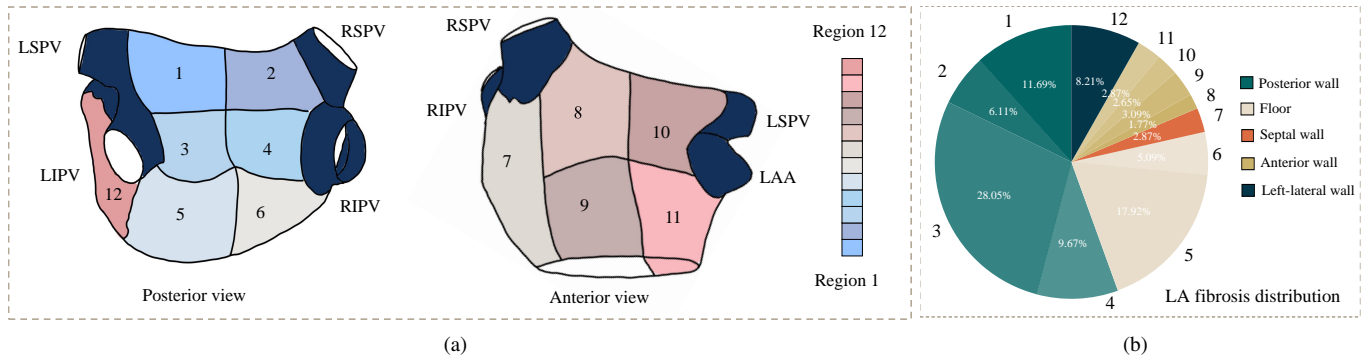
Fig. 13. Example of a LA parcellation and its corresponding fibrosis distribution: (a) the LA surface template parcellated from anatomical landmarks; (b) the regional distribution of LA fibrosis (the same color with different transparencies refers to the same region category), where the values were obtained from Benito et al. (2018). Illustrations designed referring to Benito et al. (2018).

and distribution of gaps are regarded to be positively associated with AF recurrence. The identification and localization of ablation gaps from LGE MRI have been used to predict AF recurrence and further guide repeated PVI procedures (Bisbal et al., 2014).

6. Discussion and future perspectives

LGE MRI has attracted increasing attention in the assessment of AF before and after an ablation procedure. Automatic segmentation and quantification algorithms of LA structures and tissues can facilitate the diagnosis and therapy of AF patients. However, the translation of current algorithms into the clinical environment remains challenging. In this section, we summarize existing major challenges in the field of LA LGE MRI computing and the solutions recently proposed. The exploration of these challenges and related works is expected to provide useful information for developing novel methods and applications for AF analysis.

6.1. Surface projection and LA unfolding mapping

Recent studies have shown that the success of AF treatment highly relies on the formation of contiguous and transmural scars on the LA wall (Glover et al., 2018). However, the wall thickness is difficult to measure based on current LGE MRI techniques. In clinical practice, the location and extent of scars are believed to have greater clinical significance and can be used to predict outcomes of AF ablation procedures (Arujuna et al., 2012). Therefore, several studies have been proposed to project scars onto the LA surface to perform scar quantification (Ravanelli et al., 2014; Tao et al., 2016a; Li et al., 2020b, 2021c). Fig. 14 (a) presents an example of scar projection achieved by MIP. By projection, the errors due to LA wall thickness can be mitigated, and the computational complexity of algorithms can be drastically reduced.

Nevertheless, the cross-subject comparison of 3D surface data is still arduous. To solve this, (Roney et al., 2019) developed a universal atrial coordinate mapping system for 2D visualization of both the LA and right atrium. Williams et al. (2017) created a 2D LA standardized unfolding mapping (LA-SUM) template where the MV was mapped to a disk, the PVs

to circles, and the LAA to an ellipse, as presented in Fig. 14 (b). The target 3D LA will be registered to a 3D template and then transferred to the 2D template via a 3D-2D template mapping. The LA flattening of LA-SUM may result in undesired information loss between 3D and 2D LA representations due to the possible inaccurate registration between LA surfaces with high shape variability. Instead of relying on a 3D registration step, (Nuñez-Garcia et al., 2020) proposed a quasi-conformal LA flatting scheme and employed additional regional constraints to overcome undesired mesh self-folding. The advantages of these LA unfolding mapping techniques include 2D visualization, LA regional assessment, and multi-modal data combination. However, their templates were generally designed for the most common LA topology with four PVs. We therefore expect that more flexible templates can be developed to adapt for the LA topological variants.

6.2. Joint optimization and independent analysis of the AF-related tasks

The target regions of the four tasks reviewed in Section 3 are all inherently related, particularly in the spatial information of images, as shown in Fig. 2. Several studies employed multi-task learning for simultaneous LA cavity segmentation and scar segmentation/ quantification and proved the effectiveness of joint optimization (Chen et al., 2018b; Li et al., 2020a). The spatial information between the LA cavity and scars could simply be learned via spatial attention, i.e., multiplying the LA cavity feature map by the scar feature map (Chen et al., 2018b), or projecting the scars onto the LA endocardial surface (Li et al., 2020a). At the same time, several studies have been devoted to reducing the correlation between the accuracy of related tasks in LA LGE MRI computing, i.e., their conditional dependencies. For instance, MIP schemes have been widely used in LA scar and gap quantification to mitigate the effect of inaccurate LA cavity segmentation (Knowles et al., 2010; Tao et al., 2016a; Razeghi et al., 2020; Bisbal et al., 2014). Patch shift scheme was developed to apply a random shift along the LA boundary when performing surface projection (Li et al., 2020b). Li et al. (2020a) learned the spatial information around the LA boundary to reduce the dependence on accurate LA cavity segmentation. Despite these advances, the joint optimization and inde-

Maximal intensity projection
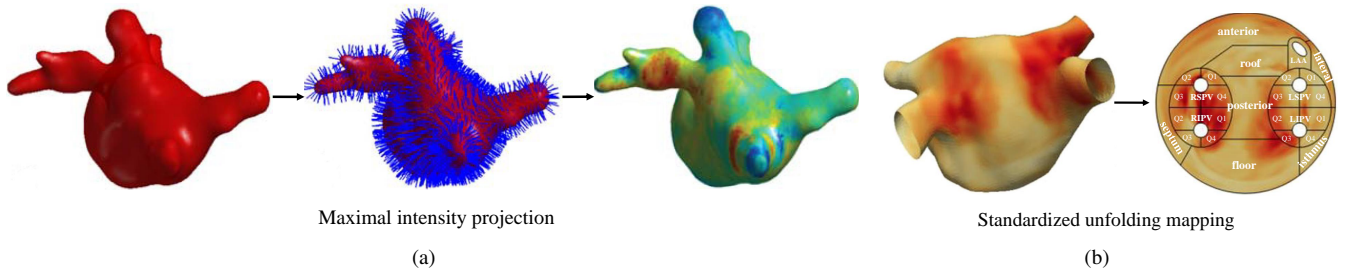
(a)

Standardized unfolding mapping

(b)

Fig. 14. Alternative visualizations and representations of LA scars via: (a) maximum intensity projection (images adapted from Tao et al. (2016a) with permission); (b) LA standardized unfolding mapping (images adapted from Williams et al. (2017) with permission).

pendence analysis of the AF-related tasks are yet to be explored in further depth in the future.

6.3. Challenges with deep learning in LA LGE MRI computing

It is evident that DL-based methods have obtained promising results on the LA cavity and scar segmentation and quantification. It is mainly attributed to the release of related public datasets and the emerge of advanced network architectures. With the release of public datasets, the research on the LA cavity and scar segmentation from LGE MRI started to increase, as Fig. 4 shows. Despite the promising results, deep neural networks still confront a number of challenges, such as poor interpretability, scarcity of annotated data, class imbalance problems, limited domain generalization ability, and catastrophic forgetting. One may refer to the review papers (Chen et al., 2020; Hesamian et al., 2019) to follow these challenges and state-of-the-art solutions for DL-based medical image segmentation. Here, we mainly discuss the limited data (Section 6.3.1) and model generalization issues (Section 6.3.2), as there exist several unique points in the two challenges for AF studies.

6.3.1. Scarcity of (annotated) data

The scarcity of (annotated) data is a serious issue in LA LGE MRI computing. Though this is common in many other tasks, LGE imaging could be more challenging, due to the existence of contrast enhancement, its complex patterns, and the large quality and contrast variations across different patients. Especially, LGE MRI of LA wall requires substantially higher spatial resolution, patient-specific optimization of scan parameters, strict criteria for contrast dosage and delay between contrast injection and image acquisition, compared to LGE MRI of the LV (Siebermair et al., 2017; Chubb et al., 2018). These precise requirements are difficult to meet in practice, resulting in scarcity and poor image quality of LGE MRI. It is also complicated to collect many annotated cases of 3D LGE MRI. However, DL-based LA LGE MRI computing typically relies on a large number of annotated samples for training. Several schemes have been proposed to solve this. For example, Yu et al. (2019) employed a semi-supervised learning method for the LA cavity segmentation from LGE MRI, to fully utilize the unlabeled data. Li et al. (2020b) adopted a patch-wise training for the LA scar quantification from LGE MRI, which considerably increased the amount of labeled training data. Data argumentation is generally useful in deep learning with limited

training data, for example the method of partially region rotation of scars was employed for LV segmentation from LGE MRI (Campello et al., 2019). Unsupervised domain adaptation has also been proven to be capable to alleviate the problem of limited annotated data from the target domain, which has been widely used for LV LGE MRI segmentation (Zhuang et al., 2020; Wu and Zhuang, 2020, 2021; Pei et al., 2021). Finally, the methods making full use of sparse annotation (Çiçek et al., 2016) are promising for LA LGE MRI computing with limited annotated data and could be further explored in the future.

6.3.2. Limited domain generalization ability

Currently, most existing algorithms have only been evaluated on center- and vendor-specific LGE MRI. Though the *Left Atrium Fibrosis and Scar Segmentation Challenge* offered multi-center and multi-scanner data, the benchmark algorithms only tested on center- and vendor-specific images. Their suitability and performance had not been tested on data from other centers or vendors (Karim et al., 2013). Note that LGE MRIs from different centers can vary evidently in appearance, as Fig. 15 shows. This is mainly due to the absence of standardized LGE MRI acquisition protocols, leading to poor reproducibility of LGE MRI (Benito et al., 2017; Sim et al., 2019). Even in the same dataset, one could encounter a severe data mismatch problem, resulting in poor outlier results (Li et al., 2021c). Several schemes have been employed to solve this, such as data augmentation/ generation, domain-invariant representation learning, and meta-learning (Wang et al., 2021). Nevertheless, large multi-center and multi-scanner datasets are needed to validate the robustness and generalizability of current methods, which is more useful in practice. It is also worthy of promoting deep models with efficient inherent generalization abilities for the LGE MRI data processing from different centers and vendors (Li et al., 2021b). Moreover, it could be interesting to study the domain shift between pre- and post-ablation LGE MRIs from the same center, and the label variations of LGE MRIs from different centers.

7. Conclusion

We have presented and discussed the current progress of LGE MRI computing for LA studies, particularly for the four tasks, including segmentation and (or) quantification of LA cavity,
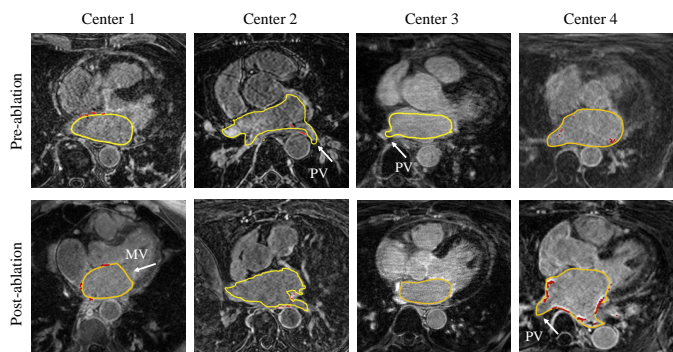
Fig. 15. Multi-center pre- and post-ablation LGE MRIs, where LA cavity contour is labeled in the yellow and scarring region is labeled in red. The images differ in contrast, enhancement as well as background, and the labels across different centers also exist variations, especially in the MV and PV regions.

wall, scars, and ablation gaps. Though LGE MRI has been proven to be a powerful diagnostic and prognostic tool in the study of AF, a standardized imaging protocol should be further investigated. Furthermore, a limited number of works have been reported focusing on image computing tasks, especially for automatic LA wall segmentation and ablation gap quantification. Most research relies on manual delineation for further analysis and clinical applications. Therefore, more accurate and robust automatic methods are desired for overall wide and intelligent use in the clinical setting. The data-driven approaches have shown great potential for the LA cavity and scar segmentation and quantification, thanks to the development of deep neural networks. The joint optimization of these related tasks can be a new direction for the utilization of their spatial relationship. To research for a broader clinical application, well-controlled and large-cohort studies are expected to better guarantee the reproducibility of measurements, refine the evaluation methods, and validate the impact on clinical outcomes as well as the computing accuracy.

Although we limit our survey related to AF analysis in the article, the described methodologies can be useful to other clinical applications. We described in detail the characteristics of targets, which motivated the methodologies. Consequently, such methods can be used for other targets sharing similar characteristics as the targets in AF studies. For instance, tumor lesions are also small and diffuse targets, so the review on the scar segmentation and quantification methods could inspire the development of methods on tumor lesion segmentation, and vice versa. We believe that this review has the potential to help researchers to design appropriate frameworks according to their problems and be aware of similar challenging issues and state-of-the-art solutions.

Acknowledgment

References

Akoum, N., Daccarett, M., McGann, C., Segerson, N., Vergara, G., Kuppahally, S., Badger, T., Burgon, N., Haslam, T., Kholmovski, E., et al., 2011. Atrial fibrosis helps select the appropriate patient and strategy in catheter ablation of atrial fibrillation: A DE-MRI guided approach. Journal of Cardiovascular Electrophysiology 22, 16–22.

Akoum, N., Wilber, D., Hindricks, G., Jais, P., Cates, J., Marchlinski, F., Kholmovski, E., Burgon, N., Hu, N., Mont, L., et al., 2015. MRI assessment of ablation-induced scarring in atrial fibrillation: analysis from the DECAAF study. Journal of Cardiovascular Electrophysiology 26, 473–480.

Ambellan, F., Tack, A., Ehlke, M., Zachow, S., 2019. Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the osteoarthritis initiative. Medical Image Analysis 52, 109–118.

Andalò, A., Fabbri, C., Valinoti, M., Masci, A., Corsi, C., 2018. Quantification of left atrium fibrosis from LGE MRI in atrial fibrillation, in: 2018 Computing in Cardiology Conference (CinC), IEEE. pp. 1–4.

Arujuna, A., Karim, R., Caulfield, D., Knowles, B., Rhode, K., Schaeffter, T., Kato, B., Rinaldi, C.A., Cooklin, M., Razavi, R., et al., 2012. Acute pulmonary vein isolation is achieved by a combination of reversible and irreversible atrial injury after catheter ablation: evidence from magnetic resonance imaging. Circulation: Arrhythmia and Electrophysiology 5, 691–700.

Badger, T.J., Daccarett, M., Akoum, N.W., Adjei-Poku, Y.A., Burgon, N.S., Haslam, T.S., Kalvaitis, S., Kuppahally, S., Vergara, G., McMullen, L., et al., 2010. Evaluation of left atrial lesions after initial and repeat atrial fibrillation ablation: lessons learned from delayed-enhancement MRI in repeat ablation procedures. Circulation: Arrhythmia and Electrophysiology 3, 249–259.

Beinart, R., Abbara, S., Blum, A., Ferencik, M., Heist, K., Ruskin, J., Mansour, M., 2011. Left atrial wall thickness variability measured by CT scans in patients undergoing pulmonary vein isolation. Journal of Cardiovascular Electrophysiology 22, 1232–1236.

Beinart, R., Khurram, I.M., Liu, S., Yarmohammadi, H., Halperin, H.R., Bluemke, D.A., Gai, N., Van Der Geest, R.J., Lima, J.A., Calkins, H., et al., 2013. Cardiac magnetic resonance t1 mapping of left atrial myocardium. Heart Rhythm 10, 1325–1331.

Benito, E.M., Cabanelas, N., Nuñez-Garcia, M., Alarcón, F., Figueras I Ventura, R.M., Soto-Iglesias, D., Guasch, E., Prat-Gonzalez, S., Perea, R.J., Borràs, R., et al., 2018. Preferential regional distribution of atrial fibrosis in posterior wall around left inferior pulmonary vein as identified by late gadolinium enhancement cardiac magnetic resonance in patients with atrial fibrillation. Ep Europace 20, 1959–1965.

Benito, E.M., Carlosena-Remirez, A., Guasch, E., Prat-González, S., Perea, R.J., Figueras, R., Borràs, R., Andreu, D., Arbelo, E., Tolosana, J.M., et al., 2017. Left atrial fibrosis quantification by late gadolinium-enhanced magnetic resonance: a new method to standardize the thresholds for reproducibility. Ep Europace 19, 1272–1279.

Berruezo, A., Tamborero, D., Mont, L., Benito, B., Tolosana, J.M., Sitges, M., Vidal, B., Arriagada, G., Méndez, F., Matiello, M., et al., 2007. Preprocedural predictors of atrial fibrillation recurrence after circumferential pulmonary vein ablation. European Heart Journal 28, 836–841.

Bhagirath, P., van der Graaf, A., Karim, R., van Driel, V., Ramanna, H., Rhode, K., de Groot, N., Götte, M., 2014. Multimodality imaging for patient evaluation and guidance of catheter ablation for atrial fibrillation—current status and future perspective. International Journal of Cardiology 175, 400–408.

Bian, C., Yang, X., Ma, J., Zheng, S., Liu, Y.A., Nezafat, R., Heng, P.A., Zheng, Y., 2018. Pyramid network with online hard example mining for accurate left atrium segmentation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 237–245.

Bisbal, F., Guiu, E., Cabanas-Grandío, P., Berruezo, A., Prat-Gonzalez, S., Vidal, B., Garrido, C., Andreu, D., Fernandez-Armenta, J., Tolosana, J.M.,

et al., 2014. Cmr-guided approach to localize and ablate gaps in repeat AF ablation procedure. JACC: Cardiovascular Imaging 7, 653–663.

Bishop, M., Rajani, R., Plank, G., Gaddum, N., Carr-White, G., Wright, M., O'Neill, M., Niederer, S., 2016. Three-dimensional atrial wall thickness maps to inform catheter ablation procedures for atrial fibrillation. Europace 18, 376–383.

Blondheim, D.S., Osipov, A., Meisel, S.R., Frimerman, A., Shochat, M., Shotan, A., 2005. Relation of left atrial size to function as determined by transesophageal echocardiography. The American Journal of Cardiology 96, 457–463.

Borra, D., Andalò, A., Paci, M., Fabbri, C., Corsi, C., 2020. A fully automated left atrium segmentation approach from late gadolinium enhanced magnetic resonance imaging based on a convolutional neural network. Quantitative Imaging in Medicine and Surgery 10, 1894.

Borra, D., Masci, A., Esposito, L., Andalò, A., Fabbri, C., Corsi, C., 2018. A semantic-wise convolutional neural network approach for 3-D left atrium segmentation from late gadolinium enhanced magnetic resonance imaging, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 329–338.

Bourier, F., Reents, T., Ammar-Busch, S., Buiatti, A., Kottmaier, M., Semmler, V., Telishevska, M., Brkic, A., Grebmer, C., Lennerz, C., et al., 2016. Evaluation of a new very low dose imaging protocol: feasibility and impact on x-ray dose levels in electrophysiology procedures. Ep Europace 18, 1406–1410.

Calkins, H., Brugada, J., Packer, D.L., Cappato, R., Chen, S.A., Crijns, H.J., Damiano Jr, R.J., Davies, D.W., Haines, D.E., Haissaguerre, M., et al., 2007. HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: Recommendations for personnel, policy, procedures and follow-up: A report of the heart rhythm society (HRS) task force on catheter and surgical ablation of atrial fibrillation developed in partnership with the European heart rhythm association (EHRA) and the European cardiac arrhythmia society (ECAS); in collaboration with the American college of cardiology (ACC), American heart association (AHA), and the society of thoracic surgeons (STS). endorsed and approved by the governing bodies of the American college of cardiology, the American heart association, the European cardiac arrhythmia society, the European heart rhythm association, the society of thoracic surgeons, and the heart rhythm society. Europace 9, 335–379.

Cameli, M., Lisi, M., Righini, F.M., Mondillo, S., 2012. Novel echocardiographic techniques to assess left atrial size, anatomy and function. Cardiovascular Ultrasound 10, 1–14.

Campello, V.M., Gkontra, P., Izquierdo, C., Martín-Isla, C., Sojoudi, A., Full, P.M., Maier-Hein, K., Zhang, Y., He, Z., Ma, J., et al., 2021. Multi-centre, multi-vendor and multi-disease cardiac segmentation: The M&Ms challenge. IEEE Transactions on Medical Imaging .

Campello, V.M., Martín-Isla, C., Izquierdo, C., Petersen, S.E., Ballester, M.A.G., Lekadir, K., 2019. Combining multi-sequence and synthetic images for improved segmentation of late gadolinium enhancement cardiac MRI, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 290–299.

Cappato, R., Negroni, S., Pecora, D., Bentivegna, S., Lupo, P.P., Carolei, A., Esposito, C., Furlanello, F., De Ambroggi, L., 2003. Prospective assessment of late conduction recurrence across radiofrequency lesions producing electrical disconnection at the pulmonary vein ostium in patients with atrial fibrillation. Circulation 108, 1599–1604.

Casaclang-Verzosa, G., Gersh, B.J., Tsang, T.S., 2008. Structural and functional remodeling of the left atrium: clinical and therapeutic implications for atrial fibrillation. Journal of the American College of Cardiology 51, 1–11.

Chen, C., Bai, W., Rueckert, D., 2018a. Multi-task learning for left atrial segmentation on GE-MRI, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 292–301.

Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W., Rueckert, D., 2020. Deep learning for cardiac image segmentation: A review. Frontiers in Cardiovascular Medicine 7, 25.

Chen, J., Yang, G., Gao, Z., Ni, H., Angelini, E., Mohiaddin, R., Wong, T., Zhang, Y., Du, X., Zhang, H., et al., 2018b. Multiview two-task recursive attention model for left atrium and atrial scars segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 455–463.

Chubb, H., Karim, R., Roujol, S., Nuñez-Garcia, M., Williams, S.E., Whitaker, J., Harrison, J., Butakoff, C., Camara, O., Chiribiri, A., et al., 2018. The re-

producibility of late gadolinium enhancement cardiovascular magnetic resonance imaging of post-ablation atrial scar: a cross-over study. Journal of Cardiovascular Magnetic Resonance 20, 21.

Chugh, S.S., Havmoeller, R., Narayanan, K., Singh, D., Rienstra, M., Benjamin, E.J., Gillum, R.F., Kim, Y.H., McAnulty Jr, J.H., Zheng, Z.J., et al., 2014. Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 study. Circulation 129, 837–847.

Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 424–432.

Cochet, H., Mouries, A., Nivet, H., Sacher, F., Derval, N., Denis, A., Merle, M., Relan, J., Hocini, M., Haissaguerre, M., et al., 2015. Age, atrial fibrillation, and structural heart disease are the main determinants of left atrial fibrosis detected by delayed-enhanced magnetic resonance imaging in a general cardiology population. Journal of Cardiovascular Electrophysiology 26, 484–492.

Cox, J.L., 2003. Surgical treatment of atrial fibrillation: a review. EP Europace 5, S20–S29.

Deng, C., Zhang, X., 2016. Automatic segmentation of the left atrium from mr images via semantic information, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE. pp. 003312–003316.

Ding, W., Li, L., Zhuang, X., Huang, L., 2020. Cross-modality multi-atlas segmentation using deep neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 233–242.

Du, X., Yin, S., Tang, R., Liu, Y., Song, Y., Zhang, Y., Liu, H., Li, S., 2020. Segmentation and visualization of left atrium through a unified deep learning framework. International Journal of Computer Assisted Radiology and Surgery , 1–12.

Dzeshka, M.S., Lip, G.Y., Snezhitskiy, V., Shantsila, E., 2015. Cardiac fibrosis in patients with atrial fibrillation: mechanisms and clinical implications. Journal of the American College of Cardiology 66, 943–959.

Ellis, K., Ziada, K.M., Vivekananthan, D., Latif, A.A., Shaaraoui, M., Martin, D., Grimm, R.A., 2006. Transthoracic echocardiographic predictors of left atrial appendage thrombus. The American Journal of Cardiology 97, 421–425.

Floria, M., Radu, S., Gosav, E.M., Cozma, D., Mitu, O., Ouatu, A., Tanase, D.M., Scripcariu, V., Serban, L.I., 2020. Left atrial structural remodelling in non-valvular atrial fibrillation: what have we learnt from CMR? Diagnostics 10, 137.

Fukumoto, K., Habibi, M., Ipek, E.G., Khurram, I.M., Zimmerman, S.L., Zipunnikov, V., Spragg, D.D., Ashikaga, H., Rickard, J., Marine, J.E., et al., 2015. Comparison of preexisting and ablation-induced late gadolinium enhancement on left atrial magnetic resonance imaging. Heart Rhythm 12, 668–672.

Galand, V., Pavin, D., Behar, N., Auffret, V., Fénéon, D., Behaghel, A., Daubert, J.C., Mabo, P., Martins, R.P., 2016. Localization of gaps during redo ablations of paroxysmal atrial fibrillation: preferential patterns depending on the choice of cryoballoon ablation or radiofrequency ablation for the initial procedure. Archives of Cardiovascular Diseases 109, 591–598.

Gao, Y., Gholami, B., MacLeod, R.S., Blauer, J., Haddad, W.M., Tannenbaum, A.R., 2010. Segmentation of the endocardial wall of the left atrium using local region-based active contours and statistical shape learning, in: Medical Imaging 2010: Image Processing, International Society for Optics and Photonics. p. 76234Z.

Glover, B.M., Hong, K.L., Baranchuk, A., Bakker, D., Chacko, S., Bisleri, G., 2018. Preserved left atrial epicardial conduction in regions of endocardial "isolation". JACC: Clinical Electrophysiology 4, 557–558.

Gottlieb, I., Pinheiro, A., Brinker, J.A., Corretti, M.C., Mayer, S.A., Bluemke, D.A., Lima, J.A., Marine, J.E., Berger, R.D., Calkins, H., et al., 2008. Diagnostic accuracy of arterial phase 64-slice multidetector CT angiography for left atrial appendage thrombus in patients undergoing atrial fibrillation ablation. Journal of Cardiovascular Electrophysiology 19, 247–251.

Habibi, M., Lima, J.A., Khurram, I.M., Zimmerman, S.L., Zipunnikov, V., Fukumoto, K., Spragg, D., Ashikaga, H., Rickard, J., Marine, J.E., et al., 2015. Association of left atrial function and left atrial enhancement in patients with atrial fibrillation: cardiac magnetic resonance study. Circulation: Cardiovascular Imaging 8, e002769.

Habijan, M., Babin, D., Galić, I., Leventić, H., Romić, K., Velicki, L., Pižurica, A., 2020. Overview of the whole heart and heart chamber segmentation methods. Cardiovascular Engineering and Technology , 1–23.

Hall, B., Jeevanantham, V., Simon, R., Filippone, J., Vorobiof, G., Daubert, J., 2006. Variation in left atrial transmural wall thickness at sites commonly targeted for ablation of atrial fibrillation. Journal of Interventional Cardiac Electrophysiology 17, 127–132.

Harrison, J.L., Jensen, H.K., Peel, S.A., Chiribiri, A., Grøndal, A.K., Bloch, L.Ø., Pedersen, S.F., Bentzon, J.F., Kolbitsch, C., Karim, R., et al., 2014. Cardiac magnetic resonance and electroanatomical mapping of acute and chronic atrial ablation injury: a histological validation study. European Heart Journal 35, 1486–1495.

Harrison, J.L., Sohns, C., Linton, N.W., Karim, R., Williams, S.E., Rhode, K.S., Gill, J., Cooklin, M., Rinaldi, C.A., Wright, M., et al., 2015a. Repeat left atrial catheter ablation: cardiac magnetic resonance prediction of endocardial voltage and gaps in ablation lesion sets. Circulation: Arrhythmia and Electrophysiology 8, 270–278.

Harrison, J.L., Whitaker, J., Chubb, H., Williams, S.E., Wright, M., Razavi, R.S., O'Neill, M.D., 2015b. Advances in CMR of post-ablation atrial injury. Current Cardiovascular Imaging Reports 8, 22.

Hayashi, H., Hayashi, M., Miyauchi, Y., Takahashi, K., Uetake, S., Tsuboi, I., Yodogawa, K., Iwasaki, Y.K., Shimizu, W., 2014. Left atrial wall thickness and outcomes of catheter ablation for atrial fibrillation in patients with hypertrophic cardiomyopathy. Journal of Interventional Cardiac Electrophysiology 40, 153–160.

Hesamian, M.H., Jia, W., He, X., Kennedy, P., 2019. Deep learning techniques for medical image segmentation: achievements and challenges. Journal of Digital Imaging 32, 582–596.

Higuchi, K., Cates, J., Gardner, G., Morris, A., Burgon, N.S., Akoum, N., Marrouche, N.F., 2018. The spatial distribution of late gadolinium enhancement of left atrial magnetic resonance imaging in patients with atrial fibrillation. JACC: Clinical Electrophysiology 4, 49–58.

Holmes, D.R., Monahan, K.H., Packer, D., 2009. Pulmonary vein stenosis complicating ablation for atrial fibrillation: clinical spectrum and interventional considerations. JACC: Cardiovascular Interventions 2, 267–276.

Hsing, J., Peters, D.C., Knowles, B.R., Manning, W.J., Josephson, M.E., 2014. Cardiovascular magnetic resonance imaging of scar development following pulmonary vein isolation: a prospective study. PloS one 9, e104844.

Hunter, R.J., Jones, D.A., Boubertakh, R., MALCOLME-LAWES, L.C., Kanagaratnam, P., Juli, C.F., Davies, D.W., Peters, N.S., Baker, V., Earley, M.J., et al., 2013. Diagnostic accuracy of cardiac magnetic resonance imaging in the detection and characterization of left atrial catheter ablation lesions: a multicenter experience. Journal of Cardiovascular Electrophysiology 24, 396–403.

Inoue, J., Drangova, M., 2016. Left atrial wall segmentation using clinically correlated metrics, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 201–210.

Inoue, J., Skanes, A.C., Gula, L.J., Drangova, M., 2016. Effect of left atrial wall thickness on radiofrequency ablation success. Journal of Cardiovascular Electrophysiology 27, 1298–1303.

Inoue, J., Skanes, A.C., White, J.A., Rajchl, M., Drangova, M., 2014. Patient-specific left atrial wall-thickness measurement and visualization for radiofrequency ablation, in: Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling, International Society for Optics and Photonics. p. 90361N.

Jamart, K., Xiong, Z., Talou, G.D.M., Stiles, M.K., Zhao, J., 2020. Mini review: Deep learning for atrial segmentation from late gadolinium-enhanced MRIs. Frontiers in Cardiovascular Medicine 7.

Jamart, K., Xiong, Z., Talou, G.M., Stiles, M.K., Zhao, J., 2019. Two-stage 2D CNN for automatic atrial segmentation from LGE-MRIs, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 81–89.

Jannin, P., Krupinski, E., Warfield, S.K., 2006. Validation in medical image processing. IEEE Transactions on Medical Imaging 25, 1405–9.

Jia, S., Cadour, L., Cochet, H., Sermesant, M., 2016. Stacom-slawt challenge: left atrial wall segmentation and thickness measurement using region growing and marker-controlled geodesic active contour, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 211–219.

Jia, S., Despinasse, A., Wang, Z., Delingette, H., Pennec, X., Jaïs, P., Cochet, H., Sermesant, M., 2018. Automatically segmenting the left atrium from cardiac images using successive 3D U-nets and a contour loss, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 221–229.

Jongbloed, M., Schalij, M., Zeppenfeld, K., Oemrawsingh, P., van der Wall,

E., Bax, J., 2005a. Clinical applications of intracardiac echocardiography in interventional procedures. Heart 91, 981–990.

Jongbloed, M.R., Bax, J.J., Lamb, H.J., Dirksen, M.S., Zeppenfeld, K., van der Wall, E.E., de Roos, A., Schalij, M.J., 2005b. Multislice computed tomography versus intracardiac echocardiography to evaluate the pulmonary veins before radiofrequency catheter ablation of atrial fibrillation: a head-to-head comparison. Journal of the American College of Cardiology 45, 343–350.

Karim, R., Arujuna, A., Brazier, A., Gill, J., Rinaldi, C.A., O'Neill, M., Razavi, R., Schaeffter, T., Rueckert, D., Rhode, K.S., 2011. Automatic segmentation of left atrial scar from delayed-enhancement magnetic resonance imaging, in: International Conference on Functional Imaging and Modeling of the Heart, Springer. pp. 63–70.

Karim, R., Arujuna, A., Housden, R.J., Gill, J., Cliffe, H., Matharu, K., Gill, J., Rindaldi, C.A., O'Neill, M., Rueckert, D., et al., 2014. A method to standardize quantification of left atrial scar from delayed-enhancement MR images. IEEE Journal of Translational Engineering in Health and Medicine 2, 1–15.

Karim, R., Blake, L.E., Inoue, J., Tao, Q., Jia, S., Housden, R.J., Bhagirath, P., Duval, J.L., Varela, M., Behar, J.M., et al., 2018. Algorithms for left atrial wall segmentation and thickness–evaluation on an open-source CT and MRI image database. Medical Image Analysis 50, 36–53.

Karim, R., Housden, R.J., Balasubramaniam, M., Chen, Z., Perry, D., Uddin, A., Al-Beyatti, Y., Palkhi, E., Acheampong, P., Obom, S., et al., 2013. 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. Journal of Cardiovascular Magnetic Resonance 15, 105.

Knowles, B.R., Caulfield, D., Cooklin, M., Rinaldi, C.A., Gill, J., Bostock, J., Razavi, R., Schaeffter, T., Rhode, K.S., 2010. 3-D visualization of acute rf ablation lesions using MRI for the simultaneous determination of the patterns of necrosis and edema. IEEE Transactions on Biomedical Engineering 57, 1467–1475.

Kuchynka, P., Podzimkova, J., Masek, M., Lambert, L., Cerny, V., Danek, B., Palecek, T., 2015. The role of magnetic resonance imaging and cardiac computed tomography in the assessment of left atrial anatomy, size, and function. BioMed Research International 2015.

Kutra, D., Saalbach, A., Lehmann, H., Groth, A., Dries, S.P., Krueger, M.W., Dössel, O., Weese, J., 2012. Automatic multi-model-based segmentation of the left atrium in cardiac MRI scans, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 1–8.

Lee, D.K., Shim, J., Choi, J.i., Kim, Y.H., Oh, Y.W., Hwang, S.H., 2019. Left atrial fibrosis assessed with cardiac MRI in patients with paroxysmal and those with persistent atrial fibrillation. Radiology 292, 575–582.

Li, C., Tong, Q., Liao, X., Si, W., Sun, Y., Wang, Q., Heng, P.A., 2018a. Attention based hierarchical aggregation network for 3D left atrial segmentation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 255–264.

Li, L., Weng, X., Schnabel, J.A., Zhuang, X., 2020a. Joint left atrial segmentation and scar quantification based on a DNN with spatial encoding and shape attention, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 118–127.

Li, L., Wu, F., Yang, G., Xu, L., Wong, T., Mohiaddin, R., Firmin, D., Keegan, J., Zhuang, X., 2020b. Atrial scar quantification via multi-scale CNN in the graph-cuts framework. Medical Image Analysis 60, 101595.

Li, L., Yang, G., Wu, F., Wong, T., Mohiaddin, R., Firmin, D., Keegan, J., Xu, L., Zhuang, X., 2018b. Atrial scar segmentation via potential learning in the graph-cut framework, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 152–160.

Li, L., Zimmer, V.A., Ding, W., Wu, F., Huang, L., Schnabel, J.A., Zhuang, X., 2021a. Random style transfer based domain generalization networks integrating shape and spatial information, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 208–218.

Li, L., Zimmer, V.A., Schnabel, J.A., Zhuang, X., 2021b. AtrialGeneral: Domain generalization for left atrial segmentation of multi-center LGE MRIs, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 557–566.

Li, L., Zimmer, V.A., Schnabel, J.A., Zhuang, X., 2021c. AtrialJSQnet: A new framework for joint segmentation and quantification of left atrium and scars incorporating spatial and shape information. Medical Image Analysis , 102303.

Linhart, M., Alarcon, F., Borras, R., Benito, E.M., Chipa, F., Cozzari, J., Caixal, G., Enomoto, N., Carlosena, A., Guasch, E., et al., 2018. Delayed gadolinium enhancement magnetic resonance imaging detected anatomic gap length in wide circumferential pulmonary vein ablation lesions is associated with recurrence of atrial fibrillation. Circulation: Arrhythmia and Electrophysiology 11, e006659.

Liu, Y., Dai, Y., Yan, C., Wang, K., 2018. Deep learning based method for left atrial segmentation in GE-MRI, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 311–318.

Lu, Y., Yang, Y., Connelly, K.A., Wright, G.A., Radau, P.E., 2012. Automated quantification of myocardial infarction using graph cuts on contrast delayed enhanced magnetic resonance images. Quantitative Imaging in Medicine and Surgery 2, 81.

Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., Yang, X., Martel, A.L., 2021. Loss odyssey in medical image segmentation. Medical Image Analysis 11, 102035.

Mahnkopf, C., Badger, T.J., Burgon, N.S., Daccarett, M., Haslam, T.S., Badger, C.T., McGann, C.J., Akoum, N., Kholmovski, E., Macleod, R.S., et al., 2010. Evaluation of the left atrial substrate in patients with lone atrial fibrillation using delayed-enhanced MRI: implications for disease progression and response to catheter ablation. Heart Rhythm 7, 1475–1481.

Malcolme-Lawes, L., Juli, C., Karim, R., Bai, W., Quest, R., Lim, P.B., Jamil-Copley, S., Kojodjojo, P., Ariff, B., Davies, D., et al., 2013. Automated analysis of atrial late gadolinium enhancement imaging that correlates with endocardial voltage and clinical outcomes: a 2-center study. Heart Rhythm 10, 1184–1191.

Marrouche, N.F., Wilber, D., Hindricks, G., Jais, P., Akoum, N., Marchlinski, F., Kholmovski, E., Burgon, N., Hu, N., Mont, L., et al., 2014. Association of atrial tissue fibrosis identified by delayed enhancement MRI and atrial fibrillation catheter ablation: the DECAAF study. JAMA 311, 498–506.

McGann, C., Akoum, N., Patel, A., Kholmovski, E., Revelo, P., Damal, K., Wilson, B., Cates, J., Harrison, A., Ranjan, R., et al., 2014. Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI. Circulation: Arrhythmia and Electrophysiology 7, 23–30.

McGann, C., Kholmovski, E., Blauer, J., Vijayakumar, S., Haslam, T., Cates, J., DiBella, E., Burgon, N., Wilson, B., Alexander, A., et al., 2011. Dark regions of no-reflow on late gadolinium enhancement magnetic resonance imaging result in scar formation after atrial fibrillation ablation. Journal of the American College of Cardiology 58, 177–185.

McGann, C.J., Kholmovski, E.G., Oakes, R.S., Blauer, J.J., Daccarett, M., Segerson, N., Airey, K.J., Akoum, N., Fish, E., Badger, T.J., et al., 2008. New magnetic resonance imaging-based method for defining the extent of left atrial wall injury after the ablation of atrial fibrillation. Journal of the American College of Cardiology 52, 1263–1271.

Miller, M.A., d'Avila, A., Dukkipati, S.R., Koruth, J.S., Viles-Gonzalez, J., Napolitano, C., Eggert, C., Fischer, A., Gomes, J.A., Reddy, V.Y., 2012. Acute electrical isolation is a necessary but insufficient endpoint for achieving durable PV isolation: the importance of closing the visual gap. Europace 14, 653–660.

Mishima, T., Miyamoto, K., Morita, Y., Kamakura, T., Nakajima, K., Yamagata, K., Wada, M., Ishibashi, K., Inoue, Y., Nagase, S., et al., 2019. Feasibility of late gadolinium enhancement magnetic resonance imaging to detect ablation lesion gaps in patients undergoing cryoballoon ablation of paroxysmal atrial fibrillation. Journal of Arrhythmia 35, 190–196.

Mohrs, O.K., Nowak, B., Petersen, S.E., Welsner, M., Rubel, C., Magedanz, A., Kauczor, H.U., Voigtlaender, T., 2006. Thrombus detection in the left atrial appendage using contrast-enhanced MRI: a pilot study. American Journal of Roentgenology 186, 198–205.

Nakamura, K., Funabashi, N., Uehara, M., Ueda, M., Murayama, T., Takaoka, H., Komuro, I., 2011. Left atrial wall thickness in paroxysmal atrial fibrillation by multislice-CT is initial marker of structural remodeling and predictor of transition from paroxysmal to chronic form. International Journal of Cardiology 148, 139–147.

Ngo, T.A., Lu, Z., Carneiro, G., 2017. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Medical Image Analysis 35, 159–171.

Njoku, A., Kannabhiran, M., Arora, R., Reddy, P., Gopinathannair, R., Lakkireddy, D., Dominic, P., 2018. Left atrial volume predicts atrial fibrillation recurrence after radiofrequency ablation: a meta-analysis. Ep Europace 20, 33–42.

Nuñez-Garcia, M., Bernardino, G., Alarcón, F., Caixal, G., Mont, L., Camara, O., Butakoff, C., 2020. Fast quasi-conformal regional flattening of the left atrium. IEEE Transactions on Visualization and Computer Graphics 26, 2591–2602.

Nuñez-Garcia, M., Camara, O., O'Neill, M.D., Razavi, R., Chubb, H., Butakoff, C., 2019. Mind the gap: quantification of incomplete ablation patterns after pulmonary vein isolation using minimum path search. Medical Image Analysis 51, 1–12.

Nuñez-Garcia, M., Zhuang, X., Sanroma, G., Li, L., Xu, L., Butakoff, C., Camara, O., 2018. Left atrial segmentation combining multi-atlas whole heart labeling and shape-based atlas selection, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 302–310.

Núñez García, M., et al., 2018. Left atrial parameterisation and multi-modal data analysis: application to atrial fibrillation. Ph.D. thesis. Universitat Pompeu Fabra.

Oakes, R.S., Badger, T.J., Kholmovski, E.G., Akoum, N., Burgon, N.S., Fish, E.N., Blauer, J.J., Rao, S.N., DiBella, E.V., Segerson, N.M., et al., 2009. Detection and quantification of left atrial structural remodeling using delayed enhancement MRI in patients with atrial fibrillation. Circulation 119, 1758.

Obeng-Gyimah, E., Nazarian, S., 2020. Advancements in imaging for atrial fibrillation ablation: Is there a potential to improve procedural outcomes? The Journal of Innovations in Cardiac Rhythm Management 11, 4172.

Ouyang, F., Antz, M., Ernst, S., Hachiya, H., Mavrakis, H., Deger, F.T., Schaumann, A., Chun, J., Falk, P., Hennig, D., et al., 2005. Recovered pulmonary vein conduction as a dominant factor for recurrent atrial tachyarrhythmias after complete circular isolation of the pulmonary veins: lessons from double lasso technique. Circulation 111, 127–135.

O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Riordan, D., Walsh, J., 2019. Deep learning vs. traditional computer vision, in: Science and Information Conference, Springer. pp. 128–144.

Pan, N.H., Tsao, H.M., Chang, N.C., Chen, Y.J., Chen, S.A., 2008. Aging dilates atrium and pulmonary veins: implications for the genesis of atrial fibrillation. Chest 133, 190–196.

Pathan, F., Hecht, H., Narula, J., Marwick, T.H., 2018. Roles of transesophageal echocardiography and cardiac computed tomography for evaluation of left atrial thrombus and associated pathology: a review and critical analysis. JACC: Cardiovascular Imaging 11, 616–627.

Pei, C., Wu, F., Huang, L., Zhuang, X., 2021. Disentangle domain features for cross-modality cardiac image segmentation. Medical Image Analysis 71, 102078.

Peng, P., Lekadir, K., Gooya, A., Shao, L., Petersen, S.E., Frangi, A.F., 2016. A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. Magnetic Resonance Materials in Physics, Biology and Medicine 29, 155–195.

Perry, D., Morris, A., Burgon, N., McGann, C., MacLeod, R., Cates, J., 2012. Automatic classification of scar tissue in late gadolinium enhancement cardiac MRI for the assessment of left-atrial wall injury after radiofrequency ablation, in: Medical Imaging 2012: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 83151D.

Peters, D.C., Wylie, J.V., Hauser, T.H., Nezafat, R., Han, Y., Woo, J.J., Taclas, J., Kissinger, K.V., Goddu, B., Josephson, M.E., et al., 2009. Recurrence of atrial fibrillation correlates with the extent of post-procedural late gadolinium enhancement: a pilot study. JACC: Cardiovascular Imaging 2, 308–316.

Polaczek, M., Szaro, P., Baranska, I., Burakowska, B., Ciszek, B., 2019. Morphology and morphometry of pulmonary veins and the left atrium in multislice computed tomography. Surgical and Radiologic Anatomy 41, 721–730.

Pontecorboli, G., Figueras i Ventura, R.M., Carlosena, A., Benito, E., Prat-Gonzales, S., Padeletti, L., Mont, L., 2017. Use of delayed-enhancement magnetic resonance imaging for fibrosis detection in the atria: a review. EP Europace 19, 180–189.

Prasanna, L., Praveena, R., AS, D., Kumar, M., 2014. Variations in the pulmonary venous ostium in the left atrium and its clinical importance. Journal of Clinical and Diagnostic Research: JCDR 8, 10.

Preetha, C.J., Haridasan, S., Abdi, V., Engelhardt, S., 2018. Segmentation of the left atrium from 3D gadolinium-enhanced mr images with convolutional neural networks, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 265–272.

Puybareau, É., Zhao, Z., Khoudli, Y., Carlinet, E., Xu, Y., Lacotte, J., Géraud, T., 2018. Left atrial segmentation in a few seconds using fully convolutional network and transfer learning, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 339–347.

Qiao, M., Wang, Y., van der Geest, R.J., Tao, Q., 2018. Fully automated left

atrium cavity segmentation from 3D GE-MRI by multi-atlas selection and registration, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 230–236.

Ranjan, R., Kato, R., Zviman, M.M., Dickfeld, T.M., Roguin, A., Berger, R.D., Tomaselli, G.F., Halperin, H.R., 2011. Gaps in the ablation line as a potential cause of recovery from electrical isolation and their visualization using MRI. Circulation: Arrhythmia and Electrophysiology 4, 279–286.

Ranjan, R., Kholmovski, E.G., Blauer, J., Vijayakumar, S., Volland, N.A., Salama, M.E., Parker, D.L., MacLeod, R., Marrouche, N.F., 2012. Identification and acute targeting of gaps in atrial ablation lesion sets using a real-time magnetic resonance imaging system. Circulation: Arrhythmia and Electrophysiology 5, 1130–1135.

Ravanelli, D., dal Piaz, E.C., Centonze, M., Casagranda, G., Marini, M., Del Greco, M., Karim, R., Rhode, K., Valentini, A., 2014. A novel skeleton based quantification and 3-D volumetric visualization of left atrium fibrosis using late gadolinium enhancement magnetic resonance imaging. IEEE Transactions on Medical Imaging 33, 566–576.

Razeghi, O., Sim, I., Roney, C.H., Karim, R., Chubb, H., Whitaker, J., O'Neill, L., Mukherjee, R., Wright, M., O'neill, M., et al., 2020. Fully automatic atrial fibrosis assessment using a multilabel convolutional neural network. Circulation: Cardiovascular Imaging 13, e011512.

Regazzoli, D., Ancona, F., Trevisi, N., Guarracini, F., Radinovic, A., Oppizzi, M., Marzi, A., Sora, N.C., Della Bella, P., Mazzone, P., et al., 2015. Left atrial appendage: physiology, pathology, and role as a therapeutic target. BioMed Research International 2015.

Rhode, K., Karim, R., 2012. ISBI 2012: Left atrium fibrosis and scar segmentation challenge. http://www.cardiacatlas.org/challenges/left-atrium-fibrosis-and-scar-segmentation-challenge/.

Rolf, S., Hindricks, G., Sommer, P., Richter, S., Arya, A., Bollmann, A., Kosiuk, J., Koutalas, E., 2014. Electroanatomical mapping of atrial fibrillation: Review of the current techniques and advances. Journal of Atrial Fibrillation 7.

Romero, J., Husain, S.A., Kelesidis, I., Sanz, J., Medina, H.M., Garcia, M.J., 2013. Detection of left atrial appendage thrombus by cardiac computed tomography in patients with atrial fibrillation: a meta-analysis. Circulation: Cardiovascular Imaging 6, 185–194.

Roney, C.H., Beach, M.L., Mehta, A.M., Sim, I., Corrado, C., Bendikas, R., Solis-Lemus, J.A., Razeghi, O., Whitaker, J., O'Neill, L., et al., 2020. In silico comparison of left atrial ablation techniques that target the anatomical, structural, and electrical substrates of atrial fibrillation. Frontiers in Physiology 11, 1145.

Roney, C.H., Pashaei, A., Meo, M., Dubois, R., Boyle, P.M., Trayanova, N.A., Cochet, H., Niederer, S.A., Vigmond, E.J., 2019. Universal atrial coordinates applied to visualisation, registration and construction of patient specific meshes. Medical Image Analysis 55, 65–75.

Saad, E.B., Cole, C.R., Marrouche, N.F., Dresing, T.J., PEREZ-LUGONES, A., Saliba, W.I., Schweikert, R.A., Klein, A., Rodriguez, L., Grimm, R., et al., 2002. Use of intracardiac echocardiography for prediction of chronic pulmonary vein stenosis after ablation of atrial fibrillation. Journal of Cardiovascular Electrophysiology 13, 986–989.

Utah, 2012. Cardiac MRI data from the comprehensive arrhythmia research and management (CARMA) center at the University of Utah. http://insight-journal.org/midas/%20collection/view/197/.

Savioli, N., Montana, G., Lamata, P., 2018. V-fcnn: volumetric fully convolution neural network for automatic atrial segmentation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 273–281.

Sibley, C.T., Noureldin, R.A., Gai, N., Nacif, M.S., Liu, S., Turkbey, E.B., Mudd, J.O., Van Der Geest, R.J., Lima, J.A., Halushka, M.K., et al., 2012. T1 mapping in cardiomyopathy at cardiac MR: comparison with endomyocardial biopsy. Radiology 265, 724–732.

Siebermair, J., Kholmovski, E.G., Marrouche, N., 2017. Assessment of left atrial fibrosis by late gadolinium enhancement magnetic resonance imaging: methodology and clinical implications. JACC: Clinical Electrophysiology 3, 791–802.

Sim, I., Razeghi, O., Karim, R., Chubb, H., Whitaker, J., O'Neill, L., Mukherjee, R.K., Roney, C.H., Razavi, R., Wright, M., et al., 2019. Reproducibility of atrial fibrosis assessment using CMR imaging and an open source platform. JACC: Cardiovascular Imaging 12, 2076–2077.

Spragg, D.D., Khurram, I., Zimmerman, S.L., Yarmohammadi, H., Barcelon, B., Needleman, M., Edwards, D., Marine, J.E., Calkins, H., Nazarian, S., 2012. Initial experience with magnetic resonance imaging of atrial scar and co-registration with electroanatomic voltage mapping during atrial fibrillation: success and limitations. Heart Rhythm 9, 2003–2009.

Suenari, K., Nakano, Y., Hirai, Y., Ogi, H., Oda, N., Makita, Y., Ueda, S., Kaji-hara, K., Tokuyama, T., Motoda, C., et al., 2013. Left atrial thickness under the catheter ablation lines in patients with paroxysmal atrial fibrillation: insights from 64-slice multidetector computed tomography. Heart and Vessels 28, 360–368.

Taclas, J.E., Nezafat, R., Wylie, J.V., Josephson, M.E., Hsing, J., Manning, W.J., Peters, D.C., 2010. Relationship between intended sites of rf ablation and post-procedural scar in AF patients, using late gadolinium enhancement cardiovascular magnetic resonance. Heart Rhythm 7, 489–496.

Taha, A.A., Hanbury, A., 2015. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Medical Imaging 15, 29.

Tao, Q., Ipek, E.G., Shahzad, R., Berendsen, F.F., Nazarian, S., van der Geest, R.J., 2016a. Fully automatic segmentation of left atrium and pulmonary veins in late gadolinium-enhanced MRI: Towards objective atrial scar assessment. Journal of Magnetic Resonance Imaging 44, 346–354.

Tao, Q., Shahzad, R., Berendsen, F.F., van der Geest, R.J., 2016b. Automatic left atrial wall segmentation from contrast-enhanced CT angiography images, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 220–227.

Taylor, A.J., Salerno, M., Dharmakumar, R., Jerosch-Herold, M., 2016. T1 mapping: basic techniques and clinical applications. JACC: Cardiovascular Imaging 9, 67–81.

Thiagalingam, A., Manzke, R., D'AVILA, A., Ho, I., Locke, A.H., Ruskin, J.N., Chan, R.C., Reddy, V.Y., 2008. Intraprocedural volume imaging of the left atrium and pulmonary veins with rotational X-ray angiography: implications for catheter ablation of atrial fibrillation. Journal of Cardiovascular Electrophysiology 19, 293–300.

Tobon-Gomez, C., Geers, A.J., Peters, J., Weese, J., Pinto, K., Karim, R., Ammar, M., Daoudi, A., Margeta, J., Sandoval, Z., et al., 2015. Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. IEEE Transactions on Medical Imaging 34, 1460–1473.

Toffanin, G., Scarabeo, V., Verlato, R., De Conti, F., Zampiero, A.A., Piovesana, P., 2006. Transoesophageal echocardiographic evaluation of pulmonary vein anatomy in patients undergoing ostial radiofrequency catheter ablation for atrial fibrillation: a comparison with magnetic resonance angiography. Journal of Cardiovascular Medicine 7, 748–752.

Tops, L.F., Schalij, M.J., Bax, J.J., 2010. Imaging and atrial fibrillation: the role of multimodality imaging in patient evaluation and management of atrial fibrillation. European Heart Journal 31, 542–551.

Tops, L.F., van der Wall, E.E., Schalij, M.J., Bax, J.J., 2007. Multi-modality imaging to assess left atrial size, anatomy and function. Heart 93, 1461–1470.

Toupin, S., Pezel, T., Bustin, A., Cochet, H., 2021. Whole-heart high-resolution late gadolinium enhancement: Techniques and clinical applications. Journal of Magnetic Resonance Imaging: JMRI , 1053–1807.

Tsao, H., Wu, M., Huang, B., Lee, S., Lee, K., Tai, C., Lin, Y., Hsieh, M., Kuo, J., Lei, M., et al., 2005. Morphologic remodeling of pulmonary veins and left atrium after catheter ablation of atrial fibrillation: Insight from long-term follow-up of three-dimensional magnetic resonance imaging. Journal of Cardiovascular Electrophysiology 16, 7–12.

Veni, G., Elhabian, S.Y., Whitaker, R.T., 2017. ShapeCut: Bayesian surface estimation using shape-driven graph. Medical Image Analysis 40, 11–29.

de Vente, C., Veta, M., Razeghi, O., Niederer, S., Pluim, J., Rhode, K., Karim, R., 2018. Convolutional neural networks for segmentation of the left atrium from gadolinium-enhancement MRI images, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 348–356.

Verma, A., Kilicaslan, F., Pisano, E., Marrouche, N.F., Fanelli, R., Brachmann, J., Geunther, J., Potenza, D., Martin, D.O., Cummings, J., et al., 2005a. Response of atrial fibrillation to pulmonary vein antrum isolation is directly related to resumption and delay of pulmonary vein conduction. Circulation 112, 627–635.

Verma, A., Wazni, O.M., Marrouche, N.F., Martin, D.O., Kilicaslan, F., Minor, S., Schweikert, R.A., Saliba, W., Cummings, J., Burkhardt, J.D., et al., 2005b. Pre-existent left atrial scarring in patients undergoing pulmonary vein antrum isolation: an independent predictor of procedural failure. Journal of the American College of Cardiology 45, 285–292.

Vesal, S., Ravikumar, N., Maier, A., 2018. 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. pp. 319–328.

Vinnakota, K.C., Bassingthwaighte, J.B., 2004. Myocardial density and composition: a basis for calculating intracellular metabolite concentrations. American Journal of Physiology-Heart and Circulatory Physiology 286, H1742–H1749.

Wang, C., Rajchl, M., Chan, A.D., Ukwatta, E., 2019a. An ensemble of U-net architecture variants for left atrial segmentation, in: Medical Imaging 2019: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 109500M.

Wang, J., Lan, C., Liu, C., Ouyang, Y., Zeng, W., Qin, T., 2021. Generalizing to unseen domains: A survey on domain generalization. arXiv preprint arXiv:2103.03097 .

Wang, Y., Xiong, Z., Nalar, A., Hansen, B.J., Kharche, S., Seemann, G., Loewe, A., Fedorov, V.V., Zhao, J., 2019b. A robust computational framework for estimating 3D bi-atrial chamber wall thickness. Computers in Biology and Medicine 114, 103444.

Wang, Y., Xiong, Z., Nalar, A., Hansen, B.J., Kharche, S., Seemann, G., Loewe, A., Fedorov, V.V., Zhao, J., 2019c. A robust computational framework for estimating 3d bi-atrial chamber wall thickness. Computers in Biology and Medicine 114, 103444.

Whitaker, J., Rajani, R., Chubb, H., Gabrawi, M., Varela, M., Wright, M., Niederer, S., O'Neill, M.D., 2016. The role of myocardial wall thickness in atrial arrhythmogenesis. Ep Europace 18, 1758–1772.

Williams, S.E., Tobon-Gomez, C., Zuluaga, M.A., Chubb, H., Butakoff, C., Karim, R., Ahmed, E., Camara, O., Rhode, K.S., 2017. Standardized unfold mapping: a technique to permit left atrial regional data display and analysis. Journal of Interventional Cardiac Electrophysiology 50, 125–131.

Wood, M.A., Wittkamp, M., Henry, D., Martin, R., Nixon, J., Shepard, R.K., Ellenbogen, K.A., 2004. A comparison of pulmonary vein ostial anatomy by computerized tomography, echocardiography, and venography in patients with atrial fibrillation having radiofrequency catheter ablation. The American Journal of Cardiology 93, 49–53.

Wu, F., Li, L., Yang, G., Wong, T., Mohiaddin, R., Firmin, D., Keegan, J., Xu, L., Zhuang, X., 2018. Atrial fibrosis quantification based on maximum likelihood estimator of multivariate images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 604–612.

Wu, F., Zhuang, X., 2020. Cf distance: A new domain discrepancy metric and application to explicit domain adaptation for cross-modality cardiac image segmentation. IEEE Transactions on Medical Imaging 39, 4274–4285.

Wu, F., Zhuang, X., 2021. Unsupervised domain adaptation with variational approximation for cardiac segmentation. IEEE Transactions on Medical Imaging .

Xia, Q., Yao, Y., Hu, Z., Hao, A., 2018. 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. pp. 211–220.

Xiao, J., Si, D., Wu, Y., Li, M., Yin, J., Ding, H., 2020. Multi-view learning for 3d lge-mri left atrial cavity segmentation, in: Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition, pp. 84–87.

Xiong, Z., Fedorov, V.V., Fu, X., Cheng, E., Macleod, R., Zhao, J., 2018. Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network. IEEE Transactions on Medical Imaging 38, 515–524.

Xiong, Z., Xia, Q., Hu, Z., Huang, N., Bian, C., Zheng, Y., Vesal, S., Ravikumar, N., Maier, A., Yang, X., et al., 2020. A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. Medical Image Analysis 67, 101832.

Yamada, T., Murakami, Y., Okada, T., Okamoto, M., Shimizu, T., Toyama, J., Yoshida, Y., Tsuboi, N., Ito, T., Muto, M., et al., 2006. Incidence, location, and cause of recovery of electrical connections between the pulmonary veins and the left atrium after pulmonary vein isolation. Europace 8, 182–188.

Yang, G., Chen, J., Gao, Z., Li, S., Ni, H., Angelini, E., Wong, T., Mohiaddin, R., Nyktari, E., Wage, R., et al., 2020. Simultaneous left atrium anatomy and scar segmentations via deep learning in multiview information with attention. Future Generation Computer Systems 107, 215–228.

Yang, G., Zhuang, X., Khan, H., Haldar, S., Nyktari, E., Li, L., Wage, R., Ye, X., Slabaugh, G., Mohiaddin, R., et al., 2018a. Fully automatic segmentation and objective assessment of atrial scars for long-standing persistent atrial fibrillation patients using late gadolinium-enhanced MRI. Medical Physics 45, 1562–1576.

Yang, G., Zhuang, X., Khan, H., Haldar, S., Nyktari, E., Li, L., Ye, X., Slabaugh, G., Wong, T., Mohiaddin, R., et al., 2017a. Differentiation of pre-ablation and post-ablation late gadolinium-enhanced cardiac MRI scans of longstanding persistent atrial fibrillation patients, in: Medical Imaging 2017: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 101340O.

Yang, G., Zhuang, X., Khan, H., Haldar, S., Nyktari, E., Ye, X., Slabaugh, G., Wong, T., Mohiaddin, R., Keegan, J., et al., 2017b. 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), IEEE. pp. 844–848.

Yang, X., Wang, N., Wang, Y., Wang, X., Nezafat, R., Ni, D., Heng, P.A., 2018b. Combating uncertainty with novel losses for automatic left atrium segmentation, in: International Workshop on Statistical Atlases and Computational Models of the Heart, Springer. pp. 246–254.

Yu, L., Wang, S., Li, X., Fu, C.W., Heng, P.A., 2019. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer. pp. 605–613.

Zghaib, T., Nazarian, S., 2018. New insights into the use of cardiac magnetic resonance imaging to guide decision making in atrial fibrillation management. Canadian Journal of Cardiology 34, 1461–1470.

Zhang, Q., Wang, J.f., Dong, Q.q., Yan, Q., Luo, X.h., Wu, X.y., Liu, J., Sun, Y.p., 2017. Evaluation of left atrial volume and function using single-beat real-time three-dimensional echocardiography in atrial fibrillation patients. BMC Medical Imaging 17, 44.

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., 2017. Three-dimensional integrated functional, structural, and computational mapping to define the structural "fingerprints" of heart-specific atrial fibrillation drivers in human heart ex vivo. Journal of the American Heart Association 6, e005922.

Zhao, J., Xiong, Z., 2018. MICCAI 2018: Atrial segmentation challenge. http://atriaseg2018.cardiacatlas.org/.

Zhao, Z., Puybareau, E., Boutry, N., Géraud, T., 2021. Do not treat boundaries and regions differently: An example on heart left atrial segmentation, in: 2020 25th International Conference on Pattern Recognition (ICPR), IEEE. pp. 7447–7453.

Zhu, L., Gao, Y., Yezzi, A., Tannenbaum, A., 2013. Automatic segmentation of the left atrium from MR images via variational region growing with a moments-based shape prior. IEEE Transactions on Image Processing 22, 5111–5122.

Zhuang, X., 2019. Multivariate mixture model for myocardial segmentation combining multi-source images. IEEE Transactions on Pattern Analysis and Machine Intelligence 41, 2933 – 2946.

Zhuang, X., Li, L., Payer, C., Štern, D., Urschler, M., Heinrich, M.P., Oster, J., Wang, C., Smedby, Ö., Bian, C., et al., 2019. Evaluation of algorithms for multi-modality whole heart segmentation: an open-access grand challenge. Medical Image Analysis 58, 101537.

Zhuang, X., Xu, J., Luo, X., Chen, C., Ouyang, C., Rueckert, D., Campello, V.M., Lekadir, K., Vesal, S., RaviKumar, N., et al., 2020. Cardiac segmentation on late gadolinium enhancement MRI: a benchmark study from multi-sequence cardiac MR segmentation challenge. arXiv preprint arXiv:2006.12434 .