Left Ventricle Segmentation in Cardiac MR: A Systematic Mapping of the Past Decade

MATHEUS A. O. RIBEIRO and FÁTIMA L. S. NUNES, Universidade de São Paulo

Left ventricle segmentation in short-axis cardiac magnetic resonance images is important to diagnose heart disease. However, repetitive manual segmentation of these images requires considerable human effort and can decrease diagnostic accuracy. In recent years, several fully and semi-automatic approaches have been proposed, mainly using image-based, atlas, graph, deformable model, and artificial intelligence methods. This article presents a systematic mapping on left ventricle segmentation, considering 74 studies published in the past decade. The main contributions of this review are definition of the main segmentation challenges in these images; proposal of a new schematization, dividing the segmentation process into stages; categorization and analysis of the segmentation methods, including hybrid combinations; and analysis of the evaluation process, metrics, and databases. The performance of the methods in the most used public database is assessed, and the main limitations, weaknesses, and strengths of each method category are presented. Finally, trends, challenges, and research opportunities are discussed. The analysis indicates that methods from all categories can achieve good performance, and hybrid methods combining deep learning and deformable models obtain the best results. Methods still fail in specific slices, segment wrong regions, and produce anatomically impossible segmentations.

1 INTRODUCTION

Cardiovascular diseases (heart diseases) lead the world ranking as the highest cause of mortality, and their incidences are increasing [9, 13, 112]. To decrease these indices, there are studies pursuing new biomarkers—that is, accurate metrics capable of diagnosing heart diseases [87]. Clinical metrics such as **ejection fraction (EF)**, as well as mass and shape of the **left ventricle (LV)**, are examples of biomarkers capable of diagnosing various diseases [74, 109], such as cardiomyopathies [89], cardiac dyssynchrony [33], cardiac insufficiency [10], and coronary artery disease

[77, 92]. Segmentation of the LV in medical images can be used for obtaining these metrics [5], allowing their reconstruction, visualization, and manipulation in a 3D environment.

Among the different modalities of medical images, **cardiac magnetic resonance (CMR)** is considered the gold standard in the analysis of the heart function because it can discriminate the different types of cardiac tissues [34, 49, 96]. The manual segmentation of the LV by an expert is a process that requires substantial human effort, usually being restricted to the **end-diastole (ED)** and **end-systole (ES)** phases, which are the most important for obtaining biomarkers [78]. Continuous and repetitive image analysis, such as that required in this process, can lead to visual fatigue, loss of concentration, and, consequently, diagnostic accuracy decrease [104]. Thus, much effort has been employed in the creation of fully and semi-automatic segmentation methods capable of achieving the same quality provided by an expert. However, due to the heterogeneity of the exams, fully automatic segmentation of the LV is still an open problem in the literature [11].

The main objective of this systematic mapping is to assemble, analyze, and determine the current strategies used in the segmentation of the LV in short-axis CMR images, as well as their limitations, evaluation metrics, and image databases used. A total of 74 studies, published over the past decade, were considered in this study.

Although other reviews have been carried out on this matter [40, 76, 78, 129], the state of the art has presented considerable changes in the past decade, mainly due to the expressive increase in **artificial intelligence (AI)** techniques applied in medical image processing. Among all AI approaches, **deep learning (DL)** methods have grown in popularity since 2017, achieving good performance in public databases [11] but still failing in many cases. This stimulated the creation of new hybrid approaches, combining DL with other classic segmentation methods. These new approaches, datasets, and metrics used to evaluate them, as well as a general schematization of the steps involved in the segmentation process, are not explored in the previous reviews. Thus, considering the complexity and importance of the LV segmentation problem and the evolution of the state of the art, the following contributions are offered in the present systematic mapping: (i) a new schematization of the segmentation process and classification of the proposed methods; (ii) the main public image databases are pooled, presented, and discussed; (iii) the main metrics employed to evaluate segmentation methods are categorized and presented; (iv) the performance of methods using public databases is assessed; (v) the main limitations, weaknesses, and strengths of all method categories are presented; (vi) the advantages that hybrid approaches and shape restrictions can provide are discussed; and (vii) trends and research opportunities related to the challenges identified in the included studies are highlighted.

The rest of the article is organized as follows. A description of the CMR examination and the challenges in the segmentation of the LV are provided in Section 2. The procedure and criteria for the systematic mapping are presented in Section 3, whereas the overview of the analyzed studies is presented in Section 4. The stages of pre-processing, **region of interest (ROI)** extraction, segmentation itself, and post-processing of the analyzed studies are detailed in Sections 5, 6, 7, and 8, respectively. Evaluation metrics are described in Section 9, and the public image databases are presented in Section 10. Finally, a discussion, limitations, and conclusion respectively are carried out in Sections 11, 12, and 13.

2 CMR EXAMINATION

The CMR examination is an important method used in the diagnosis of heart disease. Due to the high contrast between regions where there is—or not—blood (e.g., inside the LV cavity), CMR images can discriminate the internal region from the myocardial region, favoring anatomical analysis and obtaining biomarkers associated with heart function [34, 49, 96].
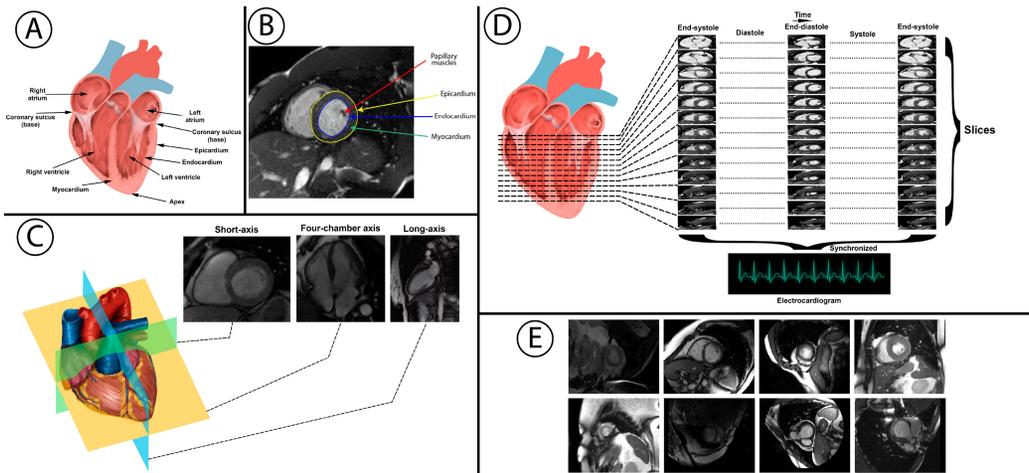
Fig. 1. Summary of CMR examination of the LV. (A) Anatomy of the heart. (B) Example of a short-axis CMR image. (C) Orientation used to obtain the images. (D) Example of a CMR examination. (E) Examples of the variability present in CMR images, including variations regarding the shape, size, intensity, rotation, and location of the LV. Heart images were inspired by Marieb et al. [66], and CMR images were taken from the public databases Sunnybrook, Sata-13, Kaggle, and ACDC (Section 10).

Generally, CMR images are represented using grayscale. The intensity of each pixel is related to the resonance of hydrogen atoms in different tissues of the human body when exposed to radio frequency waves [67]. Thus, in CMR images, regions with higher blood concentrations, such as the inside of the LV, appear brighter, whereas regions where there are muscles or air, such as the myocardium and lungs, appear darker (Figure 1(B)).

Analysis of the CMR examination usually considers one of three planes, defined according to three axes: short-axis, long-axis, and four-chamber axis (Figure 1(C)). Slices are defined considering planes parallel to the one chosen. An image is obtained for each slice, usually comprising a region of a few millimeters in thickness [67]. The short-axis orientation is the most used for the LV analysis because it does not include the atria, veins, and arteries, which have similar intensities and can hinder identifying the structures. The images are obtained between the base and apex of the heart [78].

Utilizing an electrocardiogram (ECG), the obtained images can be synchronized with the heart rate, enabling the analysis of specific phases of the cardiac cycle (Figure 1(D)). Due to the long time spent in the segmentation task, combined with fatigue and loss of concentration caused by continuous and repetitive image analysis [104], the delineation of the LV is usually restricted only to the ED and ES phases, considering the regions delimited by the endocardium and epicardium [78]. Based on segmentation of the LV in these two phases, relevant diagnostic biomarkers can be obtained. This entire process requires a time period varying between 10 and 25 minutes [22, 98, 106].

2.1 CMR Image Segmentation Challenges

One of the factors hindering the fully and semi-automatic segmentation of the LV is the presence of great variability between different CMR images. Ideally, in short-axis images, the bright region delimited by the endocardium has a circular shape, whereas the darker region of the myocardium, between the endocardium and the epicardium, has a ring-like shape. However, several artifacts and variations contribute to altering the patterns and increasing the segmentation difficulty (Figure 1(E)):

Table 1. Keywords and Related Terms

Keywords	Related Terms
segmentation	edge, boundary
left ventricle	myocardium, heart, cardiac
MRI	magnetic resonance imaging, MR imaging, MR images

- Presence of papillary muscles and trabeculations in the LV internal region, which have similar tones to the myocardium and can be confusing [11, 61];
- Blur caused by movement of the heart, the patient, or breathing [106, 111];
- Inclusion of the atrium and the outflow tract in slices close to the base, with intensities similar the internal region of the LV [118];
- Very small LV size in slices near the apex [126];
- Similar intensity between the myocardium and nearby regions, such as the liver and lungs, which can hinder epicardial extraction [21, 63, 115];
- Shape variation of the LV between slices and between patients [37, 64];
- Presence of cardiomyopathies that lead to deformations in the LV shape and increase shape variation;
- Low resolution and noise caused by MRI machines [50].

3 METHODS

The systematic mapping followed four main stages: planning, selection of primary studies, data extraction, and interpretation of results. The databases used were defined according to studies obtained in a previous exploratory analysis, including ScienceDirect,[1] ACM,[2] PubMed,[3] and IEEE.[4]

To achieve the proposed objective in this systematic mapping, the following research questions were proposed:

(1) What methods have been used to segment the LV in short-axis CMR images?
(2) What are the limitations of these methods?
(3) What are the main image databases used in the evaluation?
(4) What are the main evaluation metrics used?

Table 1 shows the keywords and related terms searched in the databases. This review considered only primary studies that propose methods for segmenting the LV in short-axis CMR images. The searches were performed considering the titles and keywords of the studies.

Of all databases, only ScienceDirect does not have the option to restrict the search to the title and keywords. In this case, the search was performed considering the abstract and the result was subsequently reviewed manually.

Table 2 describes the inclusion (I), exclusion (E), and quality (Q) criteria used. Only studies that satisfy the inclusion criteria and do not satisfy any exclusion criteria were accepted. In the first and second stages of the selection, studies prior to 2010 and with fewer than 5 pages were excluded, as in the selected databases they refer to ongoing research, generally without conclusive results. In the following stages, the titles and abstracts were evaluated, and the contents of the remaining studies were analyzed in full. Studies not written in English or that did not process short-axis CMR images of human hearts were excluded. Studies that used examinations with the application of

---

[1]https://www.sciencedirect.com.
[2]https://dl.acm.org.
[3]https://www.ncbi.nlm.nih.gov/pubmed.
[4]https://ieeexplore.ieee.org/Xplore/home.jsp.

Table 2. Inclusion, Exclusion, and Quality Criteria Used in the Selection of Primary Studies

Criteria	Description
I1	Proposition of methods for segmentation of the LV in CMR images
E1	The study was published before 2010
E2	The study has fewer than five pages
E3	The images used are not exclusively of short-axis CMR
E4	The methods are not applied in 2D objects
E5	Does not propose/evaluate methods for segmentation of the LV
E6	Secondary study (e.g., reviews)
E7	Study not written in English
E8	The images are not of human hearts
E9	Use of medications to increase contrast
E10	Use of special CMR images
E11	Does not perform segmentation of the endocardium
E12	Quality score less than 3
Q1	The study describes a pre-processing stage of the images
Q2	The segmentation method is proposed and presented clearly (e.g., formulas, flowcharts)
Q3	The method is clearly evaluated (e.g., tables, graphs)
Q4	Evaluation was carried out with public image databases
Q5	The results were compared with other approaches in the literature
Q6	The limitations of the method are presented and discussed
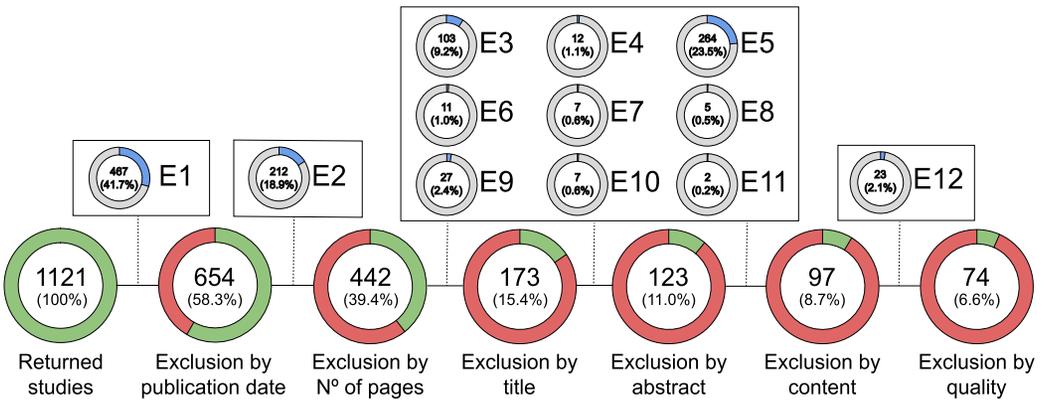
Fig. 2. Search results for each stage of the selection. Percentages refer to the total number of studies returned.

previous medications, such as contrasts, or that did not perform segmentation of the endocardium were also excluded. Figure 2 summarizes the stages for selecting and applying exclusion criteria. Each circle represents a stage and indicates the amount of studies that remain after applying the exclusion criteria, represented by the smaller circles. The numbers within the exclusion criteria show the number of studies excluded by each one.

Finally, the quality of all remaining studies was measured according to the criteria presented in Table 2. The quality criteria were defined in accordance with the research objectives and review questions. For each criterion, a score from the set {0, 0.5, 1} was assigned, respectively indicating that the study does not comply or partially or totally complies with the criterion. A criterion was
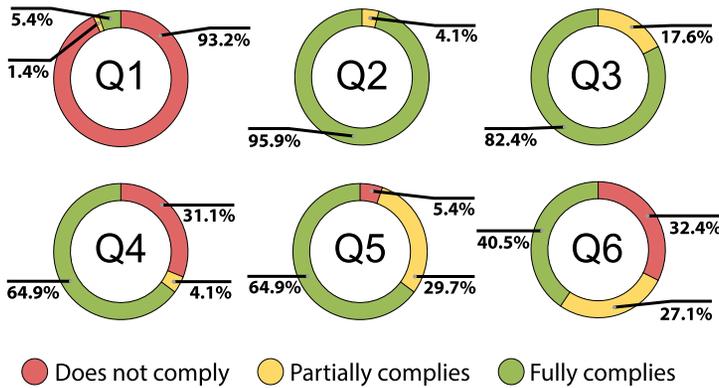
Fig. 3. Percentage of studies that meet the quality criteria. Percentages refer to the total number of studies included.

defined as "partially fulfilled" if it is cited but not detailed. For example, in studies where the comparison of results was only done with a specific approach or subjectively (e.g., only visual comparisons), the Q5 criterion was classified as partially fulfilled. Of all 1,121 returned studies, 74 (6.6%) satisfied all of the criteria and were accepted for the next stage (data extraction).

4 OVERVIEW OF THE INCLUDED STUDIES

Figure 3 shows the amount of compliance of the selected studies regarding quality criteria. Almost all studies include methods that are presented (Q2) and evaluated (Q3) clearly, with the presence of diagrams, graphs, tables, and formulas. About a third of the methods (31.1%) were evaluated using private image databases (Q4), and the majority (94.6%) compared the results with other approaches (Q5). Despite the variability and the challenging contents of CMR images (Section 2), only 6.8% of the studies mention some pre-processing stage (Q1) and about a third (32.4%) does not present the limitations of the proposed approach (Q6).

4.1 Processing Stages

Generically, the segmentation procedure can be defined as the process that receives images containing or not objects of interest as input and returns a set of contours or regions indicating the location of these objects. For the segmentation of the LV, this procedure can be divided into the following stages: pre-processing, ROI extraction, segmentation itself, and post-processing. Figure 4 illustrates these stages, including, at each stage, the techniques used in the analyzed studies. Only the segmentation stage itself is mandatory, although the others help to reduce the complexity of the problem and rectify the contours. Non-mandatory stages are represented with dashed lines.

The pre-processing stage (Section 5) refers to procedures performed on the images to reduce variability between images and facilitate the segmentation process. Only 12.1% of the studies describe this stage.

The ROI extraction stage is responsible for limiting the region of the image where the segmentation method will be applied, reducing the complexity of the problem and execution time. The methods presented in this stage are divided into two categories: manual and automatic. In the first group, the approaches are built based on the delimitation of an initial contour (IC), reference points selection (RPS), sub-image selection (SIS), or first frame segmentation (FFS). In the second group, it is common to use AI or image processing (IP) techniques or to employ a fixed region selection
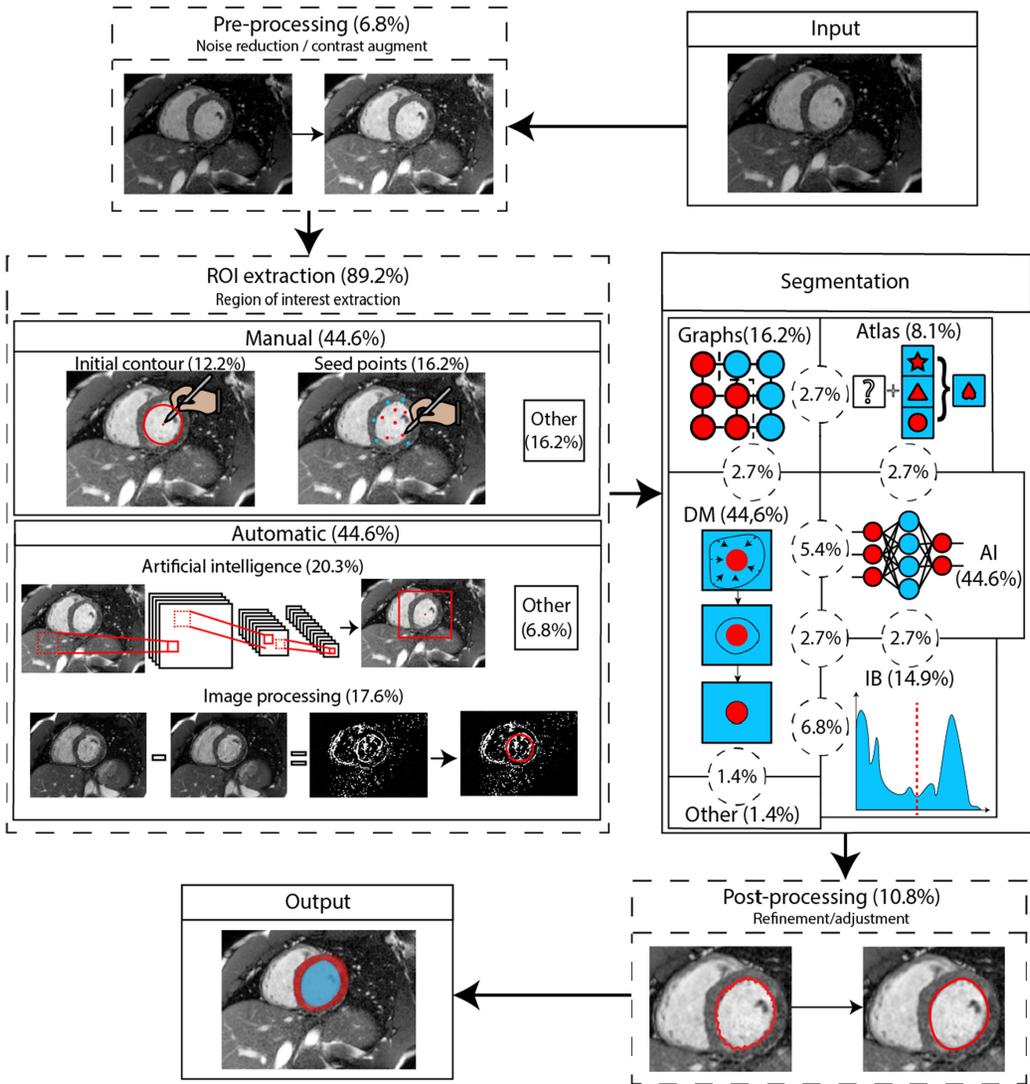
Fig. 4. Schematization of the processes used in the segmentation of the 74 analyzed studies. The percentages refer to the total number of studies analyzed. Dashed lines indicate non-mandatory stages.

(FRS). The details of the techniques are presented in Section 6. Although optional, this stage is found in most studies (89.2%).

In the segmentation stage (Section 7), the percentages of each category in Figure 4 and the intersections respectively indicate the proportion of studies that use methods in each category and the proportion of hybrid methods. The segmentation task was divided into five main categories, according to how frequently it was used in studies: atlas (Section 7.1), graphs (Section 7.2), **deformable models (DM)** (Section 7.3), **image-based (IB)** (Section 7.4), and AI (Section 7.5). Another approach, which does not fall into any of the categories, is discussed in Section 7.7. About 27.1% of the studies present hybrid methods that combine strategies from two or more categories. The methods that combine DM with AI and DM with IB methods correspond to a fifth and a

quarter of all hybrid methods, respectively. Only two studies combined these three approaches, and none of the analyzed studies were classified into more than three categories.

It is important to notice that even though some approaches use methods from different categories for ROI extraction and segmentation stages, they are only classified as hybrid if the combination is used in the segmentation stage. For example, in the work of Hu et al. [39], although DL is used alongside IB methods, since it is applied to only extract the ROI through an initial contour, this study is not classified as hybrid.

After obtaining the segmentation, some studies (10.8%) perform an additional post-processing stage (Section 8). In this stage, the contours are refined, smoothness is increased, and the failures caused in the previous stages are reduced.

4.2 Segmentation Process

Table 3 shows the summary contour of primary studies. The studies were classified into each of the five segmentation categories in Figure 4, in which hybrid categories have the plus sign. The column "Basic Principle" describes more specifically the techniques of each category used in the segmentation stage of each study, whereas the "ROI" column indicates the category that belongs to the method utilized to extract the ROI. Other considerations made are detailed in the following.

Table 3. Summary of the 74 Investigated Studies

Category	Ref.	Basic Principle	ROI	A/S	EPI	AD	Evaluation Databases
Atlas	[130]	Atlas	–	Fully	Yes	Yes	Private
	[8]	Multi-atlas	–	Fully	Yes	Yes	Private
DM	[115]	Active contour	IP	Fully	Yes	No	Sunnybrook
	[44]	Active contour	IP	Fully	Yes	No	York, Sunnybrook
	[102]	Active contour	IP	Fully	Yes	No	Private
	[43]	Active contour	IC	Semi	Yes	No	York
	[1]	Active contour	IC	Semi	Yes	No	York
	[75]	Active shape model	RPS	Semi	Yes	Yes	York
	[2]	Active shape model	–	Fully	Yes	Yes	Private
	[83]	Active shape model	FFS	Semi	Yes	Yes	York
	[91]	Elastic deformable mesh	IC	Semi	Yes	Yes	Sunnybrook
	[81]	Deformable mesh	FFS	Semi	No	No	Private
	[61]	Level set	IC	Semi	Yes	No	Sunnybrook
	[95]	Level set	IC	Semi	Yes	No	Sunnybrook
	[118]	Level set	RPS	Semi	Yes	No	Sunnybrook
	[4]	Level set	RPS	Semi	No	No	Private
	[111]	Level set	RPS	Semi	Yes	No	Private
	[107]	Level set	IP	Fully	Yes	No	Private
	[79]	Level set	IC	Semi	Yes	Yes	York
	[42]	Level set	–	Fully	Yes	Yes	Sunnybrook
	[21]	Deformable surface	IP	Fully	Yes	Yes	Private
Graph	[82]	Graph cut	FFS	Semi	No	No	Private
	[65]	Graph cut	RPS	Semi	No	Yes	LV-11
	[6]	Graph cut	FFS	Semi	Yes	Yes	Private
	[64]	Graph cut	FFS	Semi	Yes	No	LV-11
	[70]	Graph cut	FFS	Semi	Yes	Yes	Private
	[125]	Random walk	RPS	Semi	No	No	Sunnybrook
	[32]	Random walk	RPS	Semi	Yes	Yes	Private
	[15]	Sinc. superpixels with Graph	–	Fully	No	No	York, Sunnybrook
AI	[14]	C-means	IP	Fully	Yes	Yes	The heart database
	[35]	C-means	IP	Fully	No	No	Private
	[124]	C-means	IP	Fully	No	Yes	Sunnybrook

(Continued)

Table 3. Continued

Category	Ref.	Basic Principle	ROI	A/S	EPI	AD	Evaluation Databases
	[127]	Convolutional neural network	AI	Fully	No	Yes	Sunybrook
	[98]	Convolutional neural network	AI	Fully	Yes	Yes	LV-11, Kaggle
	[101]	Convolutional neural network	AI	Fully	No	Yes	Sunnybrook
	[131]	Convolutional neural network	–	Fully	Yes	Yes	ACDC, Sunnybrook
	[31]	DBN	SIS	Semi	Yes	Yes	Private
	[30]	Encoder-decoder network	RPS	Semi	Yes	Yes	Private
	[29]	Encoder-decoder network	SIS	Semi	Yes	Yes	Private
	[56]	Fully convolutional network	AI	Fully	No	Yes	Sunnybrook, Kaggle
	[45]	Fully convolutional network	IP	Fully	Yes	Yes	ACDC, Kaggle, LV-11
	[85]	Fully convolutional network	–	Fully	Yes	Yes	Sunnybrook
	[47]	Random/Mondrian forest	FFS	Semi	No	No	ACDC
	[23]	U-Net	AI	Fully	Yes	Yes	Sunnybrook
	[103]	U-Net	AI	Fully	Yes	Yes	ACDC
	[90]	U-Net	AI	Fully	No	Yes	Private
	[128]	U-Net	AI	Fully	Yes	Yes	Sunnybrook, ACDC
	[126]	U-Net	AI	Fully	No	Yes	York, Sunnybrook
	[100]	U-Net	AI	Fully	Yes	Yes	ACDC
	[20]	U-Net	FRS	Fully	Yes	Yes	Sunnybrook, ACDC
	[22]	U-Net	SIS	Semi	Yes	Yes	Sunnybrook
	[18]	U-Net	SIS	Semi	No	Yes	ACDC
IB	[39]	Thresholding + DP	AI	Fully	Yes	Yes	Sunnybrook
	[106]	Thresholding	IP	Fully	No	No	Private
Atlas + AI	[122]	Multi-atlas + fusion network	RPS	Semi	Yes	Yes	SATA-13, Sunnybrook
	[7]	Multi-atlas + SVM/KNN	RPS	Semi	Yes	Yes	SATA-13
DM + Graph	[26]	Level set + Graph cut	RPS	Semi	Yes	No	Sunnybrook
	[25]	Level set + Graph cut	RPS	Semi	Yes	No	Sunnybrook, LV-11, York
DM + AI	[119]	Active contour + U-Net	AI	Fully	Yes	Yes	Private
	[5]	Level set + Autoenconders	AI	Fully	No	Yes	Sunnybrook
	[105]	Level set + C-means	FRS	Fully	Yes	No	Sunnybrook
	[71]	Level set + DBN	IC	Semi	No	Yes	Sunnybrook
DM + AI + IB	[73]	Level set + DBN + Thresholding	AI	Fully	Yes	Yes	Sunnybrook
	[72]	Level set + DBN + Thresholding	AI	Fully	No	Yes	Sunnybrook
DM + IB	[50]	Active contour + Thresholding	IP	Fully	Yes	No	Private
	[108]	Active contour + Thresholding	SIS	Semi	No	No	Sunnybrook
	[37]	Active contour + DP	FRS	Fully	Yes	No	Sunnybrook
	[88]	Active contour + DP	IC	Semi	No	No	York
	[89]	Active contour + DP	IC	Semi	No	No	Sunnybrook, York
DM + Other	[63]	Active contour + PCNN	IP	Fully	Yes	Yes	Sunnybrook
Graph + Atlas	[36]	Graph cut + Multi-atlas	IP	Fully	Yes	Yes	Private
	[46]	Graph cut + Multi-atlas	–	Fully	Yes	Yes	Private
AI + IB	[38]	GMM + DP	FRS	Fully	Yes	No	Sunnybrook
	[60]	K-means+ Thresholding + DP	FRS	Fully	Yes	No	Sunnybrook

ROI, method used to extract the region of interest; A/S, fully automatic/semi-automatic; EPI, method also segments the epicardium; AD, needs annotated databases. Other acronyms can be found in Section 7.

Automaticity (A/S). The methods are defined as semi-automatic if they need any user inter-action during the ROI extraction stage (Figure 4). In all studies, user interactions refer to man-ual procedures used to select the LV region or initialize the main method in the ventricular re-gion (Section 6). Thus, in the analyzed studies, any manual procedure is contained in the ROI extraction stage. Although presented as fully automatic, methods in which the images used need to be cut or resized manually before segmentation were also considered as semi-automatic, since this step can be considered as manual ROI extraction. This practice is common in DL methods, in which networks usually need images with a specific size. In some cases [23, 101], automatic ROI
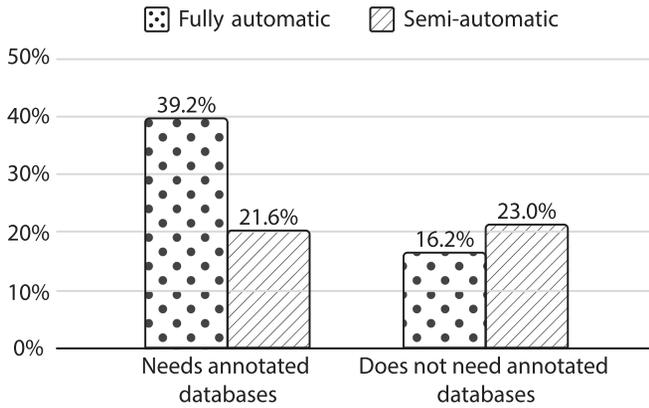
Fig. 5. Need for annotated databases according to the automaticity of methods.

extraction procedures were presented. Studies that do not define automatic methods or report only that there have been changes in the images are classified as semi-automatic.

Epicardium segmentation (EPI). All methods proposed in the studies included perform the segmentation of the endocardium, used in the calculation of the volume of blood pumped by the heart. However, only 70.3% of these also segment the epicardium. For individual methods, the strategy used in the segmentation of the epicardium is similar to that used in the segmentation of the endocardium. However, for some hybrid methods, methods from different categories are used to obtain each contour. Such strategies are detailed in Section 7.8.

Use of annotated databases (AD). This column shows the methods that need sets of examples previously segmented for their execution. These examples are used, for example, to create statistical models, to train AI techniques, as atlases, or to assist in the ROI extraction stage. Most of the approaches (60.8%) use annotated databases. Fully automatic methods have, in general, a greater need for annotated examples, as illustrated in Figure 5.

Databases used in the evaluation. This column indicates which image databases are used to evaluate the methods. The past decade has seen an increase in the number of public databases (Section 10), mainly directed to competitions in the segmentation area. Although increasingly common, many approaches still use private databases (Q4 in Figure 3), making it difficult to compare performance between different approaches. It is important to note that databases that mix public and private images were not considered as public databases, as, for example, in the work of Woo et al. [111].

5 PRE-PROCESSING

Pre-processing is the first stage before segmentation, mainly used to reduce the variability of images. Any process carried out prior to the ROI extraction stage and the application of the segmentation method was interpreted as pre-processing of the main method.

In the analyzed studies, the pre-processing stage reduces noise, normalizes the image color intensities, and enhances contrast to better identify the endocardium region. In the work of Wu et al. [115], partial differential equations of fourth order are used to remove noise in homogeneous regions, preserving edges and highlighting the myocardium, whereas Dakua [25] uses an anisotropic filter based on statistical models.

Other options include the use of the median filter [91] to reduce blur and image normalization [45, 107], which performs pixel intensity redistribution over a predetermined interval.

Table 4. Approaches Used for Extracting the ROI

Type	Category	Freq. (%)
Manual	Reference points selection (RPS)	16.2
	Initial contour (IC)	12.2
	First frame segmentation (FFS)	9.5
	Sub-image selection (SIS)	6.8
Auto	Artificial intelligence (AI)	20.3
	Image processing (IP)	17.6
	Fixed region selection (FRS)	6.8
N/A	–	10.8

6 ROI EXTRACTION

ROI extraction is the second stage before segmentation. In contrast to the pre-processing, almost all analyzed methods (89.2%) have this stage, which is important for reducing complexity, computational cost, and initialization of segmentation methods in subsequent stages. In the present review, "extracting the ROI" is defined as any procedure used to locate the LV in the image, either by selecting a sub-image, by selecting reference points, or by defining an initial contour.

The approaches were separated into categories, according to the strategy employed, as shown in Table 4. Each category was classified as automatic or manual. Descriptions for each approach are given in the following.

6.1 Manual ROI Extraction

Reference points selection. In this approach, the user selects points in the LV region. Points can be used to initialize a DM [4, 111, 118] or to align atlases [7, 122]. They can also indicate regions inside and outside the endocardium, helping to define edge weights in graph-based approaches [25, 26, 32, 65, 125].

Initial contour. An initial contour is given by the user in the region close to the endocardium, usually with primitive shapes such as circles or rectangles. This strategy is mainly used in the initialization of DM [43, 61, 71, 91, 95] and in methods that use polar coordinates, such as **dynamic programming (DP)** [88, 89].

First frame segmentation. In this category, the user performs manual segmentation of the images in the first frame. Segmentation can be propagated to the next frames, creating point-to-point relationships [81], or can be used to define edge weights in graph-based approaches [6, 64, 82]. It can also be used in the construction of specific patient models, such as point dictionaries [83], distributions of radial distances and intensities [70], or pixel classifier training [47].

Sub-image selection. In this category, the sub-image containing the LV is manually selected [108]. It is the procedure commonly used in DL methods, in which fixed-size manual cuts are necessary due to the model restrictions [18, 22, 29]. In some cases, rotations and centralizations [31] are also employed.

6.2 Automatic ROI Extraction

Artificial intelligence. In this approach, an AI model is trained utilizing annotation examples. The model can be used to obtain the initial contour, the center of the LV, or the sub-image (e.g., rectangle). The initial contour is used in the initialization of a more precise method, such as a DM [119] or a more specific AI model [103]. The sub-image is primarily obtained based on regression, in which the coordinates of the delimiting rectangle are estimated [5, 38, 72, 90, 101, 126, 127]. Regression is also used to obtain the center of the LV [98]. The model can also be trained to classify

possible sub-images [56] or to classify pixels that belong or not to the object of interest, in which the sub-image is obtained around the classified region [39, 100, 128].

Image processing. Image processing includes methods that use classic IB techniques such as edge detection, image subtraction, thresholding, and mathematical morphology. In general, the approaches utilize two characteristics of CMR images: (i) difference in intensity between the myocardium and the LV in the endocardial region, and (ii) heartbeat in different frames of the cardiac cycle. Regarding the first characteristic, clustering and thresholding techniques are commonly used to determine brighter regions of the image. Due to the lighter shade, the endocardial region is one of the highlighted regions. The ROI can be found considering the region with the largest size [102], its circularity [35], its proximity to the center of the image [14], or using these factors together [124]. Edge detection filters, such as Canny [44], can also be used to highlight the endocardium. Regarding the second characteristic, image subtraction is the most used approach since it highlights edge regions where there is movement, such as the endocardium. Subtraction can occur between consecutive frames [107, 115], or frames referring to the ED and ES phases [50, 106]. The image can also be compared with an image generated from the average of the frames [21, 63]. The LV can be identified by looking for circular objects using the Hough transform [115] or by calculating the center of mass of an image [107]. These procedures are usually performed in the middle slice (between the base and the apex) of the set of images and at the ED phase, as in this phase the internal region of the LV has a maximum size and does not suffer interference from the atrium and the outflow tract.

Fixed region selection. This includes methods that use specific restrictions from the authors themselves to select the sub-image, based on experience and *a priori* knowledge about the image databases used. Examples include selecting a fixed-size rectangle centered on the image [37, 38, 60] or where the heart is often present in the analyzed examples [105].

7 SEGMENTATION METHODS

The main segmentation method is employed after completing the ROI extraction stage. All of the methods listed here are fully automatic since any manual intervention is performed in the previous stages.

7.1 Atlas

In this category are the approaches that use segmented and labeled examples (called *atlases*) to segment the LV. The main idea of the method is to find a transformation function, or registration, capable of mapping the labeled regions of an atlas to regions of the target image and then use them to classify the pixels. Classically, only one atlas is used in the process. However, all analyzed works use multiple atlases in segmentation, a method called *multi-atlas segmentation*.

The multi-atlas segmentation strategy can be divided into four main phases: registration, selection, label propagation, and fusion of atlases. Registration can be performed using affine [8, 130] or non-rigid [46] transformations. The process is intended to minimize a cost function that can consider the intensity, gradient [7], and mutual information [8], among others. For semi-automatic methods, the registration is carried out based on reference points selected by the user.

The atlas set most similar to the target image is selected after the registration. The similarity can be achieved using the same function as the registration. DL methods can also be used to extract features and choose atlases [122].

Each atlas of the chosen set then has its labels propagated. The propagation can consider points or set of points, called *patches*. For each pixel in the target image, its label can be obtained from the combination of nearby patches contained in each selected atlas. This stage is called *atlas fusion* and can be done in several ways.

The fusion can be carried out by voting, where each patch has a weight proportional to its relevance [8]. Machine learning methods, such as support vector machines (SVM), can use information obtained from patches to classify each pixel, and deep neural networks can be trained for this purpose [122]. The problem can be mapped as a problem of minimum cut in graphs, where each pixel is represented by a node and edges connect each pixel with those corresponding to each atlas, with weight proportional to the similarity [46].

7.2 Graph-Based Approaches

In the analyzed studies in this category, the images are represented as graphs, where each pixel is represented by a vertex and edges connect each pair of 4-neighbors. Minimum graph cut is the predominant technique used in the analyzed studies.

A graph cut can be defined as a set of edges that, when removed, divide the graph into disjoint subgraphs, in such a way that each subgraph (e.g., set of pixels) is associated with a class. Each class is represented by at least one seed vertex. The cut must be made in such a way that only the vertices belonging to each class remain connected to the corresponding seeds. The seed vertices can be from the graph itself when the class of any vertex is known. This occurs, for example, when there is manual selection of points [25, 26, 65]. Another option is defining additional vertices, called *terminals*, which are initially connected with all vertices of the graph and each one represents a class [6, 36, 46].

The effectiveness of the minimum graph cut method depends on the cost function and the strategy used to define the edge weights. For a pair of neighboring vertices representing pixels, the weight of the edge connecting them must reflect the probability of the pair belonging to the same class. In this case, the similarity of pixel intensity is the most widely used factor [6, 25, 36, 46, 64, 65, 82]. On the edges connecting vertices representing pixels and terminal vertices, the weight reflects the probability that the vertex belongs to the terminal class. The distance from a vertex to a target can be considered as the weight sought, so the shorter the distance, the greater the probability that this vertex belongs to the LV region. The target can be a seed vertex [26], the LV center [6, 82], or the resulting structure in the segmentation of the anterior frame [64]. The weight can also be determined using atlases to build probabilistic models, indicating the probability that each vertex belongs to a class [36] or including new edges, connecting vertices of the target image and the corresponding atlas set [46].

The cost function can also consider the shape of the cut [64]. In segmentation methods of the two ventricles, in which two cuts are made, the shape of one can be used to rectify the other [65]. The best cut is achieved by minimizing the cost function. However, for classifications of multiple regions, obtaining the optimal result is complex and therefore the problem can be relaxed [70].

Another approach that uses a similar strategy is the random walk [32, 125]. Instead of cuts, each vertex is classified individually. Each edge receives a weight proportional to the probability of being crossed in a random walk. A vertex is classified with the same label as the first seed vertex that is reached if a random walk starts at that vertex. In the work of Cai et al. [15], spectral graphs are used. Instead of just vertices of nearby pixels connecting, all vertices are connected, and groups are obtained based on similarity, called *superpixels*. The superpixels of each image are synchronized to achieve segmentation.

7.3 Deformable Models

The DM category is the most present in the analyzed studies, parallel to AI. In DM approaches, segmentation is obtained interactively from the deformation of a contour or surface. The models need to be initialized in some region of the image, and their evolution is guided by minimizing

an energy function. The function can have several terms related to image information, such as intensities, gradients, textures, and the shape of the DM itself, ensuring its smoothness.

One of the oldest and most used types of DM is active contours, or snakes. Originally proposed by Kass et al. [41], the energy function of the active contour is divided into internal and external forces. Using variational methods to minimize the function, energies are converted into forces capable of deforming the contour. The internal forces are responsible for maintaining the smooth shape of the contour, whereas the external forces use information from the image to guide the contour. Another widely used approach is the level set. In it, the energy function is defined for larger dimensions (e.g., 3D), and the contour is obtained implicitly from the projection of the function at a certain level, usually zero. This strategy allows the segmentation of multiple objects. For example, in the work of Liu et al. [61], the contour of the endocardium and epicardium is respectively defined as the level zero and level k of a level set function.

Generally, in works that use only DM, the internal energy is defined based on restrictions that consider simple geometric shapes, such as circles or ellipses [43, 115, 118], an association between the shape of the endocardium and the epicardium [44, 63, 111], or statistical models constructed based on examples (Section 7.6). In the work of Arrieta et al. [4] and Khamechian and Saadatmand-Tarzjan [44], an additional restriction prevents the epicardium and endocardium contours from touching.

External energy is the most altered among the approaches presented in the analyzed studies. In general, external energy uses information from the image and can be composed by terms related to the edge or region of the image. Edge-based terms use information about the image's gradient, so the contour is guided to the edges of the image. In contrast, region-based terms use information about the intensity and homogeneity of regions defined in the images to guide the contour. With exception of the work of Tufvesson et al. [102], which proposes a specific region term based on balloon force, all other studies analyzed are based on models proposed previously in the literature.

The region-based term originally proposed in the work of Chan and Vese [17] is the most used [5, 79, 111]. The term favors segmentation in images with little or no edges. In some level set approaches, the term has been adapted to include topology preservation [4], avoiding the segmentation of disconnected regions, and local and global image information [26, 95], improving performance in cases with little homogeneity. Other terms include those proposed in the work of Li et al. [53], used by Hu et al. [37], and the term proposed another work by Li et al. [52], used by Wang et al. [105], that proposes alternatives for cases with little homogeneity.

The edge-based term uses two variations. The first was originally proposed in work by Li et al. [54], where an evolution method using regularized distances is presented to avoid the level set reset problem. Although the described method can also be used in region-based terms, the proposed approach uses only edge-based terms. This variation is used in work by Liu et al. [61] and in conjunction with DL methods [71–73]. The second variation is the one originally proposed in work by Xu and Prince [117], which presents a new type of snake, called *gradient vector flow snake*. Its formulation uses an edge map and increases the movement freedom of the contour, allowing the segmentation of concave regions. This variation is combined with the forces that help expand the contour, such as the balloon force [63] and the force proposed in work by Lee et al. [50]. It can also be approximated by convolution functions, as proposed by Wu et al. [115], to decrease computational processing time.

One of the problems with edge-based terms is their susceptibility to noise and their dependence on the existence of a high gradient value at the expected contour location. However, region-based terms usually require different homogeneous regions. As seen in Section 2, both problems can exist in CMR images. For this reason, some studies use a mixed approach. The DM STACS (stochastic active contour scheme), originally proposed by Pluempitiwiriyawej et al. [80], is a mixed approach

example. Its energy terms are adapted in the work of Khamechian and Saadatmand-Tarzjan [43, 44].

Three analyzed studies propose other types of DM. In the work of Schaerer et al. [91], the my-ocardium is represented by a triangular mesh that is deformed according to dynamic elastic equations. Deformation parameters are estimated by examples. One of the advantages of this method is that the use of dynamic equations reinforces temporal smoothness between frames. Cordero-Grande et al. [21] model the DM evolution problem in a framework based on Markov random fields. The framework encompasses characteristics such as image intensities, and gradients and uses maximum likelihood estimators to estimate parameters. In the work of Punithakumar et al. [81], non-rigid transformations are used to evolve a DM and point-to-point relationships are created between segmentations of adjacent frames to restrict evolution.

One of the main limitations of DM is a high sensitivity to initialization. For the contour to evolve correctly, it should start close to the correct location. A very distant start-up increases the chances of the contour getting stuck in a local minimum. For this reason, these methods depend on a good initialization (Section 6). In some approaches, the endocardial contour can be used to initialize the epicardium contour [115]. For processing a cardiac sequence, the segmentation result of a previous frame can be used in the initialization of the DM [44, 81, 118]. The temporal coherence between adjacent frames can be explored in the energy function, as in methods that use optical flow [1, 108]. In hybrid approaches (Section 7.8), the segmentation resulting from another method can be used in the DM initialization.

Other types of DM are those based on the active shape model. Analogous to the active contour, the method uses a point distribution model constructed from aligned segmented examples capable of representing the object of interest. Based on an initial contour, seed points are defined, and deformations are calculated considering the restrictions imposed by the model, to prevent generating inconsistent formats. This method was adapted for cases of abnormal hearts by Albá et al. [2]. One of the advantages of this method is the possibility of creating specific patient models that are updated with each new segmented image [83]. Different models can be built considering the spatial and temporal variations of the CMR exams and combined to perform the segmentation [75].

7.4 IB Methods

IB methods use attributes intrinsic to the image, such as color and texture, to perform the segmentation. The classification used in this category is the same as that proposed in the work of Petitjean and Dacher [78] and includes methods of thresholding, region growth, and DP.

In the context of this review, thresholding is used to classify image pixels into two classes by comparing their intensity with a fixed value, called a *threshold*. This method is used primarily to segment the endocardium and explores the property that regions with blood appear brighter than muscles and nearby regions (Section 2). Thus, the objective is to find the threshold that best highlights the internal region of the endocardium. It can be found using the Otsu algorithm [72, 73], analyzing the Fourier transform of the intensity histogram [108], or using an interactive method [50, 106].

Due to the absence of contrast in the region close to the epicardium, threshold methods are not usually applied in its segmentation. In this case, DP is normally used [39, 60]. DP-based methods aim to find an optimal path based on minimizing a cost function. This function uses an edge map, where lower costs are added in pixels representing the edges. This method is usually applied using polar coordinates, to reduce the complexity of the problem. The map can be obtained using the gradient image [88, 89] or applying the Canny edge detector [50]. In some cases, the non-maxima gradient suppression technique is applied to produce thinner edges and improve results [37, 39, 60].

7.5 Artificial Intelligence

The AI-based methods are the most used in the analyzed studies, in parallel with DM. Approximately, almost half of the studies (44.6%) employ AI to perform the segmentation. This category includes supervised and unsupervised learning methods to perform classification or regression. Of these, most use DL (75.7%). Considering the good performance achieved by DL in many fields, including medical image analysis [59], the use of AI-based methods in the segmentation of the LV has grown in popularity since 2017, although more conventional categories still remained, such as DM-based and graph-based approaches [78].

7.5.1 Deep Learning. In this review, the oldest DL method is the one presented in the work of Ngo and Carneiro [71], which uses **deep belief network (DBN)**. One of the advantages of DL methods over other methods discussed is the ability to automatically learn the characteristics inherent to the image, which are necessary to perform segmentation, without having to manually define specific cost functions for graphs, energy for DM, and registration for atlas.

Notwithstanding automatic learning, it is still necessary to define the architecture and the loss function. Of all DL architectures, those based on **convolutional neural networks (CNNs)** are the most used in cardiac image processing [5, 98, 127]. For segmentation, specifically, **fully convolutional networks (FCN)** are widely used CNN variations, because they can make pixel-by-pixel predictions, producing as output binary masks indicating the classification of each pixel in the image [29, 45, 56, 85].

The U-Net network [86] is the most used and explored FCN in the literature [58] because it achieves good results in the segmentation of medical images [103]. In the analyzed approaches, the proposals mostly include changes in the structure and use of this network. The most used loss function is the Dice loss, based on the evaluation metric of the same name (Section 9). Other functions include Jaccard [22, 23] and cross-entropy [103].

Several modifications for the U-Net are presented in the analyzed studies, to improve performance. Modifications include the addition of batch normalization and residual learning [22], and interleaved attention blocks, recurring feedback, and in-depth supervision [100]. Cong and Zhang [20] propose Invert-U-Net, a variation that has a greater number of filters in shallow layers of the network, unlike the original U-Net. According to the authors, this configuration can help the network detect more detailed patterns regarding edges and shapes, as well as improve network performance. However, in the work of Vigneault et al. [103], the Omega-Net variation has a specific module that performs image transformations, preparing them to be used as input to the U-Net network. Images of different resolutions can be used to improve accuracy in edge regions [101, 131].

One of the problems with DL methods is the number of parameters that must be learned, which are usually in the millions and consume a lot of time and computational resources for the training phase. Some studies propose optimization techniques to minimize these limitations. In the work of Curiale et al. [23], a U-Net is presented that uses sparse information to avoid overfitting and reduce the complexity of the model. In the work of Charmchi et al. [18], Fourier and gradient flow analysis are applied to the convolutions of a U-Net network, which allows to detect and remove less relevant layers or replace filters with simpler variations. In this case, the authors were able to reduce the number of parameters to be learned from 31 million to 0.5 million.

The similarity between image slices and close frames can also be explored. In the work of Yang et al. [119], a 2D U-Net performs the initial segmentation of each image of a cardiac cycle of the same slice. The resulting masks and original images are collected in volumes and used as input to a 3D U-Net, which rectify the segmentation. Savioli et al. [90] add a recurring layer that is able to retain information about previously segmented frames, promoting temporal coherence. In the approach presented in the work of Zhang et al. [128], the contextual information of slices in

sequence is used in a model adapted from the U-Net, which receives not only the image of a slice but also the image and the segmentation from the previous slice. LSTM (long short-term memory) modules, which can incorporate temporal coherence, are used in conjunction with an FCN in the work of Du et al. [30].

Another category of DL methods that does not use convolutions includes those based on autoencoders. One of the advantages of this category is the ability to carry out part of the training in an unsupervised way, requiring a smaller number of annotated examples [73]. In the analyzed approaches, autoencoders are used in DBN for pixel classification [71–73] and regression [31]. In the work of Avendi et al. [5], sparse autoencoders are trained in an unsupervised manner and used to start stacked autoencoders.

7.5.2 Other AI Methods. For unsupervised approaches, clustering methods based on fuzzy C-means are the most used. The fuzzy C-means method divides the image pixels into N classes (e.g., endocardium, myocardium, background), according to a similarity function, which indicates the probability of each pixel belonging to each class. The function can consider information such as intensity [105], distance [35], and/or format [124]. In the work of Hu et al. [38], another clustering approach uses the Gaussian mixture model (GMM), where each class is represented by a Gaussian, which is embedded in the histogram of image intensities using the expectation-maximization algorithm. The model is used to indicate the probability that each pixel belongs to every one of the classes.

Two supervised approaches are proposed in the work of Bai et al. [7], which uses SVM and k-nearest neighbors (KNN) for atlas classification (see Section 7.8), and in the work of Koné and Boulmane [47], which uses hybrid forests. In the latter, random and Mondrian forests are used together in the classification of pixels based on information about intensity and position. For each segmented slice, the model is optimized, promoting consistency between slices.

7.6 Shape Restrictions

During segmentation, several anatomical restrictions can be added to facilitate the application of methods and prevent the generation of segmentations with impossible shapes. For example, the contour of the endocardium can be restricted within the epicardium [70], or it can be used to initialize DM to segment the epicardium [26, 73]. However, the most used type of anatomical restriction is the LV shape prior.

Anatomically, the LV has a round shape in the short axis images (Section 2). This information can be used to extract the ROI, through the Hough transform, to identify circles (Section 6), for image transformation into polar coordinates, given the center of the LV, as is done in DP methods (Section 7.4) and in clustering methods, to select the most circular object [35]. In DM, a new energy term can be integrated considering circular shape prior, preventing the contour to evolve into irregular shapes due to noise [43, 44, 107, 115, 118]. Similarly, cost terms can be added to ensure a good choice of the C-means method [124] or to define weights in graphs [125]. Another way used independent of the method category and widely used to impose the circular shape prior is to obtain the convex hull. In this strategy, some contour points produced by the method are interpolated to obtain a convex contour [38, 39, 47, 60].

The circular shape prior can be understood as a weak restriction, insufficient to capture all of the anatomical characteristics of the LV and the variations found in CMR exams. Segmented examples can be used to build strong priors capable of representing the LV shape in detail. The examples can be analyzed to obtain parameters of Markov random fields [21], elasticity and deformation of DM [91], and construction of ellipses [14]. The most common approach includes aligning the example images. This is already done naturally in atlas-based methods (Section 7.1) and the active shape

model, which use them to build point distribution models (Section 7.3). Aligned images can be used to build distance distribution models between the center and the contour edges. This model can be used to define costs in graph cut [6, 82] and random walk [32] methods or energy terms in DM [73]. A variation proposed in the work of Mahapatra [64, 65] calculates distance and angle distributions between points in the left and right ventricles. Probability maps can also be created, associating the probability of each pixel belonging to the LV, and used in AI training [131], to define DM energy terms [42] and weights in graphs [36].

7.7 Other Approaches

This category only has the **pulse-coupled neural network (PCNN)** method. Despite being defined as a neural network, the PCNN cannot be classified as AI because it does not have a training phase. In this method, each neuron represents a pixel and receives a different stimulus according to its intensity. After receiving a specific amount of stimulus, a neuron is activated and sends pulses to nearby neurons, which are interpreted as stimuli. In the work of Ma et al. [63], a simplified PCNN is proposed. In this study, the synchronization of neuron pulsations is analyzed to obtain segmentation.

7.8 Hybrid Methods

Several hybrid strategies have been proposed to overcome the limitations inherent to each approach. In this review, any study that integrates methods from different categories was defined as a hybrid. Two types of hybrid approaches were considered: approaches that use the methods together to segment the endocardium and/or epicardium (h1), and approaches that use one method for the endocardium and another for the epicardium (h2). Table 5 presents the hybrid strategies used.

DM methods are the most prevalent in hybrid approaches because of their high accuracy. Due to the sensitivity of initialization, they are commonly used to rectify endocardium segmentations produced by other methods (e.g., AI) or initiated in the endocardium to segment the epicardium. AI has been used to learn characteristics relevant to atlas selection or to restrict the evolution of a DM. Multi-atlas methods have been united to methods that use graphs in the choice of examples used in defining costs in graphs or to construct them.

Only the approach proposed in the work of Ngo et al. [73] can be classified as belonging to the two types of hybrids. Proposed as an extension of the work of Ngo and Carneiro [71, 72], the hybrid method uses the Otsu threshold to segment the endocardium. This segmentation is rectified by a DM that includes energy terms based on the classification of a DBN. To segment the epicardium, rays are launched from the center of the attained endocardium to find edge points. Another DM is initialized at these points and segments the epicardium.

8 POST-PROCESSING

The optional post-processing stage is the last part of the segmentation process and is related to the refinement and adjustments applied to the segmentations. Post-processing methods can be divided into removing disconnected regions and holes, smoothing contours, and checking consistency between segmentations.

Smoothing methods seek to remove sudden variations and discrepancies in the contours. Divergent points can be found by calculating the distance between each point and its neighbors and then removed if it exceeds a threshold value [98]. The abrupt variations can be detected by analyzing the image in the frequency domain. Considering the distance between each point and the contour center, the Fourier transform is applied to detect high frequencies, which are attenuated

Table 5. Hybrid Approaches Proposed in the Analyzed Studies

Type	Ref.	Method	Strategy
h1	[46]	Graph + Atlas	Aligned multi-atlas pixels are connected using graphs; graph cut reveals segmentation
	[36]	Graph + Atlas	Probabilistic map built by multi-atlas and used to define graph cut weights
	[7]	Atlas + AI	Choice of atlases based on the classification of SVM and KNN
	[122]	Atlas + AI	DL extract features used in the multi-atlas fusion
	[71]	DM + AI	DBN segmentation included in the DM energy function
	[5]	DM + AI	Autoencoder segmentation included in the DM energy function
	[119]	DM + AI	DL segmentation rectified by DM
	[72]	DM + AI + IB	Otsu threshold segmentation rectified by DM and restricted by DBN
	[88]	DM + IB	DM applied to polar coordinates solved by DP
	[89]	DM + IB	DM applied to polar coordinates solved by DP
	[108]	DM + IB	Threshold segmentation rectified by optical flow
h2	[26]	DM + Graph	Graph cut for endocardium and DM for epicardium
	[25]	DM + Graph	Graph cut for endocardium and DM for epicardium
	[63]	DM + Other	PCNN for endocardium and DM for epicardium
	[50]	DM + IB	Threshold for endocardium and DM for epicardium
	[37]	DM + IB	DM for endocardium and DP for epicardium
	[105]	DM + AI	DM for endocardium and C-means for epicardium
	[60]	AI + IB	KVP for endocardium and DP for epicardium
	[38]	AI + IB	MMG for endocardium and DP for epicardium
h1 + h2	[73]	DM + AI + IB	Otsu threshold rectified by DM restricted by DBN for endocardium and DM for epicardium

Classification: h1, mixture of methods to segment the same region of the heart; h2, utilization of individual methods for the segmentation of the endocardium and epicardium.

[38, 60]. Along with these rectifications, the contour points can also be interpolated to generate cubic splines [98] and Bezier curves [60].

One of the possible problems, mainly in pixel classification methods, is the presence of disconnected regions and holes in the LV region. Connected component analysis is used to remove disconnected regions, adopting the permanence of the largest region and the removal of the other regions [20, 45]. In the work of Koné and Boulmane [47], the region that has the greatest intersection with the segmented region of the previous slice is chosen. Mathematical morphology operations can be used to fill holes [45, 101].

Consistency between segmentations is also a factor that can be explored after applying the main segmentation method. The analysis of connected components can be applied in the 3D domain to ensure consistency between segmentations of different slices [45]. In the work of Savioli et al. [90], energy optimization methods based on conditional random fields and semantic flow are applied to segment different frames, reinforcing the temporal consistency.

9 EVALUATION METHODS

All studies included in the review perform some type of evaluation of the proposed method (criterion Q3 in Figure 3). The evaluation is done by comparing the segmentation obtained by the
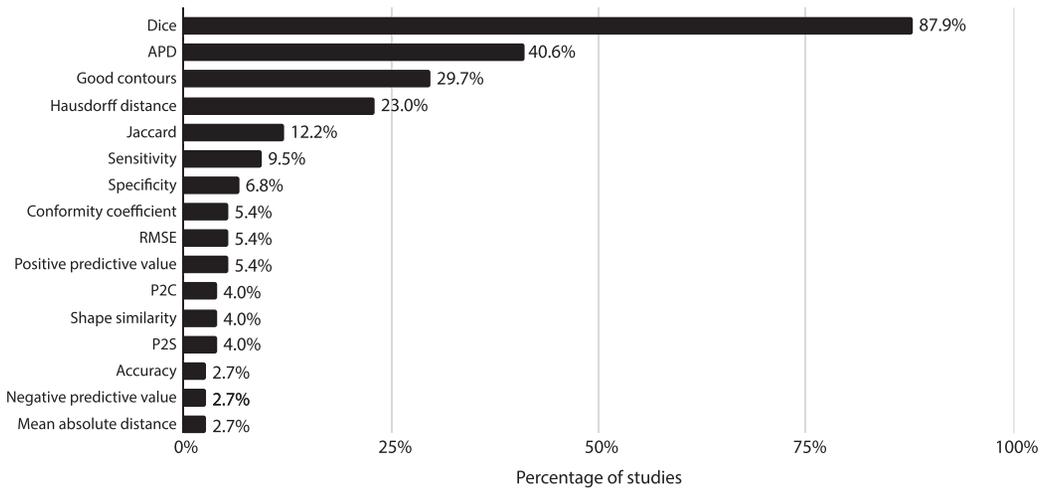
Fig. 6. The most used geometric metrics in the evaluation methods.

method with the segmentation carried out by experts (ground truth). The evaluation metrics were divided into two categories: geometric metrics and clinical metrics.

9.1 Geometric Metrics

The geometric metrics consist of values based on the joint analysis of the segmentations produced by the method and by the expert, considering geometric information such as placement in the image, distances between contours, overlap of regions, and shape. They are generic metrics that do not consider the specific context of the LV and can therefore be applied to other segmentation problems.

The graph in Figure 6 shows the most used geometric metrics in the analyzed studies. Only metrics used by at least two studies were included.

Table 6 presents the equation definitions of the metrics included in Figure 6. In it, Ω, S, and G are respectively defined as the sets of all points in the image, the points of the segmentation resulting from the proposed method, and the ground truth points. These sets can consider 2D (slice) or 3D (volume) space and are not limited to the pixels of the image. All metrics are formulated according to these sets and are classified, when relevant, as overlap or distance metrics. The "Other" category indicates that the metric does not fit the previous two categories.

9.1.1 Overlap Metrics. In the overlap metrics, the evaluation is interpreted as a classification problem, in which the sets S and G indicate the points correctly classified (i.e., belonging to the region defined by the segmentation). All metrics, with the exception of the conformity coefficient, comprise values between 0 (worst result) and 1 (best result).

The most widely used metric is the Dice, also known as the F-score. Usually, the evaluation is performed considering the entire space. However, due to the small size of the LV in relation to the whole image, there is an imbalance between the positive and negative classes, impacting the result of metrics that depend on Ω and favoring the use of more stable metrics, such as the Dice and Jaccard. A more balanced option is to consider Ω as the ROI, performed in the work of Tan et al. [98] for the calculation of specificity and negative predictive value. The downside of this approach is that it is difficult to compare results, due to the varying size and location of the ROI.

9.1.2 Distance Metrics. Used in conjunction with the overlap metrics, the distance metrics indicate how close two segmentations are in space. In this category, the sets S and G usually refer

Table 6. Formulations of the Main Geometric Metrics Used in the Studies Analyzed
for the Segmentation Evaluation

Class	Metric	Definition/Formula
Overlap	Dice	$\dfrac{2\lvert S \cap G \rvert}{\lvert S \rvert + \lvert G \rvert}$
	Jaccard	$\dfrac{\lvert S \cap G \rvert}{\lvert S \cup G \rvert}$
	Conformity coefficient	$\dfrac{3 Dice - 2}{Dice}$
	Sensitivity	$\dfrac{\lvert S \cap G \rvert}{\lvert G \rvert}$
	Specificity	$\dfrac{\lvert \Omega \rvert - \lvert S \cup G \rvert}{\lvert \Omega - G \rvert}$
	Accuracy	$1 + \dfrac{\lvert S \cap G \rvert - \lvert S \cup G \rvert}{\lvert \Omega \rvert}$
	Positive predictive value (PPV)	$\dfrac{\lvert S \cap G \rvert}{\lvert S \rvert}$
	Negative predictive value (NPV)	$\dfrac{\lvert \Omega - (S \cup G) \rvert}{\lvert \Omega - S \rvert}$
Distance	Hausdorff	$\max(\max_{s \in S} \min_{g \in G} \lvert\lvert s - g \rvert\rvert, \max_{g \in G} \min_{s \in S} \lvert\lvert s - g \rvert\rvert)$
	P2C	$\dfrac{1}{\lvert S \rvert} \sum_{s \in S} \min_{g \in G} \lvert\lvert s - g \rvert\rvert$
	APD	Similar to P2C, but only if g is in the perpendicular line of the tangent line defined by s
	RMSE	$\dfrac{1}{\lvert S \rvert} \sqrt{\sum_{s \in S} \min_{g \in G} \lvert\lvert s - g \rvert\rvert^2}$
	Mean absolute distance	$\dfrac{1}{2}\left(\dfrac{1}{\lvert S \rvert} \sum_{s \in S} \min_{g \in G} \lvert\lvert s - g \rvert\rvert + \dfrac{1}{\lvert G \rvert} \sum_{g \in G} \min_{s \in S} \lvert\lvert s - g \rvert\rvert \right)$
Other	Good contours (GC)	Proportion of the contours where APD is < 5 mm
	Shape similarity	See Section 9.1.3

G is the set of points of the ground truth, S is the set of points of the segmentation performed by the method,
$\lvert\lvert . \rvert\rvert$ is the Euclidean distance, and Ω is the set of all points.

only to the contours, disregarding the internal regions of the segmentations. In most cases, image
resolution information is used to convert the distance from pixels to millimeters.

The most used distance metric is the **average perpendicular distance (APD)**. This metric
is very similar to the **point-to-curve distance (P2C)**, which indicates the average distance from
each point in S to the closest point in G. However, APD only considers the points of G that are in the
perpendicular line of the tangent line of S defined by $s \in S$ [84]. This allows the detection of spikes
and other contour deviations that can be ignored by the P2C metric. Other metrics similar to P2C
and APD are the **root mean squared error (RMSE)**, which considers quadratic distances, and the
absolute average distance, which considers not only the distance from each point in S to G but also

from each point in G to S. Considering only the average distance can hide specific segmentation faults. Therefore, the Hausdorff distance is used to find such discrepancies. It measures the largest distance between any point in one of the segmentations and the other segmentation.

The distance calculation varies according to the space adopted. For example, for the point-to-surface distance (P2S), the P2C variant for 3D space (Table 6), two studies [2, 21] use the distance between points, whereas in the work of Zhuang et al. [130], the distance between points and triangles is calculated. The same is true for the Hausdorff distance. For the ACDC database (Section 10), it is defined between points and triangles, whereas in the work of Yang et al. [122], the distance between points is used.

9.1.3 Other Metrics. In the studied papers, the average of the metrics is usually the most prevalent value presented. When 2D space is considered, the average is calculated from all slices, whereas for 3D, the average is calculated from all volumes for each patient. To ensure that the approaches have good results across all slices, the **good contours (GC)** metric is used. Originally proposed as one of the evaluation metrics of the Sunnybrook database (Section 10) for 2D space, it indicates the proportion of contours segmented by the method where APD is < 5 mm in relation to the ground truth, which is considered a good contour.

One of the problems with distance metrics is that they do not consider segmentation shapes. To consider both shape and distance, shape similarity is used in some other approaches. Proposed in the work of Pluempitiwiriyawej et al. [80], the metric is applied in 2D space and uses distance maps, which indicate that the distance of each pixel to the contours. Based on the maps, a phase map is constructed, containing the angle of each pixel. The maps obtained for the segmented image and the ground truth image are subtracted and used in the final metric calculation.

9.2 Clinical Metrics

Unlike geometric metrics, clinical metrics are specific to the area of cardiology and are used by doctors in the diagnosis of diseases [87, 110]. They are individually quantified for each set of a patient's segmented slices, calculated based on the LV volume.

For obtaining the volumes, Simpson's method is one of the most important and most used method [87]. In it, the volume is obtained from the sum of the segmented areas of each slice, considering the thickness and spacing between slices [16]. The ED and ES volumes (V_{ED} and V_{ES}, respectively) are the most used to calculate the metrics. The EF is the most used in the analyzed studies and indicates the percentage of blood ejected from the LV during the cardiac cycle, as defined in Equation (1).

$$EF = 100\% * \frac{V_{ED} - V_{ES}}{V_{ED}} \tag{1}$$

In the analyzed studies, the evaluation of clinical metrics is done by comparing the values obtained by the proposed method with the values provided by the expert. The difference in metrics can be obtained and the Bland-Altman [12] or linear regression graphs can be used for analyzing volumes, ejection fractions, myocardial mass (MM), and so forth. Pearson's correlation coefficient is also used to find out if there is a linear correlation between the obtained values.

10   PUBLIC DATABASES

As shown in criterion Q4 of Figure 3, about two-thirds (69%) of the analyzed studies use public databases in the evaluation of methods. Table 7 summarizes the main information from these databases. With the exception of the Heart Database and York, all databases were proposed in competitions. The geometric and clinical metrics in the respective columns of Table 7 show the metrics used for evaluation in each database.

Table 7. Summary of the Public Databases Used in the Analyzed Studies

Name	Year	Freq. (%)	Patients	Ground Truth	Geometric Metrics	Clinical Metrics	Space
The Heart Database	2006	1.4	18	Endo, Epi	P2S	Volumes, MM, EF	3D
York	2008	16.2	33	Endo, Epi	Landmark error, volumetric error	Volumes	2D
Sunnybrook	2009	48.6	45	Endo, Epi	Dice, APD, GC	Volumes, MM, EF	2D
LV-2011	2011	6.8	200 (100 train and 100 test)	Myo	Jaccard, sensitivity, specificity, PPV, NPV	Volumes, MM, EF	2D
SATA-13	2013	4.1	155 (83 train and 72 test)	Myo	Dice, symmetric Hausdorff distance	–	–
Kaggle	2015	4.1	1,140	N/A	–	Volumes	N/A
ACDC	2017	10.8	150 (100 train and 50 test)	Myo	Dice, Hausdorff distance	Volumes, MM, EF	3D

The Heart Database.[5] Proposed in the work of Najman et al. [69], the database contains 18 cases of patients with symptoms of acute myocardial infarction. The images were segmented by two experts at the ED and ES phases and by an automatic method in the other phases of the cardiac cycle.

York.[6] This database contains sequences obtained from 33 patients, including cases of cardiomyopathies, aortic regurgitation, and ischemia [3]. All images from slices where the endocardium and epicardium are visible were manually segmented (5,011 of 7,980 images). In this database, the contours are represented by 32 points, available in pixel coordinates. The proposed evaluation method uses the landmark error, calculated in the same way as P2C, but only for the 32 points. The volumetric error metric indicates the average of the landmark error from each slice within a volume.

Sunnybrook.[7] Also known as the MICCAI (Medical Image Computing and Computer-Assisted Intervention) 2009 database due to the conference in which the competition took place, Sunnybrook has 45 cases including normal patients, heart failure with and without ischemia, and hypertrophic cardiomyopathy [84]. It is the most used public database in the analyzed studies. The three main metrics coincide with the most used metrics in Figure 6. During the event, the database was evenly divided into training, validation, and online datasets. The conference site is no longer active, and therefore all sets were made available.

LV-11.[8] This database was also made available by the MICCAI conference, which took place in 2011. It contains 200 exams, including cases with coronary artery disease and myocardial infarction [97]. Half of the database (100 cases) was separated for training and the rest for testing. Only training cases have ground truth. To participate in the competition, which still accepts new submissions, the segmentations for the testing set must be sent online. One of the restrictions of the

---

[5]The Heart Database: http://laurentnajman.org/heart (2006).
[6]Cardiac MRI dataset—York: http://jtl.lassonde.yorku.ca/software/datasets/ (2008).
[7]The Sunnybrook Cardiac Data: http://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/ (2009).
[8]International Workshop on Statistical Atlases and Computational Models of the Heart: http://www.cardiacatlas.org/challenges/lv-segmentation-challenge/ (2011).

database is that the images cannot be used for other purposes that do not involve the creation of segmentation methods.

SATA-13.[9] Used during the workshop entitled Multi-Atlas Labeling beyond the Cranial Vault, also from the MICCAI conference, the database contains 155 cases, 83 for training and 72 for testing. The images were taken from a larger database, called *DETERMINE* (Defibrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation), which contains cases of patients with coronary artery disease and ventricular dysfunction. Despite focusing on atlas-based methods, the event is open for submission to any type of approach.

Kaggle.[10] Kaggle is the database used in the Second Annual Data Science Bowl competition, with the participation of 192 teams, for the creation of algorithms to estimate ventricular volume. This database is the largest of those cited, containing 1,140 cases of ED and ES volumes, taken from various sources. Due to the objective of the event, the images do not have the ground truth, only the volumes. During the competition, the database was divided into training, validation, and testing sets, and users could use the Sunnybrook database to train the models.

ACDC.[11] ACDC is the most recent database, containing 150 cases (100 for training and 50 for testing), equally distributed among healthy patients, with myocardial infarction, and dilated and hypertrophic cardiomyopathies [11]. In addition to segmentation, the competition also evaluated methods for classifying pathologies, providing additional patient information such as weight, height, and age. An analysis of the competition results and submitted methods is available in [11]. New submissions can still be prepared for the testing set.

Several similarities and differences can be observed in the databases. All databases, with the exception of SATA-13, calculate the ED and ES phase volumes. However, the space considered to obtain the geometric metrics varies between 2D and 3D. For 2D spaces, metrics are calculated for each image individually, whereas for 3D spaces, they are obtained for each volume. This difference can cause discrepancies in the value of the metrics, which hinders comparing results between different databases and different methods.

The representation of the ground truth also occurs in different ways. For LV-2011, SATA-13, and ACDC databases, a mask (i.e., image with individual classification for each pixel) for the myocardium is used. For the Heart Database, York, and Sunnybrook databases, a set of pixels representing the contours are provided for the endocardium and epicardium. The region around the epicardium contains the endocardium region so that the myocardium region can be obtained by subtracting them. These characteristics can modify the values obtained by the metrics.

The York and LV-2011 databases are the only ones that provide the ground truth for all phases of the cardiac cycle. The others, with the exception of Kaggle, provide only the ED and ES phases, which are necessary for calculating the main clinical metrics.

11 DISCUSSION

11.1 Comparison between Methods

Table 8 shows the results for the geometric metrics used in the Sunnybrook database. Although other databases are presented in Section 10, only the results for the Sunnybrook database are presented, as it is the most used public database in the analyzed studies and does not have an official classification table. Due to the great variability in the use of metrics, only the studies that

---

Table 8. Results of Methods for the Sunnybrook Database

Type	Ref.	Method	Group Used	Endocardium Dice	Endocardium APD (mm)	Endocardium GC (%)	Epicardium Dice	Epicardium APD (mm)	Epicardium GC (%)
Fully	[39]	IB	val+on	0.90 (0.03)	1.95 (0.48)	96.80 (7.00)	0.93 (0.02)	1.98 (0.53)	**98.4 (6.50)**
	[124]	AI	val	0.89 (0.03)	2.23 (0.50)	–	–	–	–
	[126]	AI	val+on	0.93 (0.03)	1.46 (0.23)	**100**	–	–	–
	[127]	AI	val+on	0.92 (0.03)	2.00 (0.29)	98.35 (3.44)	–	–	–
	[101]	AI	val+on	0.91	2.08	96.80	–	–	–
	[85]	AI	all	0.92 (0.01)	2.23 (0.31)	94.19 (7.38)	0.93 (0.02)	2.13 (0.28)	95.64 (7.11)
	[42]	DM	3 cases	0.91 (0.07)	**1.21 (1.29)**	–	**0.96 (0.02)**	**0.87 (0.52)**	–
	[60]	AI + IB	all	0.88 (0.03)	2.36 (0.39)	91.17 (8.52)	0.94 (0.02)	2.19 (0.49)	90.78 (10.68)
	[38]	AI + IB	all	0.89 (0.03)	2.24 (0.40)	91.06 (9.42)	0.94 (0.02)	2.21 (0.45)	91.21 (8.52)
	[37]	DM + IB	all	0.89 (0.04)	2.30 (0.36)	92.72 (6.86)	0.93 (0.02)	2.20 (0.45)	94.33 (7.88)
	[5]	DM + AI	val+on	**0.94 (0.02)**	1.81 (0.44)	96.69 (5.70)	–	–	–
	[105]	DM + AI	7 cases	**0.94**	1.23	**100**	–	–	–
	[72]	DM + AI + IB	val	0.88 (0.03)	2.34 (0.46)	95.91 (5.28)	–	–	–
	[73]	DM + AI + IB	val	0.88 (0.03)	2.34 (0.46)	95.91 (5.28)	0.93 (0.02)	2.08 (0.28)	94.95 (6.18)
	[63]	DM + Other	all	0.86 (0.05)	2.35 (0.54)	91.20 (13.23)	0.92 (0.02)	2.41 (0.45)	93.32 (10.72)
Semi	[125]	Graph	val	0.88 (0.03)	2.41 (0.38)	93.10 (6.30)	–	–	–
	[118]	DM	all	0.92 (0.02)	1.77 (0.32)	–	0.94 (0.02)	**1.85 (0.38)**	–
	[95]	DM	val	0.95 (0.03)	–	–	**0.97 (0.01)**	–	–
	[61]	DM	val	0.92 (0.03)	**1.76 (0.57)**	97.83 (0.21)	0.95 (0.01)	1.95 (0.43)	97.48 (0.33)
	[91]	DM	on	0.87 (0.04)	2.97 (0.37)	–	0.92 (0.02)	3.14 (0.33)	–
	[122]	Atlas + AI	val	0.92 (0.04)	2.04 (0.44)	97.58 (7.12)	0.95 (0.01)	1.90 (0.22)	**100**
	[26]	DM + Graph	–	**0.97 (0.01)**	–	–	0.95 (0.01)	–	–
	[71]	DM + AI	all	0.90 (0.03)	2.08 (0.40)	**97.91 (6.18)**	–	–	–
	[108]	DM + IB	val	0.90 (0.01)	–	–	–	–	–

The values are presented in the average format (std). The best result for each metric is highlighted in bold. Fully, fully automatic; Semi, semi-automatic; val, validation dataset; on, online dataset; "–," not informed.

used the metrics originally proposed by the database are displayed. In cases where more than one result was reported (e.g., for different groups in the database), only the result that contains the best Dice is displayed. The group used column indicates which sets of cases from the database were used to calculate the metrics. The best value for each metric is highlighted in bold.

Generally, semi-automatic methods obtained higher Dice values than the fully automatic ones, indicating better overlapping of the segmentations. In contrast, fully automatic methods showed better consistency and proximity between contours, with lower APD and higher values for good contours. Manual contour initialization can contribute to greater variability in the segmentation of semi-automatic methods, whereas fully automatic methods usually include strategies to better restrict the shape of contours, providing greater consistency between segmentations.

The results presented in the analyzed studies suggest that approaches from all categories are able to obtain results close to those of the experts. All average values for Dice are above 0.85, which is a value considered excellent by some authors [4, 64]. Similarly, average APD values below 3 mm and percentages above 90% for GC were achieved by all approaches, indicating that the contours obtained are close to the ground truth in the images of several slices.
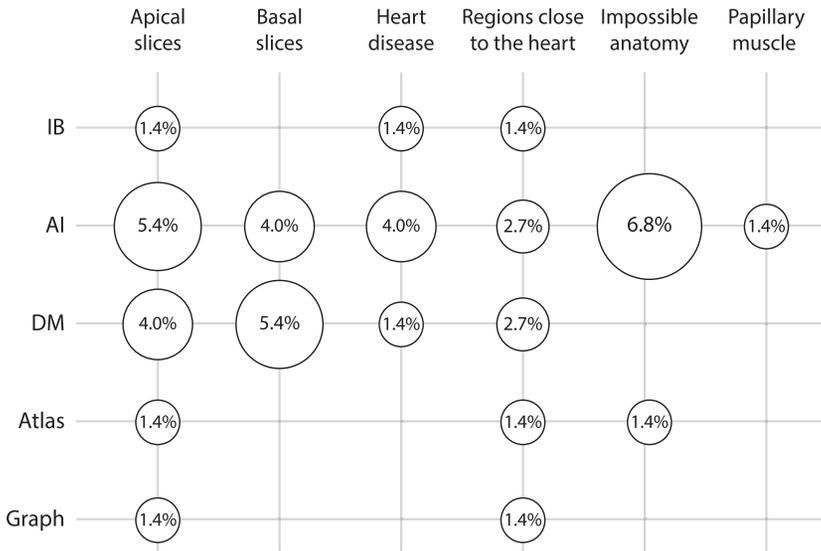
Fig. 7. Limitations most pointed out by the methods related to the segmentation problem. Only non-hybrid methods are considered.

Among all categories, DM-based methods and hybrid methods containing DM approaches achieved the best results for both endocardial and epicardial segmentations. It should be emphasized that type h2 hybrid methods (Table 5) were analyzed as individual methods for each segmentation. For fully automatic methods, the hybrid methods composed by AI and DM approaches proposed in the work of Avendi et al. [5] and Wang et al. [105] achieved the best Dice values for the endocardium (0.94). For semi-automatic methods, the graph-based method proposed in the work of Dakua [26] achieved the best Dice (0.97), followed by the DM of Soomro et al. [95] (0.95). DM methods showed very precise segmentations, proven by the lowest APD values for both the endocardium and the epicardium.

For fully automatic approaches, DL methods showed results very close to those of DM. The method proposed in work by Yang et al. [126] achieved a value just 0.01 below the best result Dice value. Despite not having the best APD, this method presents good stability when achieving a low standard deviation, of only 0.23 mm and a GC rate of 100%.

11.2   Limitations of the Approaches

To understand the main difficulties encountered by the methods of each category, the limitations presented in the studies were categorized and summarized in Figure 7. Hybrid approaches were disregarded due to the lack of clarity, in the text of the analyzed studies, regarding which constituent method is related to each limitation.

In general, the most cited difficulty in all categories refers to the segmentation of apical slices. In these slices, the reduced LV size, combined with low resolution and the presence of noise in the images, hinder defining cost/energy functions, aligning atlas and pixel classification, thus impacting the methods of all categories. An alternative to minimize this problem is to impose spatial restrictions between apical slices and nearby slices, which have a larger ventricular area.

The segmentation of basal slices is the second biggest difficulty pointed out, which is faced mainly by methods of the DM and AI categories. In these slices, regions of the atrium and the outflow tract may fuse, which have intensities similar to the LV, resulting in images with little or

no presence of edges. Thus, DM can "leak," and methods of AI and thresholding may erroneously include such regions. A possible solution is to include shape prior constraints, given that adding these regions produces distorted shapes.

Confusion with regions close to the heart is another difficulty encountered by all categories. Such regions include the lungs and liver, which have color intensities similar to the myocardium and the cardia, which has a shape and intensity similar to the endocardium. This limitation indicates the importance of choosing a good ROI extraction method.

One of the problems most indicated in studies that apply DL methods is the production of segmentations representing impossible anatomies. The absence of shape restrictions increases the creation of anatomically impossible segmentations by these methods. Anatomical errors include results in which the endocardial region is not included in the epicardial region, the presence of holes within the regions, and the classification of multiple disconnected regions for the same class. In a study for the public ACDC database [11], it was found that despite achieving results very close to those of the experts, DL methods can produce many anatomically impossible segmentations. For the best method, the manual analysis of segmentation performed by experts showed that in 41 of 50 CMR examinations (82%), the method produces anatomically incorrect segmentations in at least one image, indicating that the analysis of the overlap and distance geometric metrics is not enough to assess the anatomical coherence of the segmentations. The production of disconnected regions can also occur in atlas-based methods, depending on how the multi-atlas fusion is done. Anatomy errors are strongly related to the apical and basal slices, due to the specific characteristics of these slices, mentioned in Section 2. Other studies of DL methods [11, 99] have shown that DL methods produce worse results for these slices.

A way to improve the performance of DL and atlas methods is to increase both the quantity and variability of the segmented examples used. In a study carried out in the work of Tao et al. [99], databases of exams from different suppliers and medical centers were used to train three CNNs. In the analysis of the results, it was found that the training CNN with the greatest amount and variability of images achieved the best results. However, there is not always a sufficient number of examples to cover the entire population of interest, and there is no guarantee that the data augmentation approaches, which usually include rotations, translations, and noise addition, are able to represent all variability of the examinations. Because of this lack of coverage, many AI methods can fail in more uncommon cases, such as when the patient has a heart disease that changes the shape and function of the LV (e.g., cardiomyopathies). Another issue that hinders data augmentation is that randomly deforming the image can produce anatomically incorrect training samples. To address these issues, recent studies have proposed new data augmentation strategies that use active shape models to create valid examples [57] and generative adversarial networks to interpolate between slices [116].

In contrast to DL and atlas methods, energy-based methods (e.g., clustering, DM, graphs) require few, if any, annotated examples. One of the advantages of these methods is that several terms, including image information and shape priors, can be added and adjusted to prevent the creation of anatomically impossible segmentations and to guarantee highly accurate and good quality of the contours. However, such cost/energy functions need to be defined manually, and there is no guarantee that they will achieve good results in all possible variations. In this scenario, the definition of a function that covers all cases is a very complex problem, forcing authors to make assumptions and relaxations (e.g., approximating the distribution of intensities to a Gaussian distribution). Thus, those methods are usually designed to be case specific, and they can present performance divergence if applied in cases not considered by the function [24]. DL methods, however, are able to learn functions automatically from training samples, even when generic and simple loss functions, such as cross-entropy and Dice, are used for training.

11.3 Difficulties in Evaluating Approaches

There are several problems regarding the evaluation of the proposed approaches. Results from different databases can be impacted by the characteristics of the database, such as resolution, slice thickness, image acquisition protocols, and so forth. This complicates comparing results for private databases.

Even for methods that use the same public databases, two general questions make comparison difficult: how to use the databases and what method is used to calculate the metrics. In the Sunnybrook database (Table 8), some studies use the entire database, whereas others use one of the subdatasets (training, validation, and online), choose a random sample, or do not inform how it is used. In addition, metrics can be calculated based on different premises, such as considering a 2D or 3D space, using different units of measurement (e.g., pixel instead of millimeters), and dividing results for the ED and ES phases. Similar problems were encountered in other databases.

In addition to the general issues mentioned, there are specific difficulties observed in the analyzed studies:

(1) The various metrics are usually presented as the averages of all tested cases, which can hide potential limitations in specific regions of the LV, such as the apical and basal slices (Figure 7).
(2) The original evaluation method employed in the Sunnybrook competition only calculates the Dice and APD metrics for good contours [84]. This makes comparison difficult, especially when GC values are not presented. In addition, it is not clear if the original evaluation method was applied in all analyzed studies.
(3) Many CMR examinations include extra slices showing regions outside the range defined between the apex and the base of the heart. These slices can contain structures very similar to the LV, such as the left atrium and the outflow tract (Section 2.1), which can be mistakenly segmented, especially by automatic approaches. This is a problem because the segmentation of extra slices can significantly change the estimation of clinic metrics. However, many segmentation evaluation procedures, such as the one employed in the Sunnybrook database (Section 10), only considers the interval defined by the ground truth and ignores extra segmentations. Thus, even with good results, it is not clear if segmentation methods are able to only produce segmentations in valid slices.
(4) Few studies present specific results for each pathology. Although the databases contain different cases of heart disease, the evaluation of the methods and calculation of the metrics generally consider the whole set of cases, not allowing to individually analyze the performance of the method for each heart disease.
(5) The overlap and distance metrics together can provide clues about the proximity of segmentations but fail to detect anatomical errors. In the ACDC database (Section 10), for example, the images produced by DL methods had to be evaluated manually by experts to detect anatomical errors. As this is one of the major problems of recent DL methods, it is necessary to create metrics to automatically identify anatomical errors. Metrics such as shape similarity (Section 9.1.3) are alternatives that can be explored.
(6) Many methods and evaluation procedures analyze images individually, without considering inter-slice and inter-frame consistency. Although maintaining good performance in many images, it is not clear if those methods can maintain consistency between consecutive images. In addition, since many databases only have ground truth segmentations for ED and ES phases, it is unknown if the approaches can perform well in all of the slices generated during the entire examination.

Table 9. Comparison of Results for Two AI Methods That Used Different Training
and Testing Databases

Atlas + AI [122]					
Training Database	**Testing Database**	**3D Dice**	**3D Hausdorff**		
Sunnybrook	Sata-13	0.73 (0.06)	15.32 (3.88)		
Sata-13	Sunnybrook	0.77 (0.04)	10.68 (3.84)		
Sata-13	Sata-13	0.83 (0.04)	11.35 (3.53)		
AI [45]					
Training Database	**Testing Database**	**Jaccard**	**Sensitivity**	**Specificity**	**PPV**
ACDC	LV-2011	0.71 (0.13)	0.91 (0.06)	0.81 (0.15)	0.82 (0.12)
LV-2011	LV-2011	0.74 (0.15)	0.96 (0.03)	0.84 (0.16)	0.87 (0.10)
Training Database	**Testing Database**	**3D Dice ED**	**3D Dice ES**	**3D Hausdorff ED**	**3D Hausdorff ES**
LV-2011	ACDC	0.85	0.86	11.78	11.98
ACDC	ACDC	0.89	0.9	9.84	12.58

11.4 Generalization Ability of the Methods

The generalization ability is another important issue to be highlighted. In most of the analyzed studies, the methods are evaluated using a single database, which complicate knowing how the methods can be generalized to function on other databases, given the great variation observed in the exams.

The generalization problem can be observed in approaches that need annotated examples, such as those based on atlas and AI. Table 9 presents the results achieved when alternating training and testing between different databases for a hybrid method of atlas and AI [122] and only AI [45]. As observed, there is a decrease in the overlap metrics, indicating that the methods cannot generalize the problem and are restricted to the databases used. However, the Hausdorff distance reported is slightly shorter when there is alternation between databases, indicating that notwithstanding the difference in databases, the methods are still able to produce segmentations close to the ventricular region.

The arrival of AI methods observed in recent years, mainly driven by technology evolution and the availability of DL techniques, has been considered as the great promise for solving segmentation problems in several domains. Part of the justification for such a promise is the fact that DL frees the developer from the feature extraction phase, which is crucial for the efficient functioning of classic AI methods. It appears, however, that the large number of examples required for training these methods is still a limiting factor for the complete success of these approaches. In addition, the flexibility of these approaches, made possible by the large set of parameter combinations, can be an additional difficulty factor for their application, limiting the generalization ability of the methods. Finally, the variability of CMR images (Section 10), as already mentioned, imposes another barrier for generalizing the results.

A factor that can hinder the evaluation of generalization ability is the segmentation bias related to the database used in the training of AI methods, obtaining of atlases, or adjusting cost/energy functions in DM and graph approaches. Given that the segmentations in a database are usually done by few experts and may follow specific protocols, the methods can become biased. In the work of Bernard et al. [11], it is shown that intra- and inter-observer variability can be sometimes even greater than between experts and automatic methods. This can create uncertainty about segmentation performance. Although applied in a different domain (prostate zonal segmentation), an example of the discussion on this matter is presented in the work of Liu et al. [62].

11.5   General Critical Analysis

Based on the set of analyzed studies, it was possible to perceive the following:

- Processing stages are variable, but the success of the segmentation is highly dependent on the correct ROI extraction method. In many of the analyzed studies, this stage is still manual. Although almost half currently do it in an automated way, several methods still require some initialization by the user. This indicates that fully automatic methods for ROI extraction still need to be explored to both improve the accuracy of segmentation and decrease the human effort employed in this task.
- As mentioned in Section 11.3, even though it is not considered in some evaluation procedures, the automatic detection of valid slices is important to calculate clinic metrics. However, in most studies, it is assumed that all images contain valid slices and the LV is always present, or it is expected that the method will not make segmentations in extra slices. Thus, automatic approaches to detect the valid segmentation interval is a topic that needs to be further explored.
- Considering the segmentation stage, the AI category was the one with the highest number of studies included in the past decade, in parallel with DM. Although DL is a recent trend, the widespread application of this approach still requires the developer to define architectures and loss functions. In addition, the limited amount of available segmented examples impacts the performance of DL in specific cases, such as the presence of heart disease. The construction of new architectures, training strategies, and data augmentation approaches can produce research objects that have great potential for scientific contribution in the medical image segmentation area and in image processing as a whole.
- Few studies perform post- and pre-processing stages. The definition of specific methods for these stages is essential to reduce the high variability observed in CMR images, the presence of errors, and the complexity of the main segmentation method.
- The past decade has shown a raising availability of public databases in this area, driven mainly by competitions proposed to researchers. However, the different image characteristics, as well as the differences in the format and type of information provided in each database, still make it difficult to compare the approaches. The standardization of information and the availability of algorithms for calculating metrics are topics that can be further explored.
- The reproducibility of the approaches is an underexplored topic. As show in Section 11.4, the performance decreases when training and testing are done in different databases, with different image variations. Since most studies are adjusted to and only perform evaluation in one database (Table 3), it is not clear whether the same approaches can work in databases containing different varieties of cases and pathologies. Benchmarks and analysis can provide more evidence about the reproducibility of these approaches.
- Even though all approaches show promising results, they still have limitations. There is no method category that stands out in all situations considering the high variability of images. Thus, considering clinical practice, no method can produce reliable segmentation automatically for all slices in every type of patient cases. The development of hybrid approaches that combine the strengths of each method category, which have been little explored over the past decade, offers a promising path for the area.

12   LIMITATIONS OF THE CURRENT STUDY

Considering the systematic mapping presented in this article, the following limitations must be highlighted:

- *Searched databases*: Although we searched the main databases that index studies in the health and computer applications areas, relevant studies exclusively published in other databases could be missing in the analysis (e.g., [28, 68]). In the work of Dakua and Sahambi [28], a cantilever beam is used with random walks to segment the LV, whereas in the work of Mo et al. [68], Poincaré maps are used. This is always a limitation in systematic mapping, as the scope should be limited so as to become the research viable.
- *Hybrid methods*: In this work, a hybrid method is any approach that combines methods from two or more different segmentation categories (Section 7.8). Our classification considers only the segmentation stage (Figure 4) and not the ROI extraction stage. Thus, even if a study uses methods from other categories in the ROI extraction stage, it is not considered hybrid (e.g., [39, 45]). If other stages were considered, different interpretation and analysis of the presented results could be achieved.
- *Influence of other stages on the results*: In Sections 11.1 and 11.2, only the segmentation stage is considered in the limitations and performance analysis. Although the segmentation stage is the most important and the only one mandatory, it is not clear how much the results can be influenced by the other stages. For example, poor results in apical slices can be related to a poorly extracted ROI, and generalization performance may be influenced by pre-processing methods. Given this uncertainty, it is assumed that the main segmentation methods are responsible for the obtained results. The influence of other stages on the segmentation quality is a topic that needs to be further explored.
- *Other image types*: In this mapping, we focus on short-axis CMR images. Many other image types and views not included in the present article can be used in the segmentation of the LV, such as computerized tomography [121, 132], echocardiogram [55, 123], and long-axis CMR [51, 93]. In addition, specific CMR images, such as late gadolinium enhancement CMR used in segmentation of cardiac fibrosis and scars were not explored [114]. Although many approaches presented here can be applied in different image modalities, such as DL, methods that use specific information about images, such as DM and graphs, require different cost and energy functions, being applied in a particular image type. Again, adequate analysis requires a scope delimitation in systematic mapping, and the analysis presented in this current study cannot be extended to studies outside the scope considered here.
- *Segmentation of other cardiac regions and other body parts*: Many methods presented in the included works can be adapted and applied to segment other cardiac regions, such as the right ventricle [11] and atria [19, 120]. However, this work focuses on the LV segmentation. Given that the accurate segmentation of other structures is important in diagnosis, this is a topic that must be further explored. Likewise, articles that segment other parts of the human body (e.g., [27]) were not considered.
- *New studies not included in this review*: Given that the LV segmentation field is very active, new studies were published after the period considered in the search during this review, including new energy terms for level set approaches that impose shape restrictions [94] and new DL network architectures that explore temporal consistency [48, 113]. Although not included in the analysis, it is observed that these studies follow tendencies pointed out in this article.

13 CONCLUSION

In this review, a systematic mapping about LV segmentation methods in CMR images is presented. Dividing the LV segmentation process into stages (pre-processing, ROI extraction, segmentation, and post-processing) allowed analyzing the most used techniques and the results achieved in each processing stage.

The analysis of results for the Sunnybrook public database, the most used in the analyzed studies, indicates that methods of all categories have attained satisfactory performance, considering both semi-automatic and fully automatic approaches. Among them, DM and DL methods stand out. Difficulties in the construction of energy functions for DMs and generation of anatomically impossible segmentations of DL can be solved with the addition of shape priors and the creation of hybrid methods.

Despite the good results, there is still uncertainty about the generalization ability of methods. The variability of exams and evaluation methods makes it difficult to properly compare approaches. However, with the increased number of public databases available and the trend toward standardization of metrics, it is expected that more accurate and coherent evaluations can be carried out in the future. Furthermore, the high number of annotated examples favor the creation of new AI methods, atlases, and the construction of strong shape priors, but without abandoning precise methods like DM.

Considering the results presented in the analyzed studies and the topics covered in this review, there is considerable space for contributions to this area, especially with regard to the construction of hybrid methods, proposing new evaluation metrics, providing adequate databases, and algorithms that can take into account the range of characteristics that are included in the medical images used in the area.

REFERENCES

[1] Ahmed O. Al-Agamy, Ayman Khalifa, and Ahmed S. Fahmy. 2014. Greedy framework for optical flow tracking of myocardium contours. *IET Image Processing* 8, 4 (April 2014), 207–212.

[2] Xènia Albá, Marco Pereañez, Corné Hoogendoorn, Andrew J. Swift, Jim M. Wild, Alejandro F. Frangi, and Karim Lekadir. 2016. An algorithm for the segmentation of highly abnormal hearts using a generic statistical shape model. *IEEE Transactions on Medical Imaging* 35, 3 (March 2016), 845–859.

[3] Alexander Andreopoulos and John K. Tsotsos. 2008. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Medical Image Analysis* 12, 3 (2008), 335–357.

[4] Cristobal Arrieta, Sergio Uribe, Carlos Sing-Long, Daniel Hurtado, Marcelo Andia, Pablo Irarrazaval, and Cristian Tejos. 2017. Simultaneous left and right ventricle segmentation using topology preserving level sets. *Biomedical Signal Processing and Control* 33 (2017), 88–95.

[5] Michael R. Avendi, Arash Kheradvar, and Hamid Jafarkhani. 2016. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical Image Analysis* 30 (2016), 108–119.

[6] Ismail Ben Ayed, Hua-Mei Chen, Kumaradevan Punithakumar, Ian Ross, and Shuo Li. 2012. Max-flow segmentation of the left ventricle by recovering subject-specific distributions via a bound of the Bhattacharyya measure. *Medical Image Analysis* 16, 1 (2012), 87–100.

[7] Wenjia Bai, Wenzhe Shi, Christian Ledig, and Daniel Rueckert. 2015. Multi-atlas segmentation with augmented features for cardiac MR images. *Medical Image Analysis* 19, 1 (2015), 98–109.

[8] Wenjia Bai, Wenzhe Shi, Declan P. O'Regan, Tong Tong, Haiyan Wang, Shahnaz Jamil-Copley, Nicholas S. Peters, and Daniel Rueckert. 2013. A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images. *IEEE Transactions on Medical Imaging* 32, 7 (July 2013), 1302–1315.

[9] Emelia J. Benjamin, Paul Muntner, Alvaro Alonso, Marcio S. Bittencourt, Clifton W. Callaway, April P. Carson, Alanna M. Chamberlain, et al. 2019. Heart disease and stroke statistics—2019 update: A report from the American Heart Association. *Circulation* 139, 10 (Jan. 2019), e56–e528.

[10] Leila C. C. Bergamasco, Carlos E. Rochitte, and Fátima L. S. Nunes. 2018. 3D medical objects processing and retrieval using spherical harmonics: A case study with congestive heart failure MRI exams. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC'18)*. ACM, New York, NY, 22–29.

[11] Olivier Bernard, Alain Lalande, Clement Zotti, Frederick Cervenansky, Xin Yang, Pheng-Ann Heng, Irem Cetin, et al. 2018. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Transactions on Medical Imaging* 37, 11 (Nov. 2018), 2514–2525.

[12] John Martin Bland and Douglas G. Altman. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 8476 (1986), 307–310.

[13] Eugene Braunwald. 2013. Heart failure. *JACC: Heart Failure* 1, 1 (2013), 1–20.

[14] Monika N. Bugdol and Ewa Pietka. 2015. Mathematical model in left ventricle segmentation. *Computers in Biology and Medicine* 57 (2015), 187–200.

[15] Yunliang Cai, Ali Islam, Mousumi Bhaduri, Ian Chan, and Shuo Li. 2016. Unsupervised freeview groupwise cardiac segmentation using synchronized spectral network. *IEEE Transactions on Medical Imaging* 35, 9 (Sept. 2016), 2174–2188.

[16] Jérôme Caudron, Jeannette Fares, Fabrice Bauer, and Jean-Nicolas Dacher. 2011. Evaluation of left ventricular diastolic function with cardiac MR imaging. *RadioGraphics* 31, 1 (2011), 239–259.

[17] Tony F. Chan and Luminita A. Vese. 2001. Active contours without edges. *IEEE Transactions on Image Processing* 10, 2 (2001), 266–277.

[18] Sadegh Charmchi, Kumaradevan Punithakumar, and Pierre Boulanger. 2018. Optimizing U-Net to segment left ventricle from magnetic resonance imaging. In *Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM'18)*. IEEE, Los Alamitos, CA, 327–332.

[19] Jun Chen, Guang Yang, Habib Khan, Heye Zhang, Yanping Zhang, Shu Zhao, Raad Mohiaddin, Tom Wong, David Firmin, and Jennifer Keegan. 2021. JAS-GAN: Generative adversarial network based joint atrium and scar segmentation on unbalanced atrial targets. *IEEE Journal of Biomedical and Health Informatics* PP, 1 (2021), 1.

[20] Chao Cong and Hongmin Zhang. 2018. Invert-U-Net DNN segmentation model for MRI cardiac left ventricle segmentation. *Journal of Engineering* 2018, 16 (2018), 1463–1467.

[21] Lucilio Cordero-Grande, Gonzalo Vegas-Sánchez-Ferrero, Pablo Casaseca-de-la-Higuera, José Alberto San-Román-Calvar, Ana Revilla-Orodea, Marcos Martín-Fernández, and Carlos Alberola-López. 2011. Unsupervised 4D myocardium segmentation with a Markov random field based deformable model. *Medical Image Analysis* 15, 3 (2011), 283–301.

[22] Ariel H. Curiale, Flavio D. Colavecchia, Pablo Kaluza, Roberto A. Isoardi, and German Mato. 2017. Automatic myocardial segmentation by using a deep learning network in cardiac MRI. In *Proceedings of the 2017 XLIII Latin American Computer Conference (CLEI'17)*. IEEE, Los Alamitos, CA, 1–6.

[23] Ariel H. Curiale, Flavio D. Colavecchia, and German Mato. 2019. Automatic quantification of the LV function and mass: A deep learning approach for cardiovascular MRI. *Computer Methods and Programs in Biomedicine* 169 (2019), 37–50.

[24] Sarada P. Dakua. 2011. Performance divergence with data discrepancy: A review. *Artificial Intelligence Review* 40, 4 (Dec. 2011), 429–455.

[25] Sarada P. Dakua. 2014. AnnularCut: A graph-cut design for left ventricle segmentation from magnetic resonance images. *IET Image Processing* 8, 1 (January 2014), 1–11.

[26] Sarada P. Dakua. 2017. Towards left ventricle segmentation from magnetic resonance images. *IEEE Sensors Journal* 17, 18 (Sept. 2017), 5971–5981.

[27] Sarada P. Dakua, Julien Abinahed, and Abdulla A. Al-Ansari. 2016. Pathological liver segmentation using stochastic resonance and cellular automata. *Journal of Visual Communication and Image Representation* 34 (Jan. 2016), 89–102.

[28] Sarada P. Dakua and J. S. Sahambi. 2010. Automatic left ventricular contour extraction from cardiac magnetic resonance images using cantilever beam and random walk approach. *Cardiovascular Engineering* 10, 1 (Jan. 2010), 30–43.

[29] Xiuquan Du, Renjun Tang, Susu Yin, Yanping Zhang, and Shuo Li. 2019. Direct segmentation-based full quantification for left ventricle via deep multi-task regression learning network. *IEEE Journal of Biomedical and Health Informatics* 23, 3 (May 2019), 942–948.

[30] Xiuquan Du, Susu Yin, Renjun Tang, Yanping Zhang, and Shuo Li. 2019. Cardiac-DeepIED: Automatic pixel-level deep segmentation for cardiac bi-ventricle using improved end-to-end encoder-decoder network. *IEEE Journal of Translational Engineering in Health and Medicine* 7 (2019), 1–10.

[31] Xiuquan Du, Weiwei Zhang, Heye Zhang, Jun Chen, Yanping Zhang, James Claude Warrington, Gary Brahm, and Shuo Li. 2018. Deep regression segmentation for cardiac bi-ventricle MR images. *IEEE Access* 6 (2018), 3828–3838.

[32] Abouzar Eslami, Athanasios Karamalis, Amin Katouzian, and Nassir Navab. 2013. Segmentation by retrieval with guided random walks: Application to left ventricle segmentation in MRI. *Medical Image Analysis* 17, 2 (Feb. 2013), 236–253.

[33] Paul W. X. Foley, Kayvan Khadjooi, Joseph A. Ward, Russell E. A. Smith, Berthold Stegemann, Michael P. Frenneaux, and Francisco Leyva. 2009. Radial dyssynchrony assessed by cardiovascular magnetic resonance in relation to left ventricular function, myocardial scarring and QRS duration in patients with heart failure. *Journal of Cardiovascular Magnetic Resonance* 11, 1 (2009), 50.

[34] Morie A. Gertz, Angela Dispenzieri, and Taimur Sher. 2014. Pathophysiology and treatment of cardiac amyloidosis. *Nature Reviews Cardiology* 12, 2 (Oct. 2014), 91–102.

[35] Ayush Goyal. 2019. Image-based clustering and connected component labeling for rapid automated left and right ventricular endocardial volume extraction and segmentation in full cardiac cycle multi-frame MRI images of cardiac patients. *Medical & Biological Engineering & Computing* 57, 6 (June 2019), 1213–1228.

[36] Damien Grosgeorge, Caroline Petitjean, and Su Ruan. 2016. Multilabel statistical shape prior for image segmentation. *IET Image Processing* 10, 10 (2016), 710–716.

[37] Huaifei Hu, Zhiyong Gao, Liman Liu, Haihua Liu, Junfeng Gao, Shengzhou Xu, Wei Li, and Lu Huang. 2014. Automatic segmentation of the left ventricle in cardiac MRI using local binary fitting model and dynamic programming techniques. *PLoS ONE* 9, 12 (Dec. 2014), 1–17.

[38] Huaifei Hu, Haihua Liu, Zhiyong Gao, and Lu Huang. 2013. Hybrid segmentation of left ventricle in cardiac MRI using Gaussian-mixture model and region restricted dynamic programming. *Magnetic Resonance Imaging* 31, 4 (2013), 575–584.

[39] Huaifei Hu, Ning Pan, Jiayu Wang, Tailang Yin, and Renzhen Ye. 2019. Automatic segmentation of left ventricle from cardiac MRI via deep learning and region constrained dynamic programming. *Neurocomputing* 347 (2019), 139–148.

[40] Dongwoo Kang, Jonghye Woo, C. C. Jay Kuo, Piotr J. Slomka, Damini Dey, and Guido Germano. 2012. Heart chambers and whole heart segmentation techniques: Review. *Journal of Electronic Imaging* 21, 1 (2012), 1–17.

[41] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. 1988. Snakes: Active contour models. *International Journal of Computer Vision* 1, 4 (Jan. 1988), 321–331.

[42] Fahmi Khalifa, Garth M. Beache, Georgy Gimelrfarb, Guruprasad A. Giridharan, and Ayman El-Baz. 2012. Accurate automatic analysis of cardiac cine images. *IEEE Transactions on Biomedical Engineering* 59, 2 (Feb. 2012), 445–455.

[43] Mohammad-Bagher Khamechian and Mahdi Saadatmand-Tarzjan. 2013. A new patch-based active contour for segmentation of the myocardium of the left ventricle in cardiac magnetic resonance images. In *Proceedings of the International eConference on Computer and Knowledge Engineering (ICCKE'13)*. IEEE, Los Alamitos, CA, 223–228.

[44] Mohammad-Bagher Khamechian and Mahdi Saadatmand-Tarzjan. 2018. FoCA: A new framework of coupled geometric active contours for segmentation of 3D cardiac magnetic resonance images. *Magnetic Resonance Imaging* 51 (2018), 51–60.

[45] Mahendra Khened, Varghese Alex Kollerathu, and Ganapathy Krishnamurthi. 2019. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Medical Image Analysis* 51 (2019), 21–45.

[46] Lisa M. Koch, Martin Rajchl, Wenjia Bai, Christian F. Baumgartner, Tong Tong, Jonathan Passerat-Palmbach, Paul Aljabar, and Daniel Rueckert. 2018. Multi-atlas segmentation using partially annotated data: Methods and annotation strategies. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40, 7 (July 2018), 1683–1696.

[47] Ismael Koné and Lahsen Boulmane. 2018. Hybrid forests for left ventricle segmentation using only the first slice label. In *Proceedings of the 2018 International Conference on Intelligent Systems and Computer Vision (ISCV'18)*. IEEE, Los Alamitos, CA, 1–5.

[48] Mengmeng Kuang, Yinzhe Wu, Diego Alonso-Álvarez, David Firmin, Jennifer Keegan, Peter Gatehouse, and Guang Yang. 2021. Three-dimensional embedded attentive RNN (3D-EAR) segmentor for left ventricle delineation from myocardial velocity mapping. In *Functional Imaging and Modeling of the Heart*. Springer International, Cham, Switzerland, 55–62.

[49] Rainer P. Kunz, Florian Oellig, Frank Krummenauer, Katja Oberholzer, Bernd Romaneehsen, Toni W. Vomweg, Georg Horstick, Carmel Hayes, Manfred Thelen, and Karl-Friedrich Kreitner. 2005. Assessment of left ventricular function by breath-hold cine MR imaging: Comparison of different steady-state free precession sequences. *Journal of Magnetic Resonance Imaging* 21, 2 (2005), 140–148.

[50] Hae-Yeoun Lee, Noel C. F. Codella, Matthew D. Cham, Jonathan W. Weinsaft, and Yi Wang. 2010. Automatic left ventricle segmentation using iterative thresholding and an active contour model with adaptation on short-axis cardiac MRI. *IEEE Transactions on Biomedical Engineering* 57, 4 (April 2010), 905–913.

[51] Shuang Leng, Xulei Yang, Xiaodan Zhao, Zeng Zeng, Yi Su, Angela S. Koh, David Sim, Ju Le Tan, Ru San Tan, and Liang Zhong. 2018. Computational platform based on deep learning for segmenting ventricular endocardium in long-axis cardiac MR imaging. In *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'18)*. IEEE, Los Alamitos, CA.

[52] Chunming Li, Rui Huang, Zhaohua Ding, J. C. Gatenby, Dimitris N. Metaxas, and John C. Gore. 2011. A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE Transactions on Image Processing* 20, 7 (2011), 2007–2016.

[53] Chunming Li, Chiu-Yen Kao, John C. Gore, and Zhaohua Ding. 2008. Minimization of region-scalable fitting energy for image segmentation. *IEEE Transactions on Image Processing* 17, 10 (2008), 1940–1949.

[54] Chunming Li, Chenyang Xu, Changfeng Gui, and Martin D. Fox. 2010. Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing* 19, 12 (2010), 3243–3254.

[55] Ming Li, Chengjia Wang, Heye Zhang, and Guang Yang. 2020. MV-RAN: Multiview recurrent aggregation network for echocardiographic sequences segmentation and full cardiac cycle analysis. *Computers in Biology and Medicine* 120 (May 2020), 103728.

[56] Fangzhou Liao, Xi Chen, Xiaolin Hu, and Sen Song. 2019. Estimation of the volume of the left ventricle from MRI images using deep neural networks. *IEEE Transactions on Cybernetics* 49, 2 (Feb. 2019), 495–504.

[57] Adan Lin, Junhao Wu, and Xuan Yang. 2020. A data augmentation approach to train fully convolutional networks for left ventricle segmentation. *Magnetic Resonance Imaging* 66 (2020), 152–164.

[58] Geert Litjens, Francesco Ciompi, Jelmer M. Wolterink, Bob D. de-Vos, Tim Leiner, Jonas Teuwen, and Ivana Išgum. 2019. State-of-the-art deep learning in cardiovascular image analysis. *JACC: Cardiovascular Imaging* 12, 8, Part 1 (2019), 1549–1565.

[59] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. Laak, Bram Ginneken, and Clara I. Sánchez. 2017. A survey on deep learning in medical image analysis. *Medical Image Analysis* 42 (2017), 60–88.

[60] Hong Liu, Huaifei Hu, Xiangyang Xu, and Enmin Song. 2012. Automatic left ventricle segmentation in cardiac MRI using topological stable-state thresholding and region restricted dynamic programming. *Academic Radiology* 19, 6 (2012), 723–731.

[61] Yu Liu, Gabriella Captur, James C. Moon, Shuxu Guo, Xiaoping Yang, Shaoxiang Zhang, and Chunming Li. 2016. Distance regularized two level sets for segmentation of left and right ventricles from cine-MRI. *Magnetic Resonance Imaging* 34, 5 (2016), 699–706.

[62] Yongkai Liu, Guang Yang, Melina Hosseiny, Afshin Azadikhah, Sohrab Afshari Mirak, Qi Miao, Steven S. Raman, and Kyunghyun Sung. 2020. Exploring uncertainty measures in Bayesian deep attentive neural networks for prostate zonal segmentation. *IEEE Access* 8 (2020), 151817–151828.

[63] Yurun Ma, Li Wang, Yide Ma, Min Dong, Shiqiang Du, and Xiaoguang Sun. 2016. An SPCNN-GVF-based approach for the automatic segmentation of left ventricle in cardiac cine MR images. *International Journal of Computer Assisted Radiology and Surgery* 11, 11 (June 2016), 1951–1964.

[64] Dwarikanath Mahapatra. 2013. Cardiac image segmentation from cine cardiac MRI using graph cuts and shape priors. *Journal of Digital Imaging* 26, 4 (Jan. 2013), 721–730.

[65] Dwarikanath Mahapatra. 2013. Cardiac MRI segmentation using mutual context information from left and right ventricle. *Journal of Digital Imaging* 26, 5 (Jan. 2013), 898–908.

[66] Elaine Nicpon Marieb, Patricia Brady Wilhelm, and Jon Mallatt. 2012. *Human Anatomy* (6th ed., media update ed.). Benjamin Cummings, Boston, MA.

[67] Donald W. McRobbie, Elizabeth A. Moore, Martin J. Graves, and Martin R. Prince. 2017. *MRI from Picture to Proton* (3rd ed.). Cambridge University Press, New York, NY.

[68] Yuanhan Mo, Fangde Liu, Douglas McIlwraith, Guang Yang, Jingqing Zhang, Taigang He, and Yike Guo. 2018. The deep Poincaré map: A novel approach for left ventricle segmentation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Springer International, Cham, Switzerland, 561–568.

[69] Laurent Najman, Jean Cousty, Michel Couprie, Hugues Talbot, Stéphanie Clément-Guinaudeau, and Thomas Goissen. 2007. An open, clinically-validated database of 3D+t cine-MR images of the left ventricle with associated manual and automated segmentation. *Insight Journal* 2007 (Jan. 2007), 1–9.

[70] Mohammad Saleh Nambakhsh, Jing Yuan, Ismail Ben Ayed, Kumaradevan Punithakumar, Aashish Goela, Ali Islam, Terry Peters, and Shuo Li. 2011. A convex max-flow segmentation of LV using subject-specific distributions on cardiac MRI. In *Information Processing in Medical Imaging*. Springer, Berlin, Germany, 171–183.

[71] Tuan A. Ngo and Gustavo Carneiro. 2013. Left ventricle segmentation from cardiac MRI combining level set methods with deep belief networks. In *Proceedings of the 2013 IEEE International Conference on Image Processing*. IEEE, Los Alamitos, CA, 695–699.

[72] Tuan A. Ngo and Gustavo Carneiro. 2014. Fully automated non-rigid segmentation with distance regularized level set evolution initialized and constrained by deep-structured inference. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Los Alamitos, CA, 3118–3125.

[73] Tuan A. Ngo, Zhi Lu, and Gustavo Carneiro. 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 (2017), 159–171.

[74] R. M. Norris, H. D. White, D. B. Cross, C. J. Wild, and R. M. L. Whitlock. 1992. Prognosis after recovery from myocardial infarction: The relative importance of cardiac dilatation and coronary stenoses. *European Heart Journal* 13, 12 (Dec. 1992), 1611–1618.

[75] Stephen P. O'Brien, Ovidiu Ghita, and Paul F. Whelan. 2011. A novel model-based 3D + time left ventricular segmentation technique. *IEEE Transactions on Medical Imaging* 30, 2 (Feb. 2011), 461–474.

[76] Peng Peng, Karim Lekadir, Ali Gooya, Ling Shao, Steffen E. Petersen, and Alejandro F. Frangi. 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, 2 (Jan. 2016), 155–195.

[77] Dudley J. Pennell, S. Richard Underwood, Carla C. Manzara, R. Howard Swanton, J. Malcolm Walker, Peter J. Ell, and Donald B. Longmore. 1992. Magnetic resonance imaging during dobutamine stress in coronary artery disease. *American Journal of Cardiology* 70, 1 (1992), 34–40.

[78] Caroline Petitjean and Jean-Nicolas Dacher. 2011. A review of segmentation methods in short axis cardiac MR images. *Medical Image Analysis* 15, 2 (2011), 169–184.

[79] Van-Truong Pham and Thi-Thao Tran. 2016. Active contour model and nonlinear shape priors with application to left ventricle segmentation in cardiac MR images. *Optik* 127, 3 (2016), 991–1002.

[80] Charnchai Pluempitiwiriyawej, José M. F. Moura, Yi-Jen Lin Wu, and Chien Ho. 2005. STACS: New active contour scheme for cardiac MR image segmentation. *IEEE Transactions on Medical Imaging* 24, 5 (May 2005), 593–603.

[81] Kumaradevan Punithakumar, Ismail Ben Ayed, Mariam Afshin, Aashish Goela, Ali Islam, Shuo Li, Pierre Boulanger, Harald Becher, and Michelle Noga. 2016. Detecting left ventricular impaired relaxation in cardiac MRI using moving mesh correspondences. *Computer Methods and Programs in Biomedicine* 124 (2016), 58–66.

[82] Kumaradevan Punithakumar, Ismail B. Ayed, Ali Islam, Ian G. Ross, and Shuo Li. 2010. Tracking endocardial motion via multiple model filtering. *IEEE Transactions on Biomedical Engineering* 57, 8 (Aug. 2010), 2001–2010.

[83] Xianjing Qin, Yan Tian, and Pingkun Yan. 2015. Feature competition and partial sparse shape modeling for cardiac image sequences segmentation. *Neurocomputing* 149 (2015), 904–913.

[84] Perry Radau, Yingli Lu, Kim Connelly, Gideon Paul, Alexander J. Dick, and Graham A. Wright. 2009. Evaluation framework for algorithms segmenting short axis cardiac MRI. In *Proceedings of the MICCAI 2009 Workshop: Cardiac MR left ventricle segmentation challenge*. http://hdl.handle.net/10380/3070.

[85] Liset Vázquez Romaguera, Francisco Perdigón Romero, Cicero Ferreira Fernandes Costa Filho, and Marly Guimarães Fernandes Costa. 2018. Myocardial segmentation in cardiac magnetic resonance images using fully convolutional neural networks. *Biomedical Signal Processing and Control* 44 (2018), 48–57.

[86] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi (Eds.). Springer International, Cham, Switzerland, 234–241.

[87] Andrea Salzano, Alberto M. Marra, Roberta D'Assante, Michele Arcopinto, Eduardo Bossone, Toru Suzuki, and Antonio Cittadini. 2019. Biomarkers and imaging: Complementary or subtractive? *Heart Failure Clinics* 15, 2 (2019), 321–331.

[88] Carlos Santiago, Jacinto C. Nascimento, and Jorge S. Marques. 2017. Fast and accurate segmentation of the LV in MR volumes using a deformable model with dynamic programming. In *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP'17)*. IEEE, Los Alamitos, CA, 1747–1751.

[89] Carlos Santiago, Jacinto C. Nascimento, and Jorge S. Marques. 2018. Fast segmentation of the left ventricle in cardiac MRI using dynamic programming. *Computer Methods and Programs in Biomedicine* 154 (2018), 9–23.

[90] Nicolo Savioli, Miguel S. Vieira, Pablo Lamata, and Giovanni Montana. 2018. Automated segmentation on the entire cardiac cycle using a deep learning work-flow. In *Proceedings of the 2018 5th International Conference on Social Networks Analysis, Management, and Security (SNAMS'18)*. IEEE, Los Alamitos, CA, 153–158.

[91] Joël Schaerer, Christopher Casta, Jérôme Pousin, and Patrick Clarysse. 2010. A dynamic elastic model for segmentation and tracking of the heart in MR image sequences. *Medical Image Analysis* 14, 6 (2010), 738–749.

[92] Udo Sechtem, Frank M. Baer, Eli Voth, P. Theissen, and C. A. Schneider. 1999. Stress functional MRI: Detection of ischemic heart disease and myocardial viability. *Journal of Magnetic Resonance Imaging* 10, 5 (1999), 667–675.

[93] Rahil Shahzad, Qian Tao, Oleh Dzyubachyk, Marius Staring, Boudewijn P. F. Lelieveldt, and Rob J. van der Geest. 2017. Fully-automatic left ventricular segmentation from long-axis cardiac cine MR scans. *Medical Image Analysis* 39 (July 2017), 44–55.

[94] Xue Shi and Chunming Li. 2021. Convexity preserving level set for left ventricle segmentation. *Magnetic Resonance Imaging* 78 (2021), 109–118.

[95] Shafiullah Soomro, Farhan Akram, Asad Munir, Chang Ha Lee, and Kwang Nam Choi. 2017. Segmentation of left and right ventricles in cardiac MRI using active contours. *Computational and Mathematical Methods in Medicine* 2017 (2017), 1–16.

[96] Avan Suinesiaputra, Brett R. Cowan, Ahmed O. Al-Agamy, Mustafa A. Elattar, Nicholas Ayache, Ahmed S. Fahmy, Ayman M. Khalifa, et al. 2014. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. *Medical Image Analysis* 18, 1 (Jan. 2014), 50–62.

[97] Avan Suinesiaputra, Brett R. Cowan, Ahmed O. Al-Agamy, Mustafa A. Elattar, Nicholas Ayache, Ahmed S. Fahmy, Ayman M. Khalifa, et al. 2014. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. *Medical Image Analysis* 18, 1 (2014), 50–62.

[98] Li Kuo Tan, Yih Miin Liew, Einly Lim, and Robert A. McLaughlin. 2017. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. *Medical Image Analysis* 39 (2017), 78–86.

[99] Qian Tao, Wenjun Yan, Yuanyuan Wang, Elisabeth H. M. Paiman, Denis P. Shamonin, Pankaj Garg, et al. 2019. Deep learning–based method for fully automatic quantification of left ventricle function from cine MR images: A multi-vendor, multicenter study. *Radiology* 290, 1 (2019), 81–88.

[100] Qianqian Tong, Caizi Li, Weixin Si, Xiangyun Liao, Yaliang Tong, Zhiyong Yuan, and Pheng Ann Heng. 2019. RIANet: Recurrent interleaved attention network for cardiac MRI segmentation. *Computers in Biology and Medicine* 109 (2019), 290–302.

[101] Qianqian Tong, Zhiyong Yuan, Xiangyun Liao, Mianlun Zheng, Weixu Zhu, Guian Zhang, and Munan Ning. 2017. A joint multi-scale convolutional network for fully automatic segmentation of the left ventricle. In *Proceedings of the IEEE International Conference on Image Processing (ICIP'17)*. IEEE, Los Alamitos, CA, 3110–3114.

[102] Jane Tufvesson, Erik Hedström, Katarina Steding-Ehrenborg, Marcus Carlsson, Håkan Arheden, and Einar Heiberg. 2015. Validation and development of a new automatic algorithm for time-resolved segmentation of the left ventricle in magnetic resonance imaging. *BioMed Research International* 2015 (2015), 1–12.

[103] Davis M. Vigneault, Weidi Xie, Carolyn Y. Ho, David A. Bluemke, and J. Alison Noble. 2018. Ω-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks. *Medical Image Analysis* 48 (2018), 95–106.

[104] Stephen Waite, Srinivas Kolla, Jean Jeudy, Alan Legasto, Stephen L. Macknik, Susana Martinez-Conde, Elizabeth A. Krupinski, and Deborah L. Reede. 2017. Tired in the reading room: The influence of fatigue in radiology. *Journal of the American College of Radiology* 14, 2 (Feb. 2017), 191–197.

[105] Li Wang, Yurun Ma, Kun Zhan, and Yide Ma. 2015. Automatic left ventricle segmentation in cardiac MRI via level set and fuzzy C-means. In *Proceedings of the 2015 2nd International Conference on Recent Advances in Engineering Computational Sciences (RAECS'15)*. IEEE, Los Alamitos, CA, 1–6.

[106] Lijia Wang, Mengchao Pei, Noel C. F. Codella, Minisha Kochar, Jonathan W. Weinsaft, Jianqi Li, Martin R. Prince, and Yi Wang. 2015. Left ventricle: Fully automated segmentation based on spatiotemporal continuity and myocardium information in cine cardiac magnetic resonance imaging (LV-FAST). *BioMed Research International* 2015 (2015), 1–9.

[107] Yan Wang, Yue Zhang, Wanling Xuan, Evan Kao, Peng Cao, Bing Tian, Karen Ordovas, David Saloner, and Jing Liu. 2019. Fully automatic segmentation of 4D MRI for cardiac functional measurements. *Medical Physics* 46, 1 (2019), 180–189.

[108] Zhen Zhou Wang. 2017. Segmentation of the left ventricle in short-axis sequences by combining deformation flow and optical flow. *IET Image Processing* 11, 4 (2017), 237–244.

[109] Harvey White, R. M. Norris, Michael Brown, P. W. Brandt, R. M. Whitlock, and Christopher Wild. 1987. Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation* 76 (July 1987), 44–51.

[110] Harvey White, R. M. Norris, Michael Brown, P. W. Brandt, R. M. Whitlock, and Christopher Wild. 1987. Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation* 76 (July 1987), 44–51.

[111] Jonghye Woo, Piotr J. Slomka, C. C. Jay Kuo, and Byung-Woo Hong. 2013. Multiphase segmentation using an implicit dual shape prior: Application to detection of left ventricle in cardiac MRI. *Computer Vision and Image Understanding* 117, 9 (2013), 1084–1094.

[112] World Health Organization. 2019. The Top 10 Causes of Death. Retrieved March 1, 2022 from https://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death.

[113] Yinzhe Wu, Suzan Hatipoglu, Diego Alonso-Álvarez, Peter Gatehouse, Binghuan Li, Yikai Gao, David Firmin, Jennifer Keegan, and Guang Yang. 2021. Fast and automated segmentation for the three-directional multi-slice cine myocardial velocity mapping. *Diagnostics* 11, 2 (Feb. 2021), 346.

[114] Yinzhe Wu, Zeyu Tang, Binghuan Li, David Firmin, and Guang Yang. 2021. Recent advances in fibrosis and scar segmentation from cardiac MRI: A state-of-the-art review and future perspectives. *Frontiers in Physiology* 12 (Aug. 2021), 86–108. DOI:https://doi.org/10.3389/fphys.2021.709230

[115] Yuwei Wu, Yuanquan Wang, and Yunde Jia. 2013. Segmentation of the left ventricle in cardiac cine MRI using a shape-constrained snake model. *Computer Vision and Image Understanding* 117, 9 (2013), 990–1003.

[116] Yan Xia, Nishant Ravikumar, John P. Greenwood, Stefan Neubauer, Steffen E. Petersen, and Alejandro F. Frangi. 2021. Super-resolution of cardiac MR cine imaging using conditional GANs and unsupervised transfer learning. *Medical Image Analysis* 71 (2021), 102037.

[117] Chenyang Xu and Jerry L. Prince. 1998. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing* 7, 3 (1998), 359–369.

[118] Cong Yang, Weiguo Wu, Yuanqi Su, and Shaoxiang Zhang. 2017. Left ventricle segmentation via two-layer level sets with circular shape constraint. *Magnetic Resonance Imaging* 38 (2017), 202–213.

[119] D. Yang, Q. Huang, L. Axel, and D. Metaxas. 2018. Multi-component deformable models coupled with 2D-3D U-Net for automated probabilistic segmentation of cardiac walls and blood. In *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI'18)*. IEEE, Los Alamitos, CA, 479–483.

[120] Guang Yang, Jun Chen, Zhifan Gao, Shuo Li, Hao Ni, Elsa Angelini, Tom Wong, Raad Mohiaddin, et al. 2020. Simultaneous left atrium anatomy and scar segmentations via deep learning in multiview information with attention. *Future Generation Computer Systems* 107 (June 2020), 215–228.

[121] Guanyu Yang, Yang Chen, Lijun Tang, Huazhong Shu, and Christine Toumoulin. 2014. Automatic left ventricle segmentation based on multiatlas registration in 4D CT images. In *Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI'14)*. IEEE, Los Alamitos, CA.

[122] Heran Yang, Jian Sun, Huibin Li, Lisheng Wang, and Zongben Xu. 2018. Neural multi-atlas label fusion: Application to cardiac MR images. *Medical Image Analysis* 49 (2018), 60–75.

[123] Meijun Yang, Xiaoyan Xiao, Zhi Liu, Longkun Sun, Wei Guo, Lizhen Cui, Dianmin Sun, Pengfei Zhang, and Guang Yang. 2020. Deep RetinaNet for dynamic left ventricle detection in multiview echocardiography classification. *Scientific Programming* 2020 (Aug. 2020), 1–6.

[124] Xulei Yang, Qing Song, and Yi Su. 2017. Automatic segmentation of left ventricle cavity from short-axis cardiac magnetic resonance images. *Medical & Biological Engineering & Computing* 55, 9 (Feb. 2017), 1563–1577.

[125] Xulei Yang, Yi Su, Rubing Duan, Haijin Fan, Si Y. Yeo, Calvin Lim, Liang Zhong, and Ru S. Tan. 2016. Cardiac image segmentation by random walks with dynamic shape constraint. *IET Computer Vision* 10, 1 (2016), 79–86.

[126] Xulei Yang, Zeng Zeng, and Su Yi. 2017. Deep convolutional neural networks for automatic segmentation of left ventricle cavity from cardiac magnetic resonance images. *IET Computer Vision* 11, 8 (2017), 643–649.

[127] Tianchen Yuan, Qianqian Tong, Xiangyun Liao, Xinling Du, and Jianhui Zhao. 2018. Fully automatic segmentation of the left ventricle using multi-scale fusion learning. In *Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR'18)*. IEEE, Los Alamitos, CA, 3838–3843.

[128] Qiao Zheng, Herve Delingette, Nicolas Duchateau, and Nicholas Ayache. 2018. 3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation. *IEEE Transactions on Medical Imaging* 37, 9 (Sept. 2018), 2137–2148.

[129] Xiahai Zhuang. 2013. Challenges and methodologies of fully automatic whole heart segmentation: A review. *Journal of Healthcare Engineering* 4, 3 (Sept. 2013), 371–408. https://doi.org/10.1260/2040-2295.4.3.371

[130] X. Zhuang, K. S. Rhode, R. S. Razavi, D. J. Hawkes, and S. Ourselin. 2010. A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *IEEE Transactions on Medical Imaging* 29, 9 (Sept. 2010), 1612–1625.

[131] Clement Zotti, Zhiming Luo, Alain Lalande, and Pierre-Marc Jodoin. 2019. Convolutional neural network with shape prior applied to cardiac MRI segmentation. *IEEE Journal of Biomedical and Health Informatics* 23, 3 (May 2019), 1119–1128.

[132] Majd Zreik, Tim Leiner, Bob D. de Vos, Robbert W. van Hamersvelt, Max A. Viergever, and Ivana Isgum. 2016. Automatic segmentation of the left ventricle in cardiac CT angiography using convolutional neural networks. In *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI'16)*. IEEE, Los Alamitos, CA.