City Research Online

City St George's, University of London

Citation: Chen, C., Qin, C., Qiu, H., Tarroni, G., Duan, J., Bai, W. & Rueckert, D. (2020). Deep Learning for Cardiac Image Segmentation: A Review. Frontiers in Cardiovascular Medicine, 7, 25. doi: 10.3389/fcvm.2020.00025

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/31853/

Link to published version: https://doi.org/10.3389/fcvm.2020.00025

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

frontiers in Cardiovascular Medicine

REVIEW
published: 05 March 2020
doi: 10.3389/fcvm.2020.00025

Deep Learning for Cardiac Image Segmentation: A Review

Chen Chen[1], Chen Qin[1], Huaqi Qiu[1], Giacomo Tarroni[1,2], Jinming Duan[3], Wenjia Bai[4,5] and Daniel Rueckert[1]*

[1] Biomedical Image Analysis Group, Department of Computing, Imperial College London, London, United Kingdom, [2] CitAI Research Centre, Department of Computer Science, City University of London, London, United Kingdom, [3] School of Computer Science, University of Birmingham, Birmingham, United Kingdom, [4] Data Science Institute, Imperial College London, London, United Kingdom, [5] Department of Brain Sciences, Faculty of Medicine, Imperial College London, London, United Kingdom

Deep learning has become the most widely used approach for cardiac image segmentation in recent years. In this paper, we provide a review of over 100 cardiac image segmentation papers using deep learning, which covers common imaging modalities including magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound and major anatomical structures of interest (ventricles, atria, and vessels). In addition, a summary of publicly available cardiac image datasets and code repositories are included to provide a base for encouraging reproducible research. Finally, we discuss the challenges and limitations with current deep learning-based approaches (scarcity of labels, model generalizability across different domains, interpretability) and suggest potential directions for future research.

Keywords: artificial intelligence, deep learning, neural networks, cardiac image segmentation, cardiac image analysis, MRI, CT, ultrasound

OPEN ACCESS

Edited by:
Karim Lekadir, University of Barcelona, Spain

Reviewed by:
Jichao Zhao, The University of Auckland, New Zealand
Marta Nuñez-Garcia, Institut de Rythmologie et Modélisation Cardiaque (IHU-Liryc), France

****Correspondence:***
Chen Chen
chen.chen15@imperial.ac.uk

Specialty section:
This article was submitted to Cardiovascular Imaging, a section of the journal Frontiers in Cardiovascular Medicine

Received: *30 October 2019*
Accepted: *17 February 2020*
Published: *05 March 2020*

Citation:
Chen C, Qin C, Qiu H, Tarroni G, Duan J, Bai W and Rueckert D (2020) Deep Learning for Cardiac Image Segmentation: A Review. Front. Cardiovasc. Med. 7:25. doi: 10.3389/fcvm.2020.00025

1. INTRODUCTION

Cardiovascular diseasess (CVDs) are the leading cause of death globally according to World Health Organization (WHO). About 17.9 million people died from CVDs in 2016, from CVD, mainly from heart disease and stroke[1]. The number is still increasing annually. In recent decades, major advances have been made in cardiovascular research and practice aiming to improve diagnosis and treatment of cardiac diseases as well as reducing the mortality of CVD. Modern medical imaging techniques, such as magnetic resonance imaging (MRI), computed tomography (CT) and ultrasound are now widely used, which enable non-invasive qualitative and quantitative assessment of cardiac anatomical structures and functions and provide support for diagnosis, disease monitoring, treatment planning, and prognosis.

Of particular interest, cardiac image segmentation is an important first step in numerous applications. It partitions the image into a number of semantically (i.e., anatomically) meaningful regions, based on which quantitative measures can be extracted, such as the myocardial mass, wall thickness, left ventricle (LV) and right ventricle (RV) volume as well as ejection fraction (EF) etc. Typically, the anatomical structures of interest for cardiac image segmentation include the LV, RV, left atrium (LA), right atrium (RA), and coronary arteries. An overview of typical tasks related to cardiac image segmentation is presented in **Figure 1**, where applications for the three most commonly used modalities, i.e., MRI, CT, and ultrasound, are shown.

[1] https://www.who.int/cardiovascular_diseases/about_cvd/en/

Before the rise of deep learning, traditional machine learning techniques, such as model-based methods (e.g., active shape and appearance models) and atlas-based methods had been shown to achieve good performance in cardiac image segmentation (1–4). However, they often require significant feature engineering or prior knowledge to achieve satisfactory accuracy. In contrast, deep learning (DL)-based algorithms are good at *automatically* discovering intricate features from data for object detection and segmentation. These features are directly learned from data using a general-purpose learning procedure and in end-to-end fashion. This makes DL-based algorithms easy to apply to other image analysis applications. Benefiting from advanced computer hardware [e.g., graphical processing units (GPUs) and tensor processing units (TPUs)] as well as increased available data for training, DL-based segmentation algorithms have gradually outperformed previous state-of-the-art traditional methods, gaining more popularity in research. This trend can be observed in **Figure 2A**, which shows how the number of DL-based papers for cardiac image segmentation has increased strongly in the last years. In particular, the number of the publications for MR image segmentation is significantly higher than the numbers of the other two domains, especially in 2017. One reason, which can be observed in **Figure 2B**, is that the publicly available data for MR segmentation has increased remarkably since 2016.

In this paper, we provide an overview of state-of-the-art deep learning techniques for cardiac image segmentation in the three most commonly used modalities (i.e., MRI, CT, ultrasound) in clinical practice and discuss the advantages and remaining limitations of current deep learning-based segmentation methods that hinder widespread clinical deployment. To our knowledge, there have been several review papers that presented overviews about applications of DL-based methods for general medical image analysis (5–7), as well as some surveys dedicated to applications designed for cardiovascular image analysis (8, 9). However, none of them has provided a systematic overview focused on *cardiac segmentation applications*. This review paper aims at providing a comprehensive overview from the debut to the state-of-the-art of deep learning algorithms, focusing on a variety of cardiac image segmentation tasks (e.g., the LV, RV, and vessel segmentation) (section 3). Particularly, we aim to cover most influential DL-related works in this field published until 1st August 2019 and categorized these publications in terms of specific methodology. Besides, in addition to the basics of deep learning introduced in section 2, we also provide a summary of public datasets (see **Table 6**) as well as public code (see **Table 7**), aiming to present a good reading basis for newcomers to the topic and encourage future contributions. More importantly, we provide insightful discussions about the current research situations (section 3.4) as well as challenges and potential directions for future work (section 4).

1.1. Search Criterion

To identify related contributions, search engines like Scopus and PubMed were queried for papers containing ("convolutional" OR "deep learning") and ("cardiac") and ("image segmentation") in title or abstract. Additionally, conference proceedings for MICCAI, ISBI, and EMBC were searched based on the titles of papers. Papers which do not primarily focus on segmentation problems were excluded. The last update to the included papers was on Aug 1, 2019.

2. FUNDAMENTALS OF DEEP LEARNING

Deep learning models are deep artificial neural networks. Each neural network consists of an input layer, an output layer, and multiple hidden layers. In the following section, we will review several deep learning networks and key techniques that have been commonly used in state-of-the-art segmentation algorithms. For a more detailed and thorough illustration of the mathematical background and fundamentals of deep learning we refer the interested reader to Goodfellow (43).

2.1. Neural Networks

In this section, we first introduce basic neural network architectures and then briefly introduce building blocks which are commonly used to boost the ability of the networks to learn features that are useful for image segmentation.

2.1.1. Convolutional Neural Networks (CNNs)

In this part, we will introduce convolutional neural network (CNN), which is the most common type of deep neural networks for image analysis. CNN have been successfully applied to advance the state-of-the-art on many image classification, object detection and segmentation tasks.

Abbreviations: Imaging-related terminology: CT, computed tomography; CTA, computed tomography angiography; LAX, long-axis; MPR, multi-planar reformatted; MR, magnetic resonance; MRI, magnetic resonance imaging; LGE, late gadolinium enhancement; RFCA, radio-frequency catheter ablation; SAX, short-axis; 2CH, 2-chamber; 3CH, 3-chamber; 4CH, 4-chamber.
Cardiac structures and indexes: AF, atrial fibrillation; AS, aortic stenosis; AO, aorta; CVD, cardiovascular diseases; CAC, coronary artery calcium; DCM, dilated cardiomyopathy; ED, end-diastole; ES, end-systole; EF, ejection fraction; HCM, hypertrophic cardiomyopathy; LA, left atrium; LV, left ventricle; LVEDV, left ventricular end-diastolic volume; LVESV, left ventricular end-systolic volume; MCP, mixed-calcified plaque; MI, myocardial infarction; Myo, left ventricular myocardium; NCP, non-calcified plaque; PA, pulmonary artery; PV, pulmonary vein; RA, right atrium; RV, right ventricle; RVEDV, right ventricular end-diastolic volume; RVESV, right ventricular end-systolic volume; RVEF, right ventricular ejection fraction; WHS, whole heart segmentation.
Machine learning terminology: AE, autoencoder; ASM, active shape model; BN, batch normalization; CONV, convolution; CNN, convolutional neural network; CRF, conditional random field; DBN, deep belief network; DL, deep learning; DNN, deep neural network; EM, expectation maximization; FCN, fully convolutional neural network; GAN, generative adversarial network; GRU, gated recurrent units; MSE, mean squared error; MSL, marginal space learning; MRF, markov random field; LSTM, Long-short term memory; ReLU, rectified linear unit; RNN, recurrent neural network; ROI, region-of-interest; SMC, sequential monte carlo; SRF, structured random forest; SVM, support vector machine.
Cardiac image segmentation datasets: ACDC, Automated Cardiac Diagnosis Challenge; CETUS, Challenge on Endocardial Three-dimensional Ultrasound Segmentation; MM-WHS, Multi-Modality Whole Heart Segmentation; LASC, Left Atrium Segmentation Challenge; LVSC, Left Ventricle Segmentation Challenge; RVSC, Right Ventricle Segmentation Challenge.
Others: EMBC, The International Engineering in Medicine and Biology Conference; GDPR, The General Data Protection Regulation; GPU, graphic processing unit; FDA, United States Food and Drug Administration; ISBI, The IEEE International Symposium on Biomedical Imaging; MICCAI, International Conference on Medical Image Computing and Computer-assisted Intervention; TPU, tensor processing unit; WHO, World Health Organization.

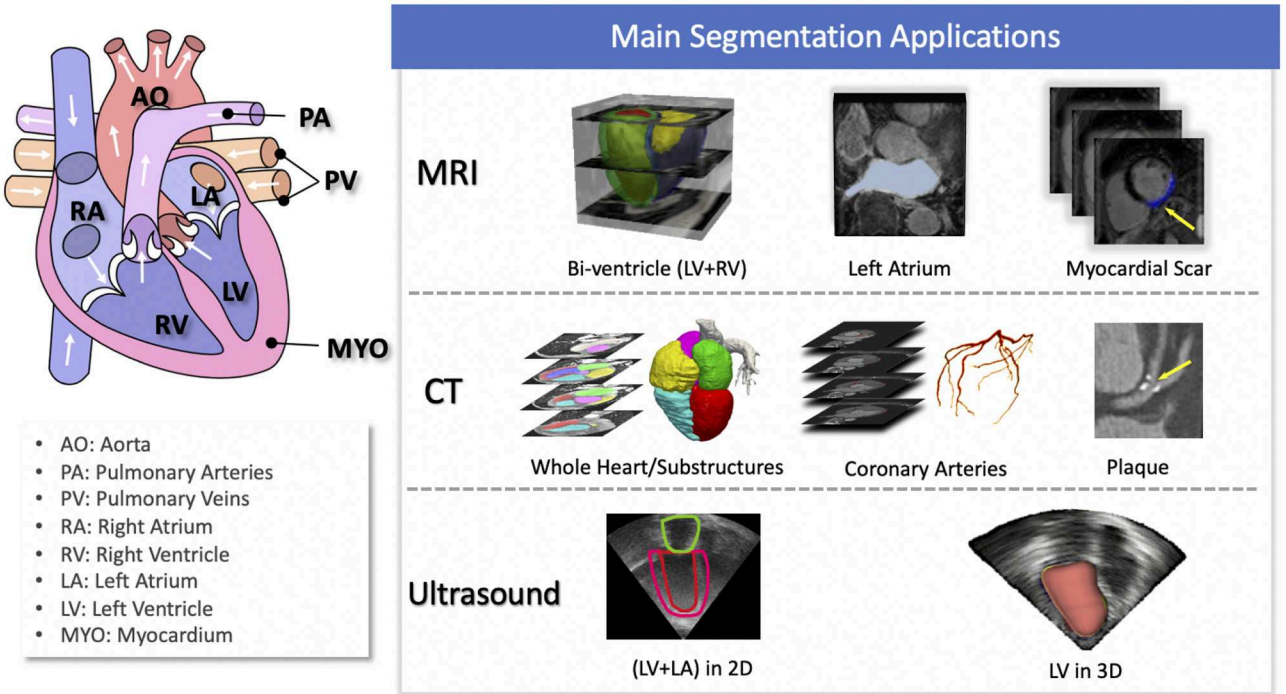

FIGURE 1 | Overview of cardiac image segmentation tasks for different imaging modalities. For better understanding, we provide the anatomy of the heart on the left (image source: Wikimedia Commons, license: CC BY-SA 3.0). Of note, for simplicity, we list the tasks for which deep learning techniques have been applied, which will be discussed in section 3.

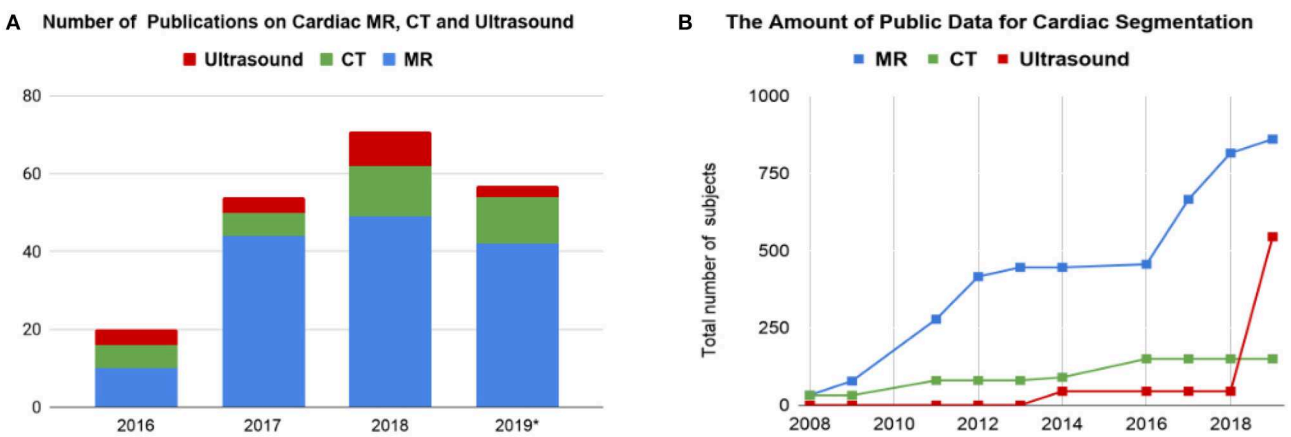

FIGURE 2 | (A) Overview of numbers of papers published from 1st January 2016 to 1st August 2019 regarding deep learning-based methods for cardiac image segmentation reviewed in this work. **(B)** The increase of public data for cardiac image segmentation in the past 10 years. A list of publicly available datasets with detailed information is provided in **Table 6**. CT, computed tomography; MR, magnetic resonance.

As shown in **Figure 3A**, a standard CNN consists of an input layer, an output layer and a stack of functional layers in between that transform an input into an output in a specific form (e.g., vectors). These functional layers often contains convolutional layers, pooling layers and/or fully-connected layers. In general, a convolutional layer CONV_l contains k_l convolution kernels/filters, which is followed by a normalization layer [e.g., batch normalization (44)] and a non-linear activation function [e.g., rectified linear unit (ReLU)] to extract k_l feature maps from the input. These feature maps are then downsampled by pooling layers, typically by a factor of 2, which remove redundant features to improve the statistical efficiency and model generalization. After that, fully connected layers are applied to reduce the dimension of features from its previous layer and

find the most task-relevant features for inference. The output of the network is a fix-sized vector where each element can be a probabilistic score for each category (for image classification), a real value for a regression task (e.g., the left ventricular volume estimation) or a set of values (e.g., the coordinates of a bounding box for object detection and localization).

A key component of CNN is the convolutional layer. Each convolutional layer has k_l convolution kernels to extract k_l feature maps and the size of each kernel n is chosen to be small in general, e.g., $n = 3$ for a 2D 3×3 kernel, to reduce the number of parameters[2]. While the kernels are small, one can increase the receptive field (the area of the input image that potentially impacts the activation of a particular convolutional kernel/neuron) by increasing the number of convolutional layers. For example, a convolutional layer with large 7×7 kernels can be replaced by three layers with small 3×3 kernels (45). The number of weights is reduced by a factor of $7^2/(3 \times (3^2)) \approx 2$ while the receptive field remains the same (7×7). An online resource[3] is referred here, which illustrates and visualizes the change of receptive field by varying the number of hidden layers and the size of kernels. In general, increasing the depth of convolution neural networks (the number of hidden layers) to enlarge the receptive field can lead to improved model performance, e.g., classification accuracy (45).

CNNs for image classification can also be employed for image segmentation applications without major adaptations to the network architecture (46), as shown in **Figure 3B**. However, this requires to divide each image into patches and then train a CNN to predict the class label of the center pixel for every patch. One major disadvantage of this patch-based approach is that, at inference time, the network has to be deployed for every patch individually despite the fact that there is a lot of redundancy due to multiple overlapping patches in the image. As a result of this inefficiency, the main application of CNNs with fully connected layers for cardiac segmentation is object localization, which aims to estimate the bounding box of the object of interest in an image. This bounding box is then used to crop the image, forming an image pre-processing step to reduce the computational cost for segmentation (47). For efficient, end-to-end pixel-wise segmentation, a variant of CNNs called fully convolutional neural network (FCN) is more commonly used, which will be discussed in the next section.

2.1.2. Fully Convolutional Neural Networks (FCNs)

The idea of FCN was first introduced by Long et al. (48) for image segmentation. FCNs are a special type of CNNs that do not have any fully connected layers. In general, as shown in **Figure 4A**, FCNs are designed to have an encoder-decoder structure such that they can take input of arbitrary size and produce the output with the same size. Given an input image, the encoder first transforms the input into high-level feature representation whereas the decoder interprets the feature maps and recovers spatial details back to the image space for pixel-wise prediction through a series of upsampling and convolution operations. Here, upsampling can be achieved by applying transposed convolutions, e.g., 3×3 transposed convolutional kernels with a stride of 2 to up-scale feature maps by a factor of 2. These transposed convolutions can also be replaced by unpooling layers and upsampling layers. Compared to a patch-based CNN for segmentation, FCN is trained and applied to the entire images, removing the need for patch selection (50).

FCN with the simple encoder-decoder structure in **Figure 4A** may be limited to capture detailed context information in an image for precise segmentation as some features may be eliminated by the pooling layers in the encoder. Several variants of FCNs have been proposed to propagate features from the encoder to the decoder, in order to boost the segmentation accuracy. The most well-known and most popular variant of FCNs for biomedical image segmentation is the U-net (49). On the basis of the vanilla FCN (48), the U-net employs skip connections between the encoder and decoder to recover spatial context loss in the down-sampling path, yielding more precise segmentation (see **Figure 4B**). Several state-of-the-art cardiac image segmentation methods have adopted the U-net or its 3D variants, the 3D U-net (51) and the 3D V-net (52), as their backbone networks, achieving promising segmentation accuracy for a number of cardiac segmentation tasks (26, 53, 54).

2.1.3. Recurrent Neural Networks (RNNs)

Recurrent neural networks (RNNs) are another type of artificial neural networks which are used for sequential data, such as cine MRI and ultrasound image sequences. An RNN can "remember" the past and use the knowledge learned from the past to make its present decision (see **Figures 5A,B**). For example, given a sequence of images, an RNN takes the first image as input, captures the information to make a prediction and then memorize this information which is then utilized to make a prediction for the next image. The two most widely used architectures in the family of RNNs are LSTM (56) and gated recurrent unit (GRU) (57), which are capable of modeling long-term memory. A use case for cardiac segmentation is to combine an RNN with a 2D FCN so that the combined network is capable of capturing information from adjacent slices to improve the inter-slice coherence of segmentation results (55).

2.1.4. Autoencoders (AE)

Autoencoders (AEs) are a type of neural networks that are designed to learn compact latent representations from data without supervision. A typical architecture of an autoencoder consists of two networks: an encoder network and a decoder network for the reconstruction of the input (see **Figure 6**). Since the learned representations contain generally useful information in the original data, many researchers have

[2]In a convolution layer l with k_l 2D $n \times n$ convolution kernels, each convolution kernel $\text{CONV}_l^{(i)}, i \in (1, k_l)$ has a weight matrix $\mathbf{w}_l^{(i)}$ and a bias term $b_l^{(i)}$ as parameters and can be formulated as: $\mathbf{y} = \mathbf{w}_l^{(i)} \circ \mathbf{x}_{in} + b_l^{(i)}$, where $\mathbf{w}_l^{(i)} \in \mathbb{R}^{n \times n \times l_{in}}$, $b_l^{(i)} \in \mathbb{R}$, $\mathbf{x}_{in} \in \mathbb{R}^{H \times W \times l_{in}}$, $\mathbf{y} \in \mathbb{R}^{H' \times W' \times k_l}$, l_{in} denotes the number of channels in the input $\mathbf{x}_{in}$ and $\circ$ denotes the convolution operation. Thus, the number of parameters in a convolutional layer is $k_l \times (n^2 \times l_{in} + 1)$. For a convolutional layer with 16 3×3 filters where the input is a $28 \times 28 \times 1$ 2D gray image, the number of parameters in this layer is $16 \times (3^2 \times 1 + 1) = 160$. For more technical details about convolutional neural networks, an online tutorial is referred here: http://cs231n.github.io/convolutional-networks.

[3]https://fomoro.com/research/article/receptive-field-calculator

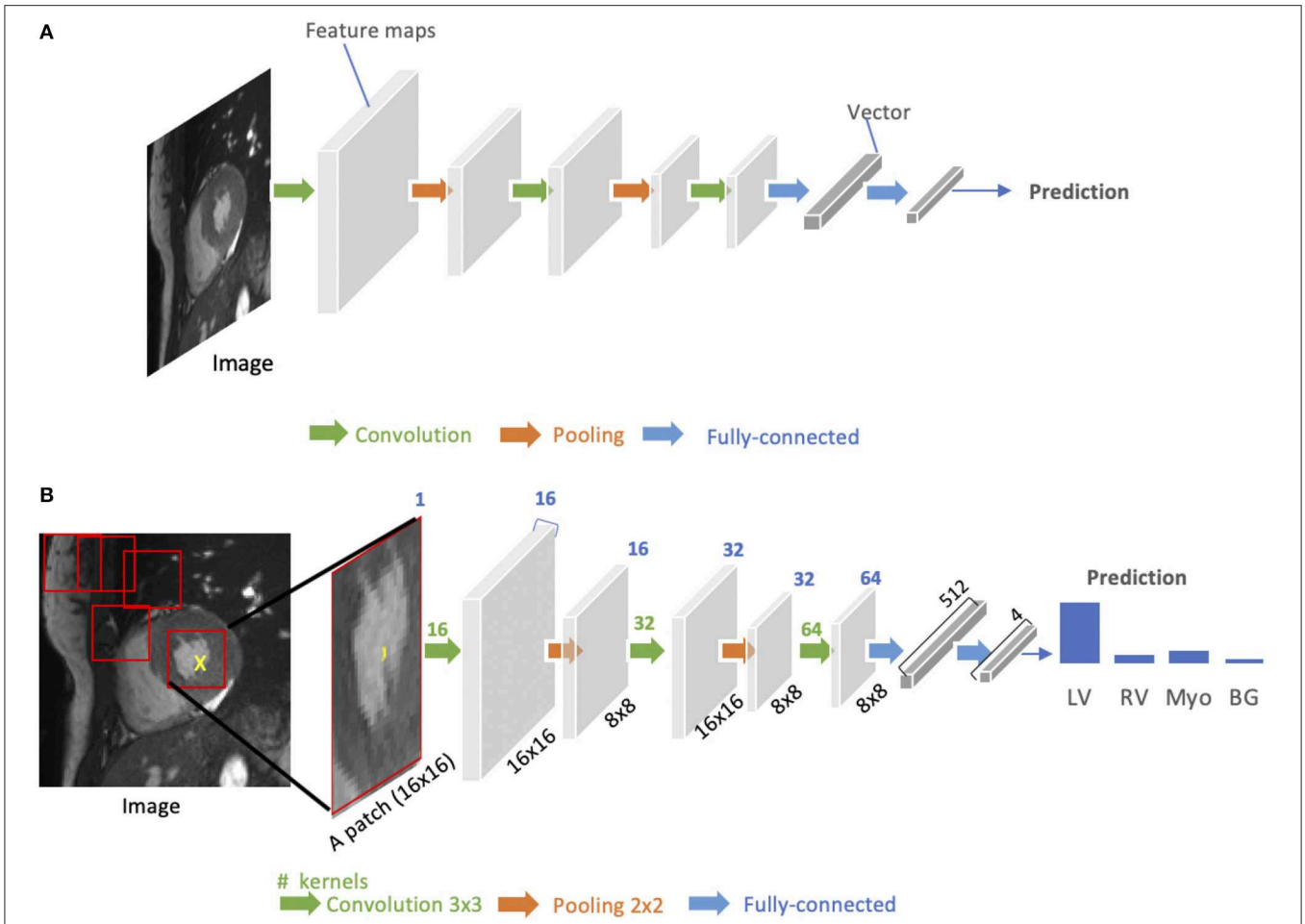

FIGURE 3 | (A) Generic architecture of convolutional neural networks (CNN). A CNN takes a cardiac MR image as input, learning hierarchical features through a stack of convolutions and pooling operations. These spatial feature maps are then flattened and reduced into a vector through fully connected layers. This vector can be in many forms, depending on the specific task. It can be probabilities for a set of classes (image classification) or coordinates of a bounding box (object localization) or a predicted label for the center pixel of the input (patch-based segmentation) or a real value for regression tasks (e.g., left ventricular volume estimation). **(B)** Patch-based segmentation method based on a CNN classifier. The CNN takes a patch as input and outputs the probabilities for four classes where the class with the highest score is the prediction for the center pixel (see the yellow cross) in this patch. By repeatedly forwarding patches located at different locations into the CNN for classification, one can finally get a pixel-wise segmentation map for the whole image. LV, left ventricle cavity; RV, right ventricle cavity; BG, Background; Myo, left ventricular myocardium. The blue number at the top indicates the number of channels of the feature maps. Here, each convolution kernel is a 3 × 3 kernel (stride = 1, padding = 1), which will produces an output feature map with the same height and width as the input.

employed autoencoders to extract general semantic features or shape information from input images or labels and then use those features to guide the cardiac image segmentation (58, 62, 63).

2.1.5. Generative Adversarial Networks (GAN)

The concept of Generative adversarial network (GAN) was proposed by Goodfellow et al. (64) for image synthesis from noise. GANs are a type of generative models that learn to model the data distribution of real data and thus are able to create new image examples. As shown in **Figure 7A**, a GAN consists of two networks: a generator network and a discriminator network. During training, the two networks are trained to compete against each other: the generator produces fake images aimed at fooling the discriminator, whereas the discriminator tries to identify real images from fake ones. This type of training is referred to as "adversarial training," since the two models are both set to win the competition. This training scheme can also be used for training a segmentation network. As shown in **Figure 7B**, the generator is replaced by a segmentation network and the discriminator is required to distinguish the generated segmentation maps from the ground truth ones (the target segmentation maps). In this way, the segmentation network is encouraged to produce more anatomically plausible segmentation maps (65, 66).

2.1.6. Advanced Building Blocks for Improved Segmentation

Medical image segmentation, as an important step for quantitative analysis and clinical research, requires high pixel-wise accuracy. Over the past years, many researchers have developed advanced building blocks to learn robust,

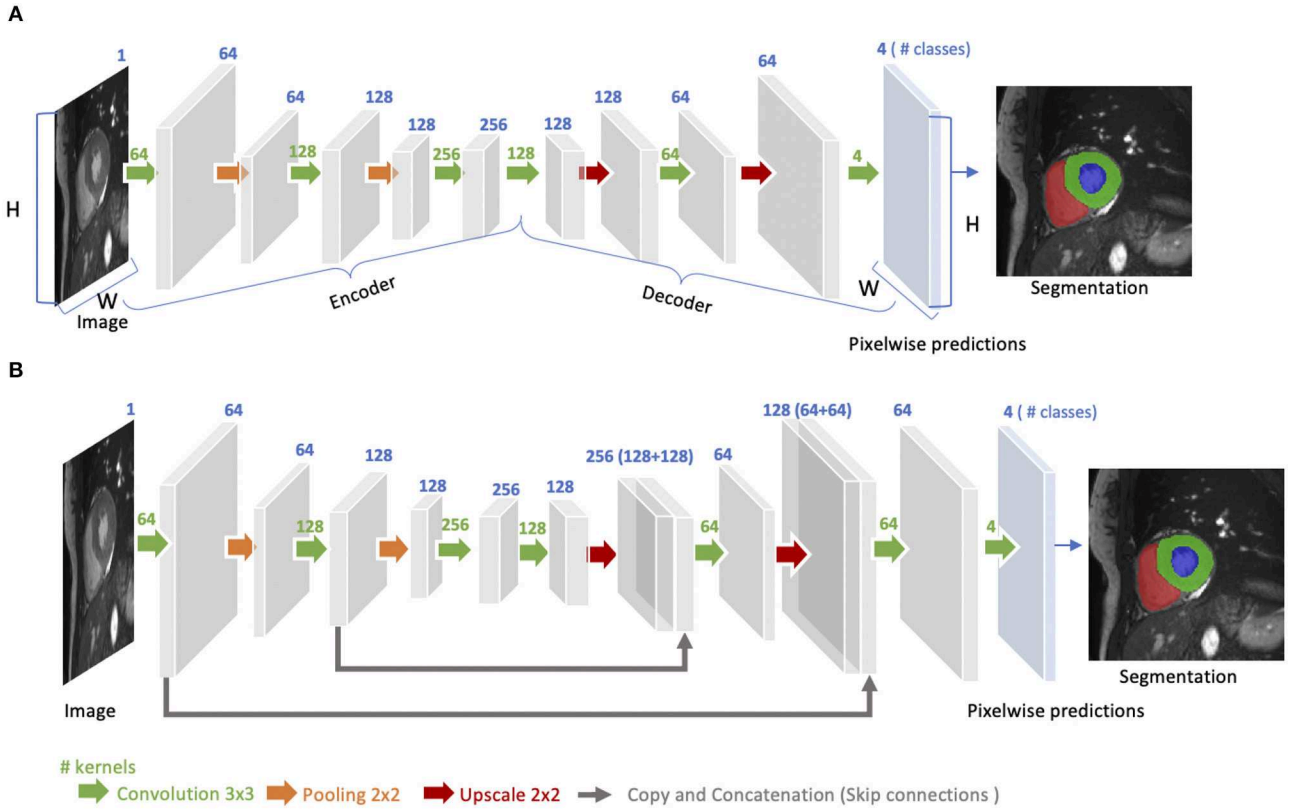

FIGURE 4 | (A) Architecture of a fully convolutional neural network (FCN). The FCN first takes the whole image as input, learns image features though the encoder, gradually recovers the spatial dimension by a series of upscaling layers (e.g., transposed convolution layers, unpooling layers) in the decoder and then produce 4-class pixel-wise probabilistic maps to predict regions of the left ventricle cavity (blue region), the left ventricular myocardium (green region) and the right ventricle cavity (red region) and background. The final segmentation map is obtained by assigning each pixel with the class of the highest probability. One use case of this FCN-based cardiac segmentation can be found in Tran (24). **(B)** Architecture of a U-net. On the basis of FCN, U-net adds "skip connections" (gray arrows) to aggregate feature maps from coarse to fine through concatenation and convolution operations. For simplicity, we reduce the number of downsampling and upsampling blocks in the diagram. For detailed information, we recommend readers to the original paper (49).

representative features for precise segmentation. These techniques have been widely applied to state-of-the-art neural networks (e.g., U-net) to improve cardiac image segmentation performance. Therefore, we identified several important techniques reported in the literature to this end and present them with corresponding references for further reading. These techniques are:

1. Advanced convolutional modules for multi-scale feature aggregation:
   - Inception modules (44, 67, 68), which concatenate multiple convolutional filter banks with different kernel sizes to extract multi-scale features in parallel (see **Figure 8A**);
   - Dilated convolutional kernels (72), which are modified convolution kernels with the same kernel size but different kernel strides to process input feature maps at larger scales;
   - Deep supervision (73), which utilizes the outputs from multiple intermediate hidden layers for multi-scale prediction;
   - Atrous spatial pyramid pooling (74), which applies spatial pyramid pooling (75) with various kernel strides to input feature maps for multi-scale feature fusion;
2. Adaptive convolutional kernels designed to focus on important features:
   - Attention units (69, 70, 76), which learn to adaptively recalibrate features spatially (see **Figure 8B**);
   - Squeeze-and-excitation blocks (77), which are used to recalibrate features with learnable weights across channels;
3. Interlayer connections designed to reuse features from previous layers:
   - Residual connections (71), which add outputs from a previous layer to the feature maps learned from the current layer (see **Figure 8C**);
   - Dense connections (78), which concatenate outputs from all preceding layers to the feature maps learned from the current layer.

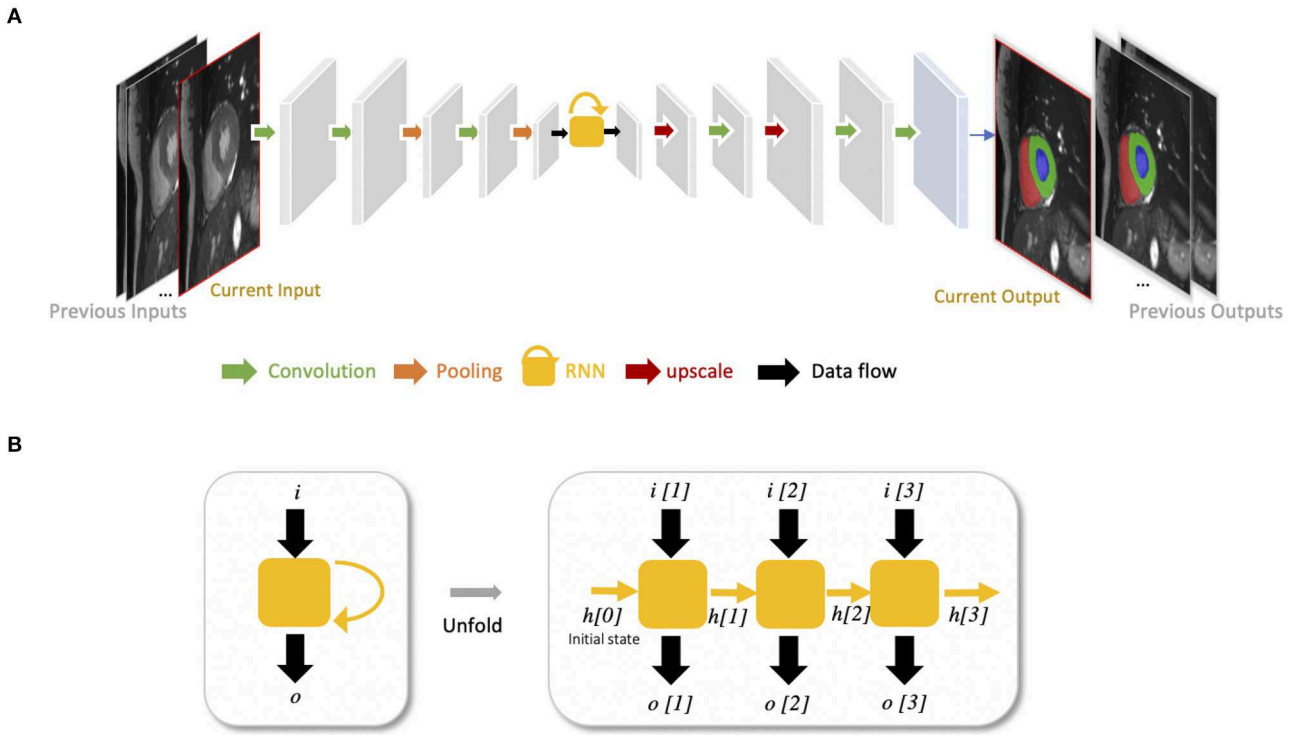

FIGURE 5 | (A) Example of FCN with an RNN for cardiac image segmentation. The yellow block with a curved arrow represents a RNN module, which utilizes the knowledge learned from the past to make the current decision. In this example, the network is used to segment cardiac ventricles from a stack of 2D cardiac MR slices, which allows propagation of contextual information from adjacent slices for better inter-slice coherence (55). This type of RNN is also suitable for sequential data, such as cine MR images and ultrasound movies to learn temporal coherence. **(B)** Unfolded schema of the RNN module for visualizing the inner process when the input is a sequence of three images. Each time, this RNN module will receive an input $i[t]$ at time step t, and produce an output $o[t]$, considering not only the input information but also the hidden state ("memory") $h[t-1]$ from the previous time step $t-1$.

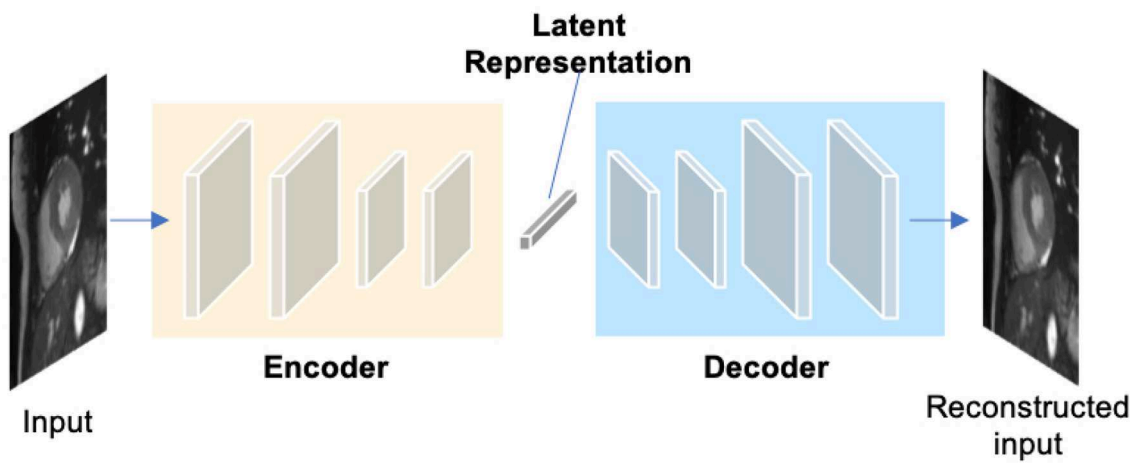

FIGURE 6 | A generic architecture of an autoencoder. An autoencoder employs an encoder-decoder structure, where the encoder maps the input data to a low-dimensional latent representation and the decoder interprets the code and reconstructs the input. The learned latent representation has been found effective for cardiac image segmentation (58, 59), cardiac shape modeling (60) and cardiac segmentation correction (61).

2.2. Training Neural Networks

Before being able to perform inference, neural networks must be trained. Standard training process requires a dataset that contains paired images and labels $\{x, y\}$ for training and testing, an optimizer (e.g., stochastic gradient descent, Adam) and a loss function to update the model parameters. This function accounts for the error of the network prediction in each iteration during training, providing signals for the optimizer to update the

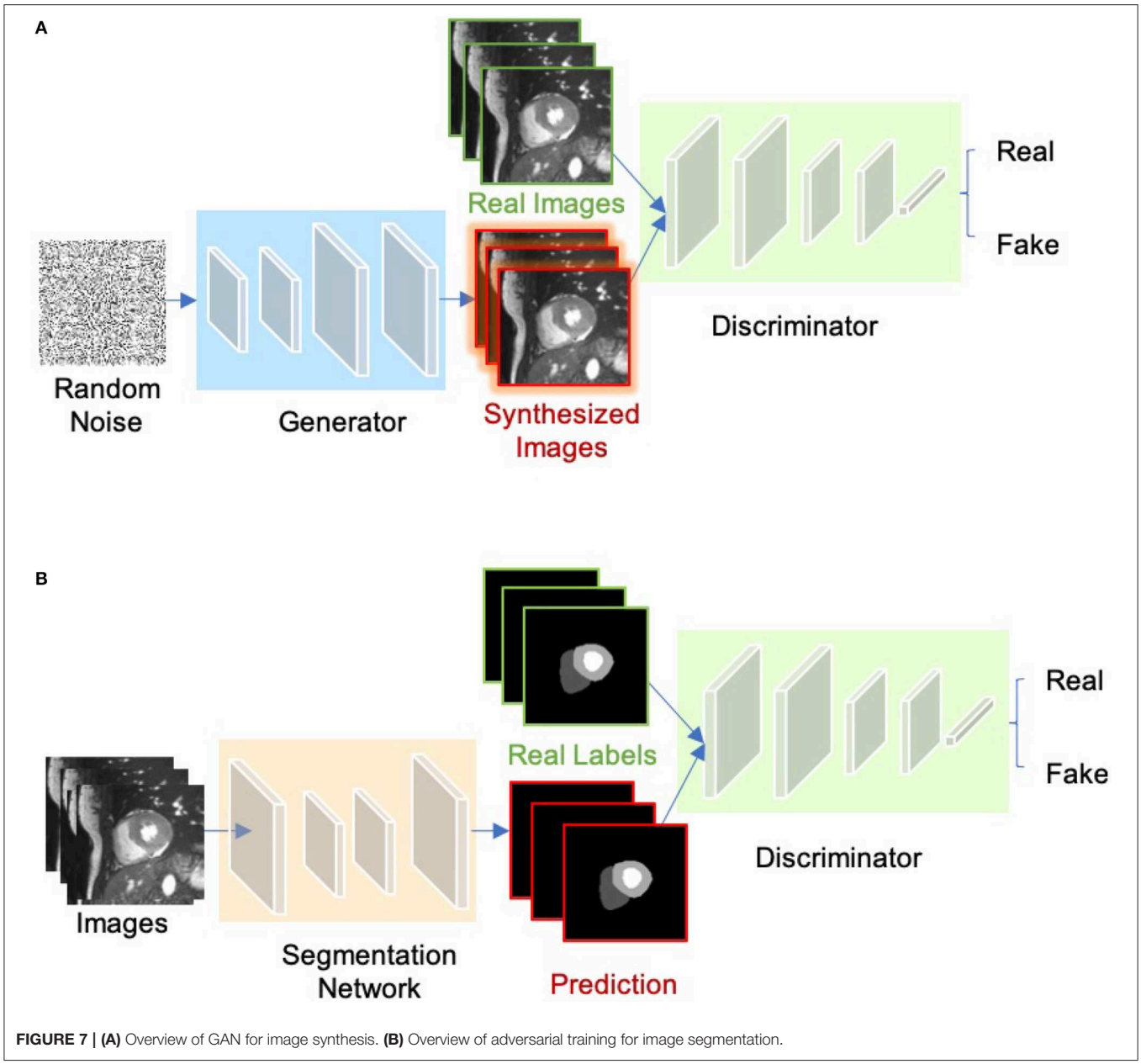

FIGURE 7 | (A) Overview of GAN for image synthesis. **(B)** Overview of adversarial training for image segmentation.

network parameters through backpropagation (43, 79). The goal of training is to find proper values of the network parameters to minimize the loss function.

2.2.1. Common Loss Functions

For regression tasks (e.g., heart localization, calcium scoring, landmark detection, image reconstruction), the simplest loss function is the mean squared error (MSE):

$$\mathcal{L}_{\text{MSE}} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{1}$$

where y_i is the vector of target values and $\hat{y}_i$ is the vector of the predicted values; n is the number of data samples at each iteration.

Cross-entropy is the most common loss for both image classification and segmentation tasks. In particular, the cross-entropy loss for segmentation summarizes pixel-wise probability errors between a predicted probabilistic output $\boldsymbol{p}_i^c$ and its corresponding target segmentation map $\boldsymbol{y}_i^c$ for each class c[4]:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{n}\sum_{i=1}^{n}\sum_{c=1}^{C}\boldsymbol{y}_i^c \log(\boldsymbol{p}_i^c), \tag{2}$$

[4]At inference time, the predicted segmentation map for each image is obtained by assigning each pixel with the class of the highest probability: $\hat{y}_i = \text{argmax}_c\, \boldsymbol{p}_i^c$.

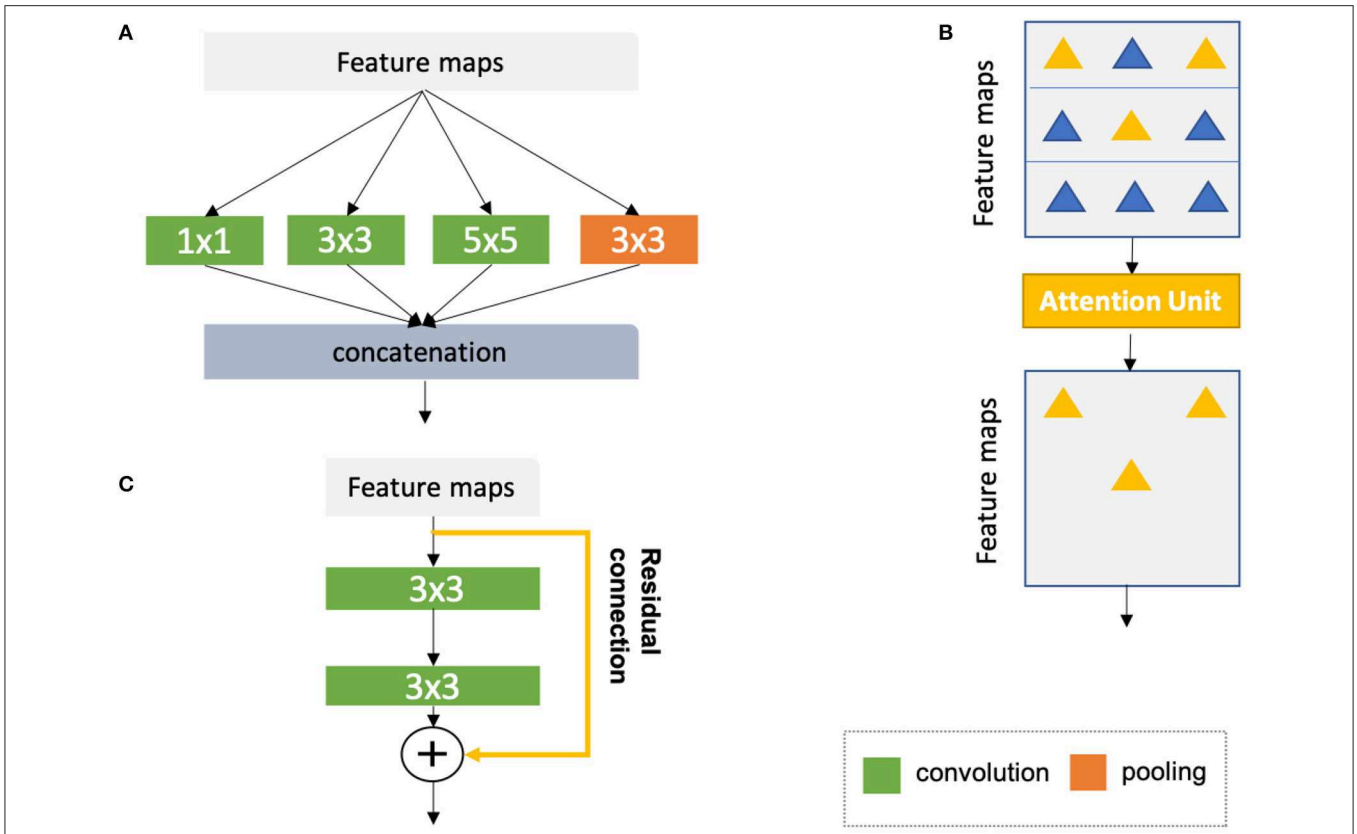

FIGURE 8 | (A) Naive version of the inception module (44). In this module, convolutional kernels with varying sizes are applied to the same input for multi-scale feature fusion. On the basis of the naive structure, a family of advanced inception modules with more complex structures have been developed (67, 68). **(B)** Schematic diagram of the attention module (69, 70). The attention module teaches the network to pay attention to important features (e.g., features relevant to anatomy) and ignore redundant features. **(C)** Schematic diagram of a residual unit (71). The yellow arrow represents a residual connection which is applied to reusing the features from a previous layer. The numbers in the green and orange blocks denote the sizes of corresponding convolutional or pooling kernels. Here, for simplicity, all diagrams have been reproduced based on the illustration in the original papers.

where C is the number of all classes. Another loss function which is specifically designed for object segmentation is called soft-Dice loss function (52), which penalizes the mismatch between a predicted segmentation map and its target map at pixel-level:

$$\mathcal{L}_{Dice} = 1 - \frac{2\sum_{i=1}^{n}\sum_{c=1}^{C} y_i^c p_i^c}{\sum_{i=1}^{n}\sum_{c=1}^{C}(y_i^c + p_i^c)}. \quad (3)$$

In addition, there are several variants of the cross-entropy or soft-Dice loss, such as the weighted cross-entropy loss (25, 80) and weighted soft-Dice loss (29, 81) that are used to address potential class imbalance problem in medical image segmentation tasks where the loss term is weighted to account for rare classes or small objects.

2.2.2. Reducing Over-Fitting

The biggest challenge of training deep networks for medical image analysis is over-fitting, due to the fact that there is often a limited number of training images in comparison with the number of learnable parameters in a deep network. A number of techniques have been developed to alleviate this problem. Some of the techniques are the following ones:

- **Weight regularization**: Weight regularization is a type of regularization techniques that add weight penalties to the loss function. Weight regularization encourages small or zero weights for less relevant or irrelevant inputs. Common methods to constrain the weights include L1 and L2 regularization, which penalize the sum of the absolute weights and the sum of the squared weights, respectively;
- **Dropout** (82): Dropout is a regularization method that randomly drops some units from the neural network during training, encouraging the network to learn a sparse representation;
- **Ensemble learning**: Ensemble learning is a type of machine learning algorithms that combine multiple trained models to obtain better predictive performance than individual models, which has been shown effective for medical image segmentation (83, 84);
- **Data augmentation**: Data augmentation is a training strategy that artificially generates more training samples to increase the diversity of the training data. This can be done via applying affine transformations (e.g., rotation, scaling), flipping or cropping to original labeled samples;

- **Transfer learning**: Transfer learning aims to transfer knowledge from one task to another related but different target task. This is often achieved by reusing the weights of a pre-trained model, to initialize the weights in a new model for the target task. Transfer learning can help to decrease the training time and achieve lower generalization error (85).

2.3. Evaluation Metrics

To quantitatively evaluate the performance of automated segmentation algorithms, three types of metrics are commonly used: (a) volume-based metrics (e.g., Dice metric, Jaccard similarity index); (b) surface distance-based metrics (e.g., mean contour distance, Hausdorff distance); (c) clinical performance metrics (e.g., ventricular volume and mass). For a detailed illustration of common used clinical indices in cardiac image analysis, we recommend the review paper by Peng et al. (2). In our paper, we mainly report the accuracy of methods in terms of the Dice metric for ease of comparison. The Dice score measures the ratio of overlap between two results (e.g., automatic segmentation vs. manual segmentation), ranging from 0 (mismatch) to 1 (perfect match). It is also important to note that the segmentation accuracy of different methods are not directly comparable in general, unless these methods are evaluated on the same dataset. This is because, even for the same segmentation task, different datasets can have different imaging modalities, different patient populations and different methods of image acquisition, which will affect the task complexities and result in different segmentation performances.

3. DEEP LEARNING FOR CARDIAC IMAGE SEGMENTATION

In this section, we provide a summary of deep learning-based applications for the three main imaging modalities: MRI, CT, and ultrasound regarding specific applications for targeted structures. In general, these deep learning-based methods provide an efficient and effective way to segmenting particular organs or tissues (e.g., the LV, coronary vessels, scars) in different modalities, facilitating follow-up quantitative analysis of cardiovascular structure and function. Among these works, a large portion of these methods are designed for ventricle segmentation, especially in MR and ultrasound domains. The objective of ventricle segmentation is to delineate the endocardium and epicardium of the LV and/or RV. These segmentation maps are important for deriving clinical indices, such as left ventricular end-diastolic volume (LVEDV), left ventricular end-systolic volume (LVESV), right ventricular end-diastolic volume (RVEDV), right ventricular end-systolic volume (RVESV), and EF. In addition, these segmentation maps are essential for 3D shape analysis (60, 86), 3D + time motion analysis (87), and survival prediction (88).

3.1. Cardiac MR Image Segmentation

Cardiac MRI is a non-invasive imaging technique that can visualize the structures within and around the heart. Compared to CT, it does not require ionizing radiation. Instead, it relies on the magnetic field in conjunction with radio-frequency waves to excite hydrogen nuclei in the heart, and then generates an image by measuring their response. By utilizing different imaging sequences, cardiac MRI allows accurate quantification of both cardiac anatomy and function (e.g., cine imaging) and pathological tissues, such as scars (late gadolinium enhancement (LGE) imaging). Accordingly, cardiac MRI is currently regarded as the gold standard for quantitative cardiac analysis (89).

A group of representative deep learning based cardiac MR segmentation methods are shown in **Table 1**. From the table, one can see that a majority of works have focused on segmenting cardiac chambers (e.g., LV, RV, LA). In contrast, there are relatively fewer works on segmenting abnormal cardiac tissue regions, such as myocardial scars and atrial fibrosis from contrast-enhanced images. This is likely due to the limited relevant public datasets as well as the difficulty of the task. In addition, to the best of our knowledge, there are very few works that apply deep learning techniques to atrial wall segmentation, as also suggested by a recent survey paper (161). In the following sections, we will describe and discuss these methods regarding different applications in detail.

3.1.1. Ventricle Segmentation

3.1.1.1. Vanilla FCN-based segmentation

Tran (24) was among the first ones to apply a FCN (50) to segment the left ventricle, myocardium and right ventricle directly on short-axis cardiac magnetic resonance (MR) images. Their end-to-end approach based on FCN achieved competitive segmentation performance, significantly outperforming traditional methods in terms of both speed and accuracy. In the following years, a number of works based on FCNs have been proposed, aiming at achieving further improvements in segmentation performance. In this regard, one stream of work focuses on optimizing the network structure to enhance the feature learning capacity for segmentation (29, 80, 91, 162–165). For example, Khened et al. (29) developed a dense U-net with inception modules to combine multi-scale features for robust segmentation across images with large anatomical variability. Jang et al. (80), Yang et al. (81), Sander et al. (166), and Chen et al. (167) investigated different loss functions, such as weighted cross-entropy, weighted Dice loss, deep supervision loss and focal loss to improve the segmentation performance. Among these FCN-based methods, the majority of approaches use 2D networks rather than 3D networks for segmentation. This is mainly due to the typical low through-plane resolution and motion artifacts of most cardiac MR scans, which limits the applicability of 3D networks (25).

3.1.1.2. Introducing spatial or temporal context

One drawback of using 2D networks for cardiac segmentation is that these networks work slice by slice, and thus they do not leverage any inter-slice dependencies. As a result, 2D networks can fail to locate and segment the heart on challenging slices, such as apical and basal slices where the contours of the ventricles are not well-defined. To address this problem, a number of works have attempted to introduce additional contextual information to guide 2D FCN. This contextual information can include shape priors learned from labels or multi-view images (109, 110, 168).

TABLE 1 | A summary of representative deep learning methods on cardiac MRI segmentation.

Application	Selected works	Description	Type of images	Structure(s)
Ventricle segmentation	**FCN-based**			
	Tran (24)	2D FCN	SAX	Bi-ventricle
	Lieman-Sifry et al. (90)	A lightweight FCN (E-Net)	SAX	Bi-ventricle
	Isensee et al. (26)	2D U-net +3D U-net (ensemble)	SAX	Bi-ventricle
	Jang et al. (80)	2D M-Net with weighted cross entropy loss	SAX	Bi-ventricle
	Baumgartner et al. (25)	2D U-net with cross entropy	SAX	Bi-ventricle
	Bai et al. (31)	2D FCN trained and verified on a large dataset (∼ 5000 subjects);	SAX, 2CH, 4CH	Four chambers
	Tao et al. (53)	2D U-net trained and verified on a multi-vendor, multi-scanner dataset	SAX	LV, Myo
	Khened et al. (29)	2D Dense U-net with inception module	SAX	Bi-ventricle
	Fahmy et al. (91)	2D FCN	SAX	LV, Myo
	Introducing spatial or temporal context			
	Poudel et al. (55)	2D FCN with RNN to model inter-slice coherency	SAX	Bi-ventricle
	Patravali et al. (92)	2D multi-channel FCN to aggregate inter-slice information	SAX	Bi-ventricle
	Wolterink et al. (93)	Dilated U-net to segment ED and ES simultaneously	SAX	Bi-ventricle
	Applying anatomical constraints			
	Oktay et al. (59)	FCN trained with additional anatomical shape-based regularization	SAX; Ultrasound	LV, Myo
	Multi-stage networks			
	Tan et al. (94)	Semi-automated method; CNN (localization) followed by another CNN to derive contour parameters	SAX	LV, Myo
	Zheng et al. (27)	FCN (localization) + FCN (segmentation); Propagate labels from adjacent slices	SAX	Bi-ventricle
	Vigneault et al. (95)	U-net (initial segmentation) + CNN (localization and transformation) + Cascaded U-net (segmentation)	SAX, 2CH, 4CH	Four chambers
	Hybrid segmentation methods			
	Avendi et al. (47, 96)	CNN (localization) + AE (shape initialization) + Deformable model	SAX	LV, Myo/RV
	Yang et al. (97)	CNN combined with Multi-atlas	SAX	LV, Myo
	Ngo et al. (98)	Level-set based segmentation with deep belief networks	SAX	LV, Myo
Atrial segmentation	Mortazi et al. (99)	Multi-view CNN with adaptive fusion strategy	3D scans	LA
	Xiong et al. (100)	Patch-based dual-stream 2D FCN	LGE MRI	LA
	Xia et al. (54)	Two-stage pipeline; 3D U-net (localization) + 3D U-net (segmentation)	LGE MRI	LA
Scar segmentation	Yang et al. (101)	Fully automated; Multi-atlas method for LA segmentation followed by an AE to find the atrial scars	LGE MRI	LA; atrial scars
	Chen et al. (102)	Fully automated; Multi-view two-task recursive attention model	LGE MRI	LA; atrial scars
	Zabihollahy et al. (103)	Semi-automated; 2D CNN for scar tissue classification	LGE MRI	Myocardial scars
	Moccia et al. (104)	Semi-automated; 2D FCN for scar segmentation	LGE MRI	Myocardial scars
	Xu et al. (105)	Fully automated; RNN for joint motion feature learning and scar segmentaion	cine MRI	Myocardial scars
Aorta segmentation	Bai et al. (32)	RNN to learn temporal coherence; Propagate labels from labeled frames to unlabeled adjacent frames for semi-supervised learning;	cine MRI	Aorta
Whole heart segmentation	Yu et al. (30)	3D U-net with deep supervision	3D scans	Blood pool + Myocardium of the heart
	Li et al. (106)	3D FCN with deep supervision	3D scans	Blood pool + Myocardium of the heart
	Wolterink et al. (107)	dilated CNN with deep supervision	3D scans	Blood pool + Myocardium of the heart

By default, LV/RV and LA/RA segmentation refer to the left/right ventricle cavity segmentation and left/right atrium cavity segmentation, respectively. The same applies to ***Tables 2–5****. SAX, short-axis view; 2CH, 2-chamber view; 4CH, 4-chamber view; ED, end-diastolic; ES, end-systolic; Myo, Left ventricular myocardium.*

Others extract spatial information from adjacent slices to assist the segmentation, using recurrent units (RNNs) or multi-slice networks (2.5D networks) (27, 55, 92, 169). These networks can also be applied to leveraging information across different temporal frames in the cardiac cycle to improve spatial and temporal consistency of segmentation results (28, 93, 169–171).

3.1.1.3. Applying anatomical constraints

Another problem that may limit the segmentation performance of both 2D and 3D FCNs is that they are typically trained with pixel-wise loss functions only (e.g., cross-entropy or soft-Dice losses). These pixel-wise loss functions may not be sufficient to learn features that represent the underlying anatomical structures. Several approaches therefore focus on designing and applying anatomical constraints to train the network to improve its prediction accuracy and robustness. These constraints are represented as regularization terms which take into account the topology (172), contour and region information (173), or shape information (59, 63), encouraging the network to generate more anatomically plausible segmentations. In addition to regularizing networks at training time (61), proposed a variational AE to correct inaccurate segmentations, at the post-processing stage.

3.1.1.4. Multi-task learning

Multi-task learning has also been explored to regularize FCN-based cardiac ventricle segmentation during training by performing auxiliary tasks that are relevant to the main segmentation task, such as motion estimation (174), estimation of cardiac function (175), ventricle size classification (176), and image reconstruction (177–179). Training a network for multiple tasks simultaneously encourages the network to extract features which are useful across these tasks, resulting in improved learning efficiency and prediction accuracy.

3.1.1.5. Multi-stage networks

Recently, there is a growing interest in applying neural networks in a multi-stage pipeline which breaks down the segmentation problem into subtasks (27, 94, 95, 108, 180). For example, Zheng et al. (27) and Li et al. (108) proposed a region-of-interest (ROI) localization network followed by a segmentation network. Likewise, Vigneault et al. (95) proposed a network called Omega-Net which consists of a U-net for cardiac chamber localization, a learnable transformation module to normalize image orientation and a series of U-nets for fine-grained segmentation. By explicitly localizing the ROI and by rotating the input image into a canonical orientation, the proposed method better generalizes to images with varying sizes and orientations.

3.1.1.6. Hybrid segmentation methods

Another stream of work aims at combining neural networks with classical segmentation approaches, e.g., level-sets (98, 181), deformable models (47, 96, 182), atlas-based methods (97, 111), and graph-cut based methods (183). Here, neural networks are applied in the feature extraction and model initialization stages, reducing the dependency on manual interactions and improving the segmentation accuracy of the conventional segmentation methods deployed afterwards. For example, Avendi et al. (47) proposed one of the first DL-based methods for LV segmentation in cardiac short-axis MR images. The authors first applied a CNN to automatically detect the LV and then used an AE to estimate the shape of the LV. The estimated shape was then used to initialize follow-up deformable models for shape refinement. As a result, the proposed integrated deformable model converges faster than conventional deformable models and the segmentation achieves higher accuracy. In their later work, the authors extended this approach to segment RV (96). While these hybrid methods demonstrated better segmentation accuracy than previous non-deep learning methods, most of them still require an iterative optimization for shape refinement. Furthermore, these methods are often designed for one particular anatomical structure. As noted in the recent benchmark study (17), most state-of-the-art segmentation algorithms for bi-ventricle segmentation are based on end-to-end FCNs, which allows the simultaneous segmentation of the LV and RV.

To better illustrate these developments for cardiac ventricle segmentation from cardiac MR images, we collate a list of bi-ventricle segmentation methods that have been trained and tested on the Automated Cardiac Diagnosis Challenge (ACDC) dataset, reported in **Table 2**. For ease of comparison, we only consider those methods which have been evaluated on the same online test set (50 subjects). As the ACDC challenge organizers keep the online evaluation platform open to the public, our comparison not only includes the methods from the original challenge participants [summarized in the benchmark study paper from Bernard et al. (17)] but also three segmentation algorithms that have been proposed after the challenge [i.e., (61, 108, 109)]. From this comparison, one can see that top algorithms are the ensemble method proposed by Isensee et al. (26) and the two-stage method proposed by Li et al. (108), both of which are based on FCNs. In particular, compared to the traditional level-set method (112), both methods achieved considerably higher accuracy even for the more challenging segmentation of the left ventricular myocardium (Myo), indicating the power of deep learning based approaches.

3.1.2. Atrial Segmentation

Atrial fibrillation (AF) is one of the most common cardiac electrical disorders, affecting around 1 million people in the UK[5]. Accordingly, atrial segmentation is of prime importance in the clinic, improving the assessment of the atrial anatomy in both pre-operative AF ablation planning and post-operative follow-up evaluations. In addition, the segmentation of atrium can be used as a basis for scar segmentation and atrial fibrosis quantification from LGE images. Traditional methods, such as region growing (184) and methods that employ strong priors [i.e., atlas-based label fusion (185) and non-rigid registration (186)] have been applied in the past for automated left atrium segmentation. However, the accuracy of these methods highly relies on good initialization and *ad-hoc* pre-processing methods, which limits the widespread adoption in the clinic.

Recently, Vigneault et al. (95) and Bai et al. (31) applied 2D FCNs to directly segment the LA and RA from standard 2D long-axis images, i.e., 2-chamber (2CH), 4-chamber (4CH)

[5]https://www.nhs.uk/conditions/atrial-fibrillation/

TABLE 2 | Segmentation accuracy of state-of-the-art segmentation methods verified on the cardiac bi-ventricular segmentation challenge (ACDC) dataset (17).

Methods	Description	LV	Myo	RV
Isensee et al. (26)	2D U-net + 3D U-net (ensemble)	**0.950**	**0.911**	0.923
Li et al. (108)	Two 2D FCNs for ROI detection and segmentation, respectively;	0.944	**0.911**	**0.926**
Zotti et al. (109)	2D GridNet-MD with registered shape prior	0.938	0.894	0.910
Khened et al. (29)	2D Dense U-net with inception module	0.941	0.894	0.907
Baumgartner et al. (25)	2D U-net with cross entropy loss	0.937	0.897	0.908
Zotti et al. (110)	2D GridNet with registered shape prior	0.931	0.890	0.912
Jang et al. (80)	2D M-Net with weighted cross entropy loss	0.940	0.885	0.907
Painchaud et al. (61)	FCN followed by an AE for shape correction	0.936	0.889	0.909
Wolterink et al. (93)	Multi-input 2D dilated FCN, segmenting paired ED and ES frames simultaneously	0.940	0.885	0.900
Patravali et al. (92)	2D U-net with a Dice loss	0.920	0.890	0.865
Rohé et al. (111)	Multi-atlas based method combined with 3D CNN for registration	0.929	0.868	0.881
Tziritas and Grinias (112)	Level-set + markov random field (MRF); *Non-deep learning method*	0.907	0.798	0.803
Yang et al. (81)	3D FCN with deep supervision	0.820	N/A	0.780

All the methods were evaluated on the same test set (50 subjects). Bold numbers are the highest overall Dice values for the corresponding structure. LV, left ventricle cavity; RV, right ventricle cavity; Myo, left ventricular myocardium; ED, end-diastolic; ES, end-systolic. Last update: 2019.8.1.
Note that for simplicity, we report the average Dice scores for each structure over ED and ES phases. More detailed comparison for different phases can be found on the public leaderboard in the post-testing part (https://acdc.creatis.insa-lyon.fr) as well as corresponding published works in this table.

views. Notably, their networks can also be trained to segment ventricles from 2D short-axis stacks without any modifications to the network architecture. Likewise, Xiong et al. (100), Preetha et al. (187), Bian et al. (188), and Chen et al. (34) applied 2D FCNs to segment the atrium from 3D LGE images in a slice-by-slice fashion, where they optimized the network structure for enhanced feature learning. 3D networks (54, 189–192) and multi-view FCN (99, 193) have also been explored to capture 3D global information from 3D LGE images for accurate atrium segmentation.

In particular, Xia et al. (54) proposed a fully automatic two-stage segmentation framework which contains a first 3D U-net to roughly locate the atrial center from down-sampled images followed by a second 3D U-net to accurately segment the atrium in the cropped portions of the original images at full resolution. Their multi-stage approach is both memory-efficient and accurate, ranking first in the left atrium segmentation challenge 2018 (LASC'18) with a mean Dice score of 0.93 evaluated on a test set of 54 cases.

3.1.3. Scar Segmentation

Scar characterization is usually performed using LGE MR imaging, a contrast-enhanced MR imaging technique. LGE MR imaging enables the identification of myocardial scars and atrial fibrosis, allowing improved management of myocardial infarction and atrial fibrillation (194). Prior to the advent of deep learning, scar segmentation was often performed using intensity thresholding-based or clustering methods which are sensitive to the local intensity changes (103). The main limitation of these methods is that they usually require the manual segmentation of the region of interest to reduce the search space and the computational costs (195). As a result, these semi-automated methods are not suitable for large-scale studies or clinical deployment.

Deep learning approaches have been combined with traditional segmentation methods for the purpose of scar segmentation: Yang et al. (101, 196) applied an atlas-based method to identify the left atrium and then applied deep neural networks to detect fibrotic tissue in that region. Relatively to end-to-end approaches, Chen et al. (102) applied deep neural networks to segment both the left atrium and the atrial scars. In particular, the authors employed a multi-view CNN with a recursive attention module to fuse features from complementary views for better segmentation accuracy. Their approach achieved a mean Dice score of 0.90 for the LA region and a mean Dice score of 0.78 for atrial scars.

In the work of Fahmy et al. (197), the authors applied a U-net based network to segment the myocardium and the scars at the same time from LGE images acquired from patients with hypertrophic cardiomyopathy (HCM), achieving a fast segmentation speed. However, the reported segmentation accuracy for the scar regions was relatively low (mean Dice: 0.58). Zabihollahy et al. (103) and Moccia et al. (104) instead adopted a semi-automated method which requires a manual segmentation of the myocardium followed by the application of a 2D network to differentiate scars from normal myocardium. They reported higher segmentation accuracy on their test sets (mean Dice >0.68). At the moment, fully-automated scar segmentation is still a challenging task since the infarcted regions in patients can lead to kinematic variabilities and abnormalities in those contrast-enhanced images. Interestingly, Xu et al. (105) developed an RNN which leverages motion patterns to automatically delineate myocardial infarction area from cine MR image sequences without contrast agents. Their method achieved a high overall Dice score of 0.90 when compared to the manual annotations on LGE MR images, providing a novel approach for infarction assessment.

3.1.4. Aorta Segmentation

The segmentation of the aortic lumen from cine MR images is essential for accurate mechanical and hemodynamic characterization of the aorta. One common challenge for this task is the typical sparsity of the annotations in aortic cine image sequences, where only a few frames have been annotated. To address the problem, Bai et al. (32) applied a non-rigid image registration method (198) to propagate the labels from the annotated frames to the unlabeled neighboring ones in the cardiac cycle, effectively generating pseudo annotated frames that could be utilized for further training. This semi-supervised method achieved an average Dice metric of 0.96 for the ascending aorta and 0.95 for the descending aorta over a test set of 100 subjects. In addition, compared to a previous approach based on deformable models (199), their approach based on FCN and RNN can directly perform the segmentation task on a whole image sequence without requiring the explicit estimation of the ROI.

3.1.5. Whole Heart Segmentation

Apart from the above mentioned segmentation applications which target one particular structure, deep learning can also be applied to segmenting the main substructures of the heart in 3D MR images (30, 106, 107, 200). An early work from Yu et al. (30) adopted a 3D dense FCN to segment the myocardium and blood pool in the heart from 3D MR scans. Recently, more and more methods began to apply deep learning pipelines to segment more specific substructures [including four chambers, aorta, pulmonary vein (PV)] in both 3D CT and MR images. This has been facilitated by the availability of a public dataset for whole heart segmentation [Multi-Modality Whole Heart Segmentation (MM-WHS)] which consists of both CT and MRI images. We will discuss these segmentation methods in the next CT section in further detail (see section 3.2.1).

3.2. Cardiac CT Image Segmentation

CT is a non-invasive imaging technique that is performed routinely for disease diagnosis and treatment planning. In particular, cardiac CT scans are used for the assessment of cardiac anatomy and specifically the coronary arteries. There are two main imaging modalities: non-contrast CT imaging and contrast-enhanced coronary CT angiography (CTA). Typically, non-contrast CT imaging exploits density of tissues to generate an image, such that different densities using various attenuation values, such as soft tissues, calcium, fat, and air can be easily distinguished, and thus allows to estimate the amount of calcium present in the coronary arteries (201). In comparison, contrast-enhanced coronary CTA, which is acquired after the injection of a contrast agent, can provide excellent visualization of cardiac chambers, vessels and coronaries, and has been shown to be effective in detecting non-calcified coronary plaques. In the following sections, we will review some of the most commonly used deep learning-based cardiac CT segmentation methods. A summary of these approaches is presented in **Table 3**.

3.2.1. Cardiac Substructure Segmentation

Accurate delineation of cardiac substructures plays a crucial role in cardiac function analysis, providing important clinical variables, such as EF, myocardial mass, wall thickness etc. Typically, the cardiac substructures that are segmented include the LV, RV, LA, RA, Myo, aorta (AO), and pulmonary artery (PA).

3.2.1.1. Two-step segmentation

One group of deep learning methods relies on a two-step segmentation procedure, where a ROI is first extracted and then fed into a CNN for subsequent classification (113, 202). For instance, Zreik et al. (113) proposed a two-step LV segmentation process where a bounding box for the LV is first detected using the method described in de Vos et al. (203), followed by a voxel classification within the defined bounding box using a patch-based CNN. More recently, FCN, especially U-net (49), has become the method of choice for cardiac CT segmentation. Zhuang et al. (19) provides a comparison of a group of methods (36, 114, 115, 117, 118, 137) for whole heart segmentation (WHS) that have been evaluated on the MM-WHS challenge. Several of these methods (37, 114–116) combine a localization network, which produces a coarse detection of the heart, with 3D FCNs applied to the detected ROI for segmentation. This allows the segmentation network to focus on the anatomically relevant regions, and has shown to be effective for whole heart segmentation. A summary of the comparison between the segmentation accuracy of the methods evaluated on MM-WHS dataset is presented in **Table 4**. These methods generally achieve better segmentation accuracy on CT images compared to that of MR images, mainly because of the smaller variations in image intensity distribution across different CT scanners and better image quality (19). For a detailed discussion on these listed methods, please refer to Zhuang et al. (19).

3.2.1.2. Multi-view CNNs

Another line of research utilizes the volumetric information of the heart by training multi-planar CNNs (axial, sagittal, and coronal views) in a 2D fashion. Examples include Wang et al. (117) and Mortazi et al. (118) where three independent orthogonal CNNs were trained to segment different views. Specifically, Wang et al. (117) additionally incorporated shape context in the framework for the segmentation refinement, while Mortazi et al. (118) adopted an adaptive fusion strategy to combine multiple outputs utilizing complementary information from different planes.

3.2.1.3. Hybrid loss

Several methods employ a hybrid loss, where different loss functions (such as focal loss, Dice loss, and weighted categorical cross-entropy) are combined to address the class imbalance issue, e.g., the volume size imbalance among different ventricular structures, and to improve the segmentation performance (36, 119).

In addition, the work of Zreik et al. (120) has proposed a method for the automatic identification of patients with significant coronary artery stenoses through the segmentation

TABLE 3 | A summary of selected deep learning methods on cardiac CT segmentation.

Application	Selected works	Description	Imaging modality	Structure(s)
Cardiac substructure segmentation	**Two-step segmentation**			
	Zreik et al. (113)	Patch-based CNN	CTA	Myo
	Payer et al. (114)	A pipeline of two FCNs	MR/CT	WHS
	Tong et al. (115)	Deeply supervised 3D U-net	MR/CT	WHS
	Wang et al. (116)	Two-stage 3D U-net with dynamic ROI extraction	MR/CT	WHS
	Xu et al. (37)	Faster RCNN and U-net	CT	WHS
	Multi-view CNNs			
	Wang and Smedby (117)	Orthogonal 2D U-nets with shape context	MR/CT	WHS
	Mortazi et al. (118)	Multi-planar FCNs with an adaptive fusion strategy	MR/CT	WHS
	Hybrid loss			
	Yang et al. (36)	3D U-net with deep supervision	MR/CT	WHS
	Ye et al. (119)	3D deeply-supervised U-net with multi-depth fusion	CT	WHS
	Others			
	Zreik et al. (120)	Multi-scale FCN	CTA	Myo
	Joyce et al. (121)	Unsupervised segmentation with GANs	MR/CT	LV, Myo, RV
Coronary artery segmentation	**End-to-end CNNs**			
	Moeskops et al. (122)	Multi-task CNN	CTA	Vessel
	Merkow et al. (38)	3D U-net with deep multi-scale supervision	CTA	Vessel
	Lee et al. (123)	Template transformer network	CTA	Vessel
	CNN as pre-/post-processing			
	Gülsün et al. (124)	CNN as path pruning	CTA	coronary artery centerline
	Guo et al. (125)	Multi-task FCN with a minimal patch extractor	CTA	Coronary artery centerline
	Shen et al. (126)	3D FCN with level set	CTA	Vessel
	Others			
	Wolterink et al. (127)	CNN to estimate direction classification and radius regression	CTA	Coronary artery centerline
	Wolterink et al. (128)	Graph convolutional network	CTA	Vessel
Coronary artery calcium and plaque segmentation	**Two-step segmentation**			
	Wolterink et al. (129)	CNN pairs	CTA	CAC
	Lessmann et al. (130)	Multi-view CNNs	CT	CAC
	Lessmann et al. (131)	Two consecutive CNNs	CT	CAC
	Liu et al. (132)	3D vessel-focused ConvNets	CTA	CAC, NCP, MCP
	Direct segmentation			
	Santini et al. (133)	Patch-based CNN	CT	CAC
	Shadmi et al. (134)	U-net and FC DenseNet	CT	CAC
	Zhang et al. (135)	U-DenseNet	CT	CAC
	Ma and Zhang (136)	DenseRAU-net	CT	CAC

and analysis of the LV myocardium. In this work, a multi-scale FCN is first employed for myocardium segmentation, and then a convolutional autoencoder is used to characterize the LV myocardium, followed by a support vector machine (SVM) to classify patients based on the extracted features.

3.2.2. Coronary Artery Segmentation

Quantitative analysis of coronary arteries is an important step for the diagnosis of cardiovascular diseases, stenosis grading, blood flow simulation and surgical planning (204). Though this topic has been studied for years (4), only a small number of works

TABLE 4 | Segmentation accuracy of methods validated on MM-WHS dataset.

Methods	LV	RV	LA	RA	Myo	AO	PA	WHS
Payer et al. (114)	91.8/91.6	**90.9**/86.8	92.9/85.5	**88.8**/88.1	88.1/77.8	93.3/88.8	84.0/73.1	**90.8**/86.3
Yang et al. (137)	92.3/75.0	85.7/75.0	**93.0**/82.6	87.1/85.9	85.6/65.8	89.4/80.9	83.5/72.6	89.0/78.3
Mortazi et al. (118)	90.4/87.1	88.3/83.0	91.6/81.1	83.6/75.9	85.1/74.7	90.7/83.9	78.4/71.5	87.9/81.8
Tong et al. (115)	89.3/70.2	81.0/68.0	88.9/67.6	81.2/65.4	83.7/62.3	86.8/59.9	69.8/47.0	84.9/67.4
Wang et al. (116)	80.0/86.3	78.6/84.9	90.4/85.2	79.4/84.0	72.9/74.4	87.4/82.4	64.8/78.8	80.6/83.2
Ye et al. (119)	**94.4**/–	89.5/–	91.6/–	87.8/–	**88.9**/–	**96.7**/–	**86.2**/–	90.7/–
Xu et al. (37)	87.9/–	90.2/–	83.2/–	84.4/–	82.2/–	91.3/–	82.1/–	85.9/–

The training set contains 20 CT and 20 MRI whereas the test set contains 40 CT and 40 MRI. Reported numbers are Dice scores (CT/MRI) for different substructures on both CT and MRI scans. For more detailed comparisons, please refer to Zhuang et al. (19). The bold number in each column represents the highest score for the corresponding structure on CT images.

investigate the use of deep learning in this context. Methods relating to coronary artery segmentation can be mainly divided into two categories: centerline extraction and lumen (i.e., vessel wall) segmentation.

3.2.2.1. CNNs as a post-/pre-processing step

Coronary centerline extraction is a challenging task due to the presence of nearby cardiac structures and coronary veins as well as motion artifacts in cardiac CT. Several deep learning approaches employ CNNs as either a post-processing or pre-processing step for traditional methods. For instance, Gülsün et al. (124) formulated centerline extraction as finding the maximum flow paths in a steady state porous media flow, with a learning-based classifier estimating anisotropic vessel orientation tensors for flow computation. A CNN classifier was then employed to distinguish true coronary centerlines from leaks into non-coronary structures. Guo et al. (125) proposed a multi-task FCN centerline extraction method that can generate a single-pixel-wide centerline, where the FCN simultaneously predicted centerline distance maps and endpoint confidence maps from coronary arteries and ascending aorta segmentation masks, which were then used as input to the subsequent minimal path extractor to obtain the final centerline extraction results. In contrast, unlike the aforementioned methods that used CNNs either as a pre-processing or post-processing step, Wolterink et al. (127) proposed to address centerline extraction via a 3D dilated CNN, where the CNN was trained on patches to directly determine a posterior probability distribution over a discrete set of possible directions as well as to estimate the radius of an artery at the given point.

3.2.2.2. End-to-end CNNs

With respect to the lumen or vessel wall segmentation, most deep learning based approaches use an end-to-end CNN segmentation scheme to predict dense segmentation probability maps (38, 122, 126, 205). In particular, Moeskops et al. (122) proposed a multi-task segmentation framework where a single CNN can be trained to perform three different tasks including coronary artery segmentation in cardiac CTA and tissue segmentation in brain MR images. They showed that such a multi-task segmentation network in multiple modalities can achieve equivalent performance as a single task network. Merkow et al. (38) introduced deep multi-scale supervision into a 3D U-net architecture, enabling efficient multi-scale feature learning and precise voxel-level predictions. Besides, shape priors can also be incorporated into the network (123, 206, 207). For instance, Lee et al. (123) explicitly enforced a roughly tubular shape prior for the vessel segments by introducing a template transformer network, through which a shape template can be deformed via network-based registration to produce an accurate segmentation of the input image, as well as to guarantee topological constraints. More recently, graph convolutional networks have also been investigated by Wolterink et al. (128) for coronary artery segmentation in CTA, where vertices on the coronary lumen surface mesh were considered as graph nodes and the locations of these tubular surface mesh vertices were directly optimized. They showed that such method significantly outperformed a baseline network that used only fully-connected layers on healthy subjects (mean Dice score: 0.75 vs. 0.67). Besides, the graph convolutional network used in their work is able to directly generate smooth surface meshes without post-processing steps.

3.2.3. Coronary Artery Calcium and Plaque Segmentation

Coronary artery calcium (CAC) is a direct risk factor for cardiovascular disease. Clinically, CAC is quantified using the Agatston score (208) which considers the lesion area and the weighted maximum density of the lesion (209). Precise detection and segmentation of CAC are thus important for the accurate prediction of the Agatston score and disease diagnosis.

3.2.3.1. Two-step segmentation

One group of deep learning approaches to segmentation and automatic calcium scoring proposed to use a two-step segmentation scheme. For example, Wolterink et al. (129) attempted to classify CAC in cardiac CTA using a pair of CNNs, where the first CNN coarsely identified voxels likely to be CAC within a ROI detected using De et al. (203) and then the second CNN further distinguished between CAC and CAC-like negatives more accurately. Similar to such a two-stage scheme, Lessmann et al. (130, 131) proposed to identify CAC in low-dose chest CT, in which a ROI of the heart or potential calcifications were first localized followed by a CAC classification process.

3.2.3.2. Direct segmentation

More recently, several approaches (133–136) have been proposed for the direct segmentation of CAC from non-contrast cardiac CT or chest CT: the majority of them employed combinations of U-net (49) and DenseNet (78) for precise quantification of CAC which showed that a sensitivity over 90% can be achieved (133). These aforementioned approaches all follow the same workflow where the CAC is first identified and then quantified. An alternative approach is to circumvent the intermediate segmentation and to perform direct quantification, such as in de Vos et al. (209) and Cano-Espinosa et al. (210), which have proven that this approach is effective and promising.

Finally, for non-calcified plaque (NCP) and mixed-calcified plaque (MCP) in coronary arteries, only a limited number of works have been reported that investigate deep learning methods for segmentation and quantification (132, 211). Yet, this is a very important task from a clinical point of view, since these plaques can potentially rupture and obstruct an artery, causing ischemic events and severe cardiac damage. In contrast to CAC segmentation, NCP and MCP segmentation are more challenging due to their similar appearances and intensities as adjacent tissues. Therefore, robust and accurate analysis often requires the generation of multi-planar reformatted (MPR) images that have been straightened along the centerline of the vessel. Recently, Liu et al. (132) proposed a vessel-focused 3D convolutional network with attention layers to segment three types of plaques on the extracted and reformatted coronary MPR volumes. Zreik et al. (211) presented an automatic method for detection and characterization of coronary artery plaques as well as determination of coronary artery stenosis significance, in which a multi-task convolutional RNN was used to perform both plaque and stenosis classification by analyzing the features extracted along the coronary artery in an MPR image.

3.3. Cardiac Ultrasound Image Segmentation

Cardiac ultrasound imaging, also known as echocardiography, is an indispensable clinical tool for the assessment of cardiovascular function. It is often used clinically as the first imaging examination owing to its portability, low cost and real-time capability. While a number of traditional methods, such as active contours, level-sets and active shape models have been employed to automate the segmentation of anatomical structures in ultrasound images (212), the achieved accuracy is limited by various problems of ultrasound imaging, such as low signal-to-noise ratio, varying speckle noise, low image contrast (especially between the myocardium and the blood pool), edge dropout and shadows cast by structures, such as dense muscle and ribs.

As in cardiac MR and CT, several DL-based methods have been recently proposed to improve the performance of cardiac ultrasound image segmentation in terms of both accuracy and speed. The majority of these DL-based approaches focus on LV segmentation, with only few addressing the problem of aortic valve and LA segmentation. A summary of the reviewed works can be found in **Table 5**.

3.3.1. 2D LV Segmentation

3.3.1.1. Deep learning combined with deformable models

The imaging quality of echocardiography makes voxel-wise tissue classification highly challenging. To address this challenge, deep learning has been combined with deformable model for LV segmentation in 2D images (138, 139, 141–145). Features extracted by trained deep neural networks were used instead of handcrafted features to improve accuracy and robustness.

Several works applied deep learning in a two-stage pipeline which first localizes the target ROI via rigid transformation of a bounding box, then segments the target structure within the ROI. This two-stage pipeline reduces the search region of the segmentation and increases robustness of the overall segmentation framework. Carneiro et al. (138, 139) first adopted this DL framework to segment the LV in apical long-axis echocardiograms. The method uses DBN (213) to predict the rigid transformation parameters for localization and the deformable model parameters for segmentation. The results demonstrated the robustness of DBN-based feature extraction to image appearance variations. Nascimento and Carneiro (140) further reduced the training and inference complexity of the DBN-based framework by using sparse manifold learning in the rigid detection step.

To further reduce the computational complexity, some works perform segmentation in one step without resorting to the two-stage approach. Nascimento and Carneiro (141, 142) applied sparse manifold learning in segmentation, showing a reduced training and search complexity compared to their previous version of the method, while maintaining the same level of segmentation accuracy. Veni et al. (143) applied a FCN to produce coarse segmentation masks, which is then further refined by a level-set based method.

3.3.1.2. Utilizing temporal coherence

Cardiac ultrasound data is often recorded as a temporal sequence of images. Several approaches aim to leverage the coherence between temporally close frames to improve the accuracy and robustness of the LV segmentation. Carneiro and Nascimento (144, 145) proposed a dynamic modeling method based on a sequential monte carlo (SMC) (or particle filtering) framework with a transition model, in which the segmentation of the current cardiac phase depends on previous phases. The results show that this approach performs better than the previous method (138) which does not take temporal information into account. In a more recent work, Jafari et al. (146) combined U-net, long-short term memory (LSTM) and inter-frame optical flow to utilize multiple frames for segmenting one target frame, demonstrating improvement in overall segmentation accuracy. The method was also shown to be more robust to image quality variations in a sequence than single-frame U-net.

3.3.1.3. Utilizing unlabeled data

Several works proposed to use non-DL based segmentation algorithms to help generating labels on unlabeled images, effectively increasing the amount of training data. To achieve this, Carneiro and Nascimento (147, 148) proposed on-line retraining strategies where segmentation network (DBN) is firstly initialized

TABLE 5 | A summary of reviewed deep learning methods for ultrasound image segmentation.

Application	Selected works	Method	Structure	Imaging modality
2D LV	**Combined with deformable models**			
	Carneiro et al. (138, 139)	DBN with two-step approach: localization and fine segmentation	LV	2D A2C, A4C
	Nascimento and Carneiro (140)	deep belief networks (DBN) and sparse manifold learning for the localization step	LV	2D A2C, A4C
	Nascimento and Carneiro (141, 142)	DBN and sparse manifold learning for one-step segmentation	LV	2D A2C, A4C
	Veni et al. (143)	FCN (U-net) followed by level-set based deformable model	LV	2D A4C
	Utilizing temporal coherence			
	Carneiro and Nascimento (144, 145)	DBN and particle filtering for dynamic modeling	LV	2D A2C, A4C
	Jafari et al. (146)	U-net and LSTM with additional optical flow input	LV	2D A4C
	Utilizing unlabeled data			
	Carneiro and Nascimento (147, 148)	DBN on-line retrain using external classifier as additional supervision	LV	2D A2C, A4C
	Smistad et al. (149)	U-Net trained using labels generated by a Kalman filter based method	LV and LA	2D A2C, A4C
	Yu et al. (150)	Dynamic CNN fine-tuning with mitral valve tracking to separate LV from LA	Fetal LV	2D
	Jafari et al. (151)	U-net with TL-net (152) based shape constraint on unannotated frames	LV	2D A4C
	Utilizing data from multiple domains			
	Chen et al. (153)	FCN trained using annotated data of multiple anatomical structures	Fetal head and LV	2D head, A2-5C
	Others			
	Smistad et al. (154)	Real time CNN view-classification and segmentation	LV	2D A2C, A4C
	Leclerc et al. (155)	U-net trained on a large heterogeneous dataset	LV, Myo	2D A4C
	Jafari et al. (156)	Real-time mobile software, lightweight U-Net, multitask and adversarial training	LV	2D A2C, A4C
3D LV	Dong et al. (157)	CNN for 2D coarse segmentation refined by 3D snake model	LV	3D (CETUS)
	Oktay et al. (59)	U-net with TL-net based shape constraint	LV	3D (CETUS)
	Dong et al. (158)	Atlas-based segmentation using DL registration and adversarial training	LV	3D
Others	Ghesu et al. (159)	Marginal space learning and adaptive sparse neural network	Aortic valves	3D
	Degel et al. (160)	V-net with TL-net based shape constraint and GAN-based domain adaptation	LA	3D
	Zhang et al. (42)	CNN for view-classification, segmentation and disease detection	Multi-chamber	2D PLAX, PSAX, A2-4C

A[X]C is short for Apical [X]-chamber view. PLAX/PSAX, parasternal long-axis/short-axis; CETUS, using the dataset from Challenge on Endocardial Three-dimensional Ultrasound Segmentation.

using a small set of labeled data and then applied to non-labeled data to propose annotations. The proposed annotations are then checked by external classifiers before being used to re-train the network. Smistad et al. (149) trained a U-net using images annotated by a Kalman filtering based method (214) and illustrated the potential of using this strategy for pre-training. Alternatively, some works proposed to exploit unlabeled data without using additional segmentation algorithm. Yu et al. (150) proposed to train a CNN on a partially labeled dataset of multiple sequences, then fine-tuned the network for each individual sequence using manual segmentation of the first frame as well as CNN-produced label of other frames. Jafari et al. (151) proposed a semi-supervised framework which enables training on both the labeled and unlabeled images. The framework uses an additional generative network, which is trained to generate ultrasound images from segmentation masks, as additional supervision for the unlabeled frames in the sequences. The generative network forces the segmentation network to predict segmentation that can be used to successfully generate the input ultrasound image.

3.3.1.4. Utilizing data from multiple domains

Apart from exploiting unlabeled data in the same domain, leveraging manually annotated data from multiple domains (e.g., different 2D ultrasound views with various anatomical

structures) can also help to improve the segmentation in one particular domain. Chen et al. (153) proposed a novel FCN-based network to utilize multi-domain data to learn generic feature representations. Combined with an iterative refinement scheme, the method has shown superior performance in detection and segmentation over traditional database-guided method (215), FCN trained on single-domain and other multi-domain training strategies.

3.3.1.5. Others

The potential of CNN in segmentation has motivated the collection and labeling of large-scale datasets. Several methods have since shown that deep learning methods, most notably CNN-based methods, are capable of performing accurate segmentation directly without complex modeling and post-processing. Leclerc et al. (155) performed a study to investigate the effect of the size of annotated data for the segmentation of the LV in 2D ultrasound images using a simple U-net. The authors demonstrated that the U-net approach significantly benefits from larger amounts of training data. In addition to performance on accuracy, some work investigated the computational efficiency of DL-based methods. Smistad et al. (154) demonstrated the efficiency of CNN-based methods by successfully performing real-time view-classification and segmentation. Jafari et al. (156) developed a software pipeline capable of real-time automated LV segmentation, landmark detection and LV ejection fraction calculation on a mobile device taking input from point-of-care ultrasound (POCUS) devices. The software uses a lightweight U-net trained using multi-task learning and adversarial training, which achieves EF prediction error that is lower than inter- and intra- observer variability.

3.3.2. 3D LV Segmentation

Segmenting cardiac structures in 3D ultrasound is even more challenging than 2D. While having the potential to derive more accurate volume-related clinical indices, 3D echocardiograms suffer from lower temporal resolution and lower image quality compared to 2D echocardiograms. Moreover, 3D images dramatically increase the dimension of parameter space of neural networks, which poses computational challenges for deep learning methods.

One way to reduce the computational cost is to avoid direct processing of 3D data in deep learning networks. Dong et al. (157) proposed a two-stage method by first applying a 2D CNN to produce coarse segmentation maps on 2D slices from a 3D volume. The coarse 2D segmentation maps are used to initialize a 3D shape model which is then refined by 3D deformable model method (216). In addition, the authors used transfer learning to side-step the limited training data problem by pre-training network on a large natural image segmentation dataset and then fine-tuning to the LV segmentation task.

Anatomical shape priors have been utilized to increase the robustness of deep learning-based segmentation methods to challenging 3D ultrasound images. Oktay et al. (59) proposed an anatomically constrained network where a shape constraint-based loss is introduced to train a 3D segmentation network. The shape constraint is based on the shape prior learned from segmentation maps using auto-encoders (152). Dong et al. (158) utilized shape prior more explicitly by combining a neural network with a conventional atlas-based segmentation framework. Adversarial training was also applied to encourage the method to produce more anatomically plausible segmentation maps, which contributes to its superior segmentation performance comparing to a standard voxel-wise classification 3D segmentation network (52).

3.3.3. Left Atrium Segmentation

Degel et al. (160) adopted the aforementioned anatomical constraints in 3D LA segmentation to tackle the domain shift problem caused by variation of imaging device, protocol and patient condition. In addition to the anatomically constraining network, the authors applied an adversarial training scheme (217) to improve the generalizability of the model to unseen domain.

3.3.4. Multi-Chamber Segmentation

Apart from LV segmentation, a few works (23, 42, 149) applied deep learning methods to perform multi-chamber (including LV and LA) segmentation. In particular, (42) demonstrated the applicability of CNNs on three tasks: view classification, multi-chamber segmentation and detection of cardiovascular diseases. Comprehensive validation on a large (non-public) clinical dataset showed that clinical metrics derived from automatic segmentation are comparable or superior than manual segmentation. To resemble real clinical situations and thus encourages the development and evaluation of robust and clinically effective segmentation methods, a large-scale dataset for 2D cardiac ultrasound has been recently made public (23). The dataset and evaluation platform were released following the preliminary data requirement investigation of deep learning methods (155). The dataset is composed of apical 4-chamber view images annotated for LV and LA segmentation, with uneven imaging quality from 500 patients with varying conditions. Notably, the initial benchmarking (23) on this dataset has shown that modern encoder-decoder CNNs resulted in lower error than inter-observer error between human cardiologists.

3.3.5. Aortic Valve Segmentation

Ghesu et al. (159) proposed a framework based on marginal space learning (MSL), Deep neural networks (DNNs) and active shape model (ASM) to segment the aortic valve in 3D cardiac ultrasound volumes. An adaptive sparsely-connected neural network with reduced number of parameters is used to predict a bounding box to locate the target structure, where the learning of the bounding box parameters is marginalized into sub-spaces to reduce computational complexity. This framework showed significant improvement over the previous non-DL MSL (218) method while achieving competitive run-time.

3.4. Discussion

So far, we have presented and discussed recent progress of deep learning-based segmentation methods in the three modalities (i.e., MR, CT, ultrasound) that are commonly used in the assessment of cardiovascular disease. To summarize, current state-of-the-art segmentation methods are mainly based on

CNNs that employ the FCN or U-net architecture. In addition, there are several commonalities in the FCN-based methods for cardiac segmentation which can be categorized into four groups: (1) enhancing network feature learning by employing advanced building blocks in networks (e.g., inception module, dilated convolutions), most of which have been mentioned earlier (section 2.1.6); (2) alleviating the problem of class imbalance with advanced loss functions (e.g., weighted loss functions); (3) improving the networks' generalization ability and robustness through a multi-stage pipeline, multi-task learning, or multi-view feature fusion; (4) forcing the network to generate more anatomically-plausible segmentation results by incorporating shape priors, applying adversarial loss or anatomical constraints to regularize the network during training. It is also worthwhile to highlight that for cardiac image sequence segmentation (e.g., cine MR images, 2D ultrasound sequences), leveraging spatial and temporal coherence from these sequences with advanced neural networks [e.g., RNN (32, 146), multi-slice FCN (27)] has been explored and shown to be beneficial for improving the segmentation accuracy and temporal consistency of the segmentation maps.

While the results reported in the literature show that neural networks have become more sophisticated and powerful, it is also clear that performance has improved with the increase of publicly available training subjects. A number of DL-based methods (especially in MRI) have been trained and tested on public challenge datasets, which not only provide large amounts of data to exploit the capabilities of deep learning in this domain, but also a platform for transparent evaluation and comparison. In addition, many of the participants in these challenges have shared their code with other researchers via open-source community websites (e.g., Github). Transparent and fair benchmarking and sharing of code are both essential for continued progress in this domain. We summarize the existing public datasets in **Table 6** and public code repositories in **Table 7** for reference.

An interesting conclusion supported by **Table 7** is that the target image type can affect the choice of network structures (i.e., 2D networks, 3D networks). For 3D imaging acquisitions, such as LGE-MRI and CT images, 3D networks are preferred whereas 2D networks are more popular approaches for segmenting cardiac cine short-axis or long-axis image stacks. One reason for using 2D networks for the segmentation of short-axis or long-axis images is their typically large slice thickness (usually around 7–8 mm) which can further exacerbated by inter-slice gaps. In addition, breath-hold related motion artifacts between different slices may negatively affect 3D networks. A study conducted by Baumgartner et al. (25) has shown that a 3D U-net performs worse than a 2D U-net when evaluated on the ACDC challenge dataset. By contrast, in the LASC'18 challenge mentioned in **Table 6**, which uses high-resolution 3D images, most participants applied 3D networks and the best performance was achieved by a cascaded network based on the 3D U-net (54).

It is well-known that training 3D networks is more difficult than training 2D networks. In general, 3D networks have significantly more parameters than 2D networks. Therefore, 3D networks are more difficult and computationally expensive to optimize as well as prone to over-fitting, especially if the training data is limited. As a result, several researchers have tried to carefully design the structure of network to reduce the number of parameters for a particular application and have also applied advanced techniques (e.g., deep supervision) to alleviate the over-fitting problem (30, 54). For this reason, 2D-based networks (e.g., 2D U-net) are still the most popular segmentation approaches for all three modalities.

In addition to 2D and 3D networks, several authors have proposed "2D+" networks that have been shown to be effective in segmenting structures from cardiac volumetric data. These "2D+" networks are mainly based on 2D networks, but are adapted with increased capacity to utilize 3D context. These networks include multi-view networks which leverage multi-planar information (i.e., coronal, sagittal, axial views) (99, 117), multi-slice networks, and 2D FCNs combined with RNNs which incorporate context across multiple slices (33, 55, 92, 169). These "2D+" networks inherit the advantages of 2D networks while still being capable of leveraging through-plane spatial context for more robust segmentation with strong 3D consistency.

Finally, it is worth to note that there is no universally optimal segmentation method. Different applications have different complexities and different requirements, meaning that customized algorithms need to be optimized. For example, while anatomical shape constraints can be applied to cardiac anatomical structure segmentation (e.g., ventricle segmentation) to boost the segmentation performance, those constraints may not be suitable for the segmentation of pathologies or lesions (e.g., scar segmentation) which can have arbitrary shapes. Also, even if the target structure in two applications are the same, the complexity of the segmentation task can vary significantly from one to another, especially when their underlying imaging modalities and patient populations are different. For example, directly segmenting the left ventricle myocardium from contrast-enhanced MR images (e.g., LGE images) is often more difficult than from MR images without contrast agents, as the anatomical structures are more attenuated by the contrast agent. For cases with certain diseases (e.g., myocardial infarction), the border between the infarcted region and blood pool appears blurry and ambiguous to delineate. As a result, a segmentation network designed for non-contrast enhanced images may not be directly applied to contrast-enhanced images (100). A more sophisticated algorithm is generally required to assist the segmentation procedure. Potential solutions include applying dedicated image pre-processing, enhancing network capacity, adding shape constraints, and integrating specific knowledge about the application.

4. CHALLENGES AND FUTURE WORK

It is evident from the literature that deep learning methods have matched or surpassed the previous state of the art in various cardiac segmentation applications, mainly benefiting from the increased size of public datasets and the emergence of advanced network architectures as well as powerful hardware for computing. Given this rapid process, one may wonder if deep learning methods can be directly deployed to real-world

TABLE 6 | Summary of public datasets on cardiac segmentation for the three modalities.

Dataset Name/ References	Year	Main modalities	# of subjects	Target(s)	Main pathology
York (10)	2008	cine MRI	33	LV, Myo	Cardiomyopathy, aortic regurgitation, enlarged ventricles and ischemia
Sunnybrook (11)	2009	cine MRI	45	LV, Myo	Hypertrophy, heart failure w./w.o infarction
LVSC (12)	2011	cine MRI	200	LV, Myo	Coronary artery disease, myocardial infarction.
RVSC (1)	2012	cine MRI	48	RV	Myocarditis, ischemic cardiomyopathy, suspicion of arrhythmogenic, right ventricular dysplasia, dilated cardiomyopathy, hypertrophic cardiomyopathy, aortic stenosis
cDEMRIS (13)	2012	LGE MRI	60	LA fibrosis and scar	Atrial fibrillation
LVIC (14)	2012	LGE MRI	30	Myocardial scars	Ischemic cardiomyopathy
LASC'13 (15)	2013	3D MRI	30	LA	N/A
HVSMR (16)	2016	3D MRI	4	Blood pool, myocardium of the heart	Congenital heart defects
ACDC (17)	2017	cine MRI	150	LV, Myo; RV	Mycardial infarction, dilated/hypertrophic cardiomyopathy, abnormal RV
LASC'18 (18)	2018	LGE MRI	154	LA	Atrial fibrillation
MM-WHS (19)	2017	CT/MRI	60/60	WHS	Myocardium infarction, atrial fibrillation, tricuspid regurgitation, aortic valve stenosis, Alagille syndrome, Williams syndrome, dilated cardiomyopathy, aortic coarctation, tetralogy of Fallot
CAT08 (20)	2008	CTA	32	Coronary artery centerline	Patients with presence of calcium scored as absent, modest or severe.
CLS12 (21)	2012	CTA	48	Coronary lumen and stenosis	Patients with different levels of coronary artery stenoses.
CETUS (22)	2014	3D Ultrasound	45	LV	Myocardial infarction, dilated cardiomyopathy
CAMUS (23)	2019	2D Ultrasound	500	LV, LA	Patients with EF $< 45\%$

Most of the datasets listed above are from the MICCAI society.

applications to reduce the workload of clinicians. The current literature suggests that there is still a long way to go. In the following paragraphs, we summarize several major challenges in the field of cardiac segmentation and some recently proposed approaches that attempt to address them. These challenges and related works also provide potential research directions for future work in this field.

4.1. Scarcity of Labels

One of the biggest challenges for deep learning approaches is the scarcity of annotated data. In this review, we found that the majority of studies uses a fully supervised approach to train their networks, which requires a large number of annotated images. In fact, annotating cardiac images is time consuming and often requires significant amounts of expertise. These methods can be divided into five classes: data augmentation, transfer learning with fine-tuning, weakly and semi-supervised learning, self-supervised learning, and unsupervised learning.

- **Data augmentation**. Data augmentation aims to increase the size and the variety of training images by artificially generating new samples from existing labeled data. Traditionally, this can be achieved by applying a stack of geometric or photometric transformations to existing image-label pairs. These transformations can be affine transformations, adding random noise to the original data, or adjusting image contrast. However, designing an effective pipeline of data augmentation often requires domain knowledge, which may not be easily extendable to different applications. And the diversity of augmented data may still be limited, failing to reflect the spectrum of real-world data distributions. Most recently, several researchers have began to investigate the use of generative models [e.g., GANs, variational AE (219)], reinforcement learning (220), and adversarial example generation (221) to directly learn task-specific augmentation strategies from existing data. In particular, the generative model-based approach has been proven to be effective for one-shot brain segmentation (222) and few-shot cardiac MR image segmentation (223) and it is thus worth exploring for more applications in the future.
- **Transfer learning with fine-tuning**. Transfer learning aims at reusing a model pre-trained on one task as a starting point to train for a second task. The key of transfer learning is to learn features in the first task that are related to the second task such that the network can quickly converge even with limited data. Several researchers have successfully demonstrated the use of transfer learning to improve the model generalization performance for cardiac ventricle segmentation, where they first trained a model on a large dataset and then fine-tuned it on a small dataset (29, 31, 85, 91, 165).
- **Weakly and semi-supervised learning**. Weakly and semi-supervised learning methods aim at improving the learning accuracy by making use of both labeled and unlabeled or weakly-labeled data (e.g., annotations in forms of scribbles

TABLE 7 | Public code for DL-based cardiac image segmentation.

Modality	Application(s)	References	Basic network	Code repo (If not specified, the repository is located under github.com)
MR (SAX)	Bi-ventricular Segmentation	Tran (24)	2D FCN	vuptran/cardiac-segmentation
MR (SAX)	Bi-ventricular Segmentation	Baumgartner et al. (25)	2D/3D U-net	baumgach/acdc_segmenter
MR (SAX)	Bi-ventricular Segmentation; 1st rank in ACDC challenge	Isensee et al. (26)	2D + 3D U-net (ensemble)	MIC-DKFZ/ACDC2017
MR (SAX)	Bi-ventricular Segmentation	Zheng et al. (27)	Cascaded 2D U-net	julien-zheng/CardiacSegmentationPropagation
MR (SAX)	Bi-ventricular segmentation and Motion Estimation	Qin et al. (28)	2D FCN, RNN	cq615
MR (SAX)	Biventricular Segmentation	Khened et al. (29)	2D U-net	mahendrakhened
MR (3D scans)	WHS	Yu et al. (30)	3D CNN	yulequan/HeartSeg
MR (Multi-view)	Four-chamber Segmentation and Aorta Segmentation	Bai et al. (31, 32)	2D FCN, RNN	baiwenjia/ukbb_cardiac
MR	Cardiac segmentation and motion tracking	Duan et al. (33)	2.5D FCN + Atlas-based	j-duan/4Dsegment
LGE MRI	Left Atrial Segmentation	Chen et al. (34)	2D U-net	cherise215/atria_segmentation_2018
LGE MRI	Left Atrial Segmentation	Yu et al. (35)	3D V-net	yulequan/UA-MT
CT	WHS	Yang et al. (36)	3D U-net	xy0806/miccai17-mmwhs-hybrid
CT	WHS	Xu et al. (37)	Faster RCNN, 3D U-net	Wuziyi616/CFUN
CT, MRI	Coronary arteries	Merkow et al. (38)	3D U-net	jmerkow/I2I
CT, MRI	WHS	Dou et al. (39, 40)	2D CNN	carrenD/Medical-Cross-Modality-Domain-Adaptation
CT, MRI	WHS	Chen et al. (41)	2D CNN	cchen-cc/SIFA
Ultrasound	View classification and four-chamber segmentation	Zhang et al. (42)	2D U-net	bitbucket.org/rahuldeo/echocv

SAX, short-axis view; WHS, whole heart segmentation.

or bounding boxes). In this context, several works have been proposed for cardiac ventricle segmentation in MR images. One approach is to estimate full labels on unlabeled or weakly labeled images for further training. For example, Qin et al. (28) and Bai et al. (32) utilized motion information to propagate labels from labeled frames to unlabeled frames in a cardiac cycle whereas (224, 225) applied the expectation maximization (EM) algorithm to predict and refine the estimated labels recursively. Others have explored different approaches to regularize the network when training on unlabeled images, applying multi-task learning (177, 178), or global constraints (226).

- **Self-supervised learning**. Another approach is self-supervised learning which aims at utilizing labels that are *generated automatically* without human intervention. These labels, designed to encode some properties or semantics of the object, can provide strong supervisory signals to pre-train a network before fine-tuning for a given task. A very recent work from Bai et al. (227) has shown the effectiveness of self-supervised learning for cardiac MR image segmentation where the authors used auto-generated anatomical position labels to pre-train a segmentation network. Compared to a network trained from scratch, networks pre-trained on the self-supervised task performed better, especially when the training data was extremely limited.
- **Unsupervised learning**. Unsupervised learning aims at learning without paired labeled data. Compared to the former four classes, there is limited literature about unsupervised learning methods for cardiac image segmentation, perhaps because of the difficulty of the task. An early attempt has been made which applied adversarial training to train a network segmenting LV and RV from CT and MR images without requiring a training set of paired images and labels (121).

In general, transfer learning and self-supervised learning allow the network to be aware of general knowledge shared across different tasks to accelerate learning procedure and to encourage model generalization. On the other hand, data augmentation, weakly and semi-supervised learning allows the network to get more labeled training data in an efficient way. In practice, the two types of methods can be integrated together to improve the model performance. For example, transfer learning can be applied at the model initialization stage whereas data augmentation can be applied at the model fine-tuning stage.

4.2. Model Generalization Across Various Imaging Modalities, Scanners, and Pathologies

Another common limitation in DL-based methods is that they still lack generalization capabilities when presented with

previously unseen samples (e.g., data from a new scanner, abnormal, and pathological cases that have not been included in the training set). In other words, deep learning models tend to be biased by their respective training datasets. This limitation prevents models to be deployed in the real world and therefore diminishes their impact for improving clinical workflows.

To improve the model performance across MR images acquired from multiple vendors and multiple scanners (53), collected a large multi-vendor, multi-center, heterogeneous labeled training set from patients with cardiovascular diseases. However, this approach may not scale to the real world, as it implies the collection and labeling of a vastly large dataset covering all possible cases. Several researchers have recently started to investigate the use of unsupervised domain adaptation techniques that aim at optimizing the model performance on a target dataset without additional labeling costs. Several works have successfully applied adversarial training to cross-modality segmentation tasks, adapting a cardiac segmentation model learned from MR images to CT images and vice versa (39–41, 228, 229). These type of approaches can also be adopted for semi-supervised learning, where the target domain is a new set of unlabeled data of the same modality (230). Of note, these domain adaptation methods often require the access to unlabeled images in the target domain (e.g., a new scanner, a different hospital), which may not be easy to obtain due to the data privacy and ethics issues. How to collect and share data safely, fairly, and legally across different sites is still an open challenge.

On the other hand, some researchers have started to develop domain generalization algorithms, *without* requiring accessing images from new sites. One stream of works aims to improve the domain generalization ability by extracting domain-independent and robust features or disentangling learned features into domain-specific and domain-invariant components from various seen domains (e.g., multi-center data, multi-modality datasets) to improve the model performance on unseen domains (221, 228, 231). Other researchers have started to adopt data augmentation techniques to simulate various possible data distributions across different domains. For instance, Chen et al. (232) have proposed a data normalization and augmentation pipeline which enables a neural network for cardiac MR image segmentation trained from a single-scanner dataset to generalize well across multi-scanner and multi-site datasets. Zhang et al. (233) applied a similar data augmentation approach to improve the model generalization ability on unseen datasets. Their method has been verified on three tasks including left atrial segmentation from 3D MRI and left ventricle segmentation from 3D ultrasound images.

One bottleneck of augmenting training data for model generalization across different sites is that it is often required to increase the model capacity to compensate for the increased dataset size and variation (232). As a result, training becomes more expensive and challenging. To address this inefficiency problem, active learning (234) has been proposed, which selects the most representative images from a large-scale dataset, reducing labeling workload as well as computational costs. This technique is also related to incremental learning, which aims to improve the model performance by adding new classes incrementally while avoiding a dramatic decrease in overall performance (235). Given the increasing size of the available medical imaging datasets and the practical challenges of collecting, labeling and storing large amounts of images from various sources, it is of great interest to combine domain generalization algorithms with active learning algorithms together to distill a large dataset into a small one but containing the most representative cases for effective and robust learning.

4.3. Lack of Model Interpretability

Unlike symbolic artificial intelligence systems, deep learning systems are difficult to interpret and not transparent. Once a network has been trained, it behaves like a "black box," providing predictions which are not directly interpretable. This issue makes the model unpredictable, intractable for model verification, and ultimately untrustworthy. Recent studies have shown that deep learning-based vision recognition systems can be attacked by images modified with nearly imperceptible perturbations (236–238). These attacks can also happen in medical scenarios, e.g., a DL-based system may make a wrong diagnosis given an image with adversarial noise or even just small rotation, as demonstrated in a very recent paper (239). Although there is no denying that deep learning has become a very powerful tool for image analysis, building resilient algorithms robust to potential attacks remains an unsolved problem. One potential solution, instead of building the resilience into the model, is raising failure awareness of the deployed networks. This can be achieved by providing users with segmentation quality scores (240) or confidence maps, such as uncertainty maps (166) and attention maps (241). These scores or maps can be used as evidence to alert users when failure happens. For example, Sander et al. (166) built a network that is able to simultaneously predict the segmentation mask over cardiac structures and its associated spatial uncertainty map, where the latter one could be used to highlight potential incorrect regions. Such uncertainty information could alert human experts for further justification and refinement in a human-in-the-loop setting.

4.4. Future Work

4.4.1. Smart Imaging

We have shown that deep learning-based methods are able to segment images in real-time with good accuracy. However, these algorithms can still fail on those image acquisitions with low image quality or significant artifacts. Although there have been several algorithms developed to avoid this problem by either checking the image quality before follow-up studies (242, 243), or predicting the segmentation quality to detect failures (240, 244, 245), the development of algorithms that can give instant feedback to correct and optimize the image acquisition process is also important despite less explored. Improving the imaging quality can greatly improve the effectiveness of medical imaging as well as the accuracy of imaging-based diagnosis. For radiologists, however, finding the optimal imaging and reconstruction parameters to scan each patient can take a great amount of time. Therefore, a DL-based system that has the potential of efficiently and effectively improving the image quality with less noise is of great need. Some researchers have utilized learning-based methods (mostly are deep learning-based) for better image

resolution (62), view planning (246), motion correction (247, 248), artifacts reduction (249), shadow detection (250), and noise reduction (251) after image acquisition. However, combining these algorithms with segmentation algorithms and seamlessly integrating them into an efficient, patient-specific imaging system for high-quality image analysis and diagnosis is still an open challenge. An alternative approach is to directly predict cardiac segmentation maps from undersampled k-space data to accelerate the whole procedure, which bypasses the image reconstruction stage (58).

4.4.2. Data Harmonization

A number of works have reported the existence of missing labels and inconsistent labeling protocols among different cardiac image datasets (27, 232). Variations have been found in defining the end of basal slices as well as the endocardial wall of myocardium (some include papillary muscles as part of the endocardial contours whereas others do not). These inconsistencies can be a major obstacle for transferring, evaluating and deploying deep learning models trained from one domain (e.g., hospital) to another. Therefore, building a standard benchmark dataset like CheXpert (252) that (1) is *large* enough to have substantial data diversity that reflects the spectrum of real-world diversity; (2) has a *standard* labeling protocol approved by experts, is indeed a need. However, directly building such a dataset from scratch is time-consuming and expensive. A more promising way might be developing an automated tool to combine existing public datasets from multiple sources and then to harmonize them to a unified, high-quality dataset. This tool can not only open the door for crowd-sourcing but also enable the rapid deployment of those DL-based segmentation models.

4.4.3. Data Privacy

As deep learning is a data-driven approach, an unavoidable and rife concern is about the data privacy. Regulations, such as The General Data Protection Regulation (GDPR) now play an important role to protect users' privacy and have forced organizations to treat data ownership seriously. On the other hand, from a technical point of view, how to store, query, and process data such that there is no privacy concerns for building deep learning systems has now become an even more difficult but interesting challenge. Building a privacy-preserving algorithm requires to combine cryptography and deep learning together and to mix techniques from a wide range of subjects, such as data analysis, distributed computing, federated learning, differential privacy, in order to achieve models with strong security, fast run time, and great generalizability (253–256). In this respect, Papernot (257) published a report for guidance, which summarized a set of best practices for improving the privacy and security of machine learning systems. Yet, this field is still in its infancy.

5. CONCLUSION

In this review paper, we provided a comprehensive overview of these deep learning techniques used in three common imaging modalities (MRI, CT, ultrasound), covering a wide range of existing deep learning approaches (mostly are CNN-based) that are designed for segmenting different cardiac anatomical structures (e.g., cardiac ventricle, atria, vessel). In particular, we presented and discussed recent progress of deep learning-based segmentation methods in the three modalities, outlined future potential and the remaining limitations of these deep learning-based cardiac segmentation methods that may hinder widespread clinical deployment. We hope that this review can provide an intuitive understanding of those deep learning-based techniques that have made a significant contribution to cardiac image segmentation and also increase the awareness of common challenges in this field that call for future contribution.

6. DATA AVAILABILITY STATEMENT

The datasets summarized in **Table 6** can be found in their corresponding websites listed below:

- York: http://www.cse.yorku.ca/~mridataset/
- Sunnybrook: http://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/
- LVSC: http://www.cardiacatlas.org/challenges/lv-segmentation-challenge/
- RVSC: http://www.litislab.fr/?projet=1rvsc
- cDEMRIS: https://www.doc.ic.ac.uk/~rkarim/la_lv_framework/fibrosis
- LVIC: https://www.doc.ic.ac.uk/~rkarim/la_lv_framework/lv_infarct
- LASC'13: www.cardiacatlas.org/challenges/left-atrium-segmentation-challenge/
- HVSMR: http://segchd.csail.mit.edu/
- ACDC: https://acdc.creatis.insa-lyon.fr/
- LASC'18: http://atriaseg2018.cardiacatlas.org/data/
- MM-WHS: http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/mmwhs17/
- CAT08: http://coronary.bigr.nl/centerlines/
- CLS12: http://coronary.bigr.nl/stenoses
- CETUS: https://www.creatis.insa-lyon.fr/Challenge/CETUS
- CAMUS: https://www.creatis.insa-lyon.fr/Challenge/camus.

AUTHOR CONTRIBUTIONS

CC, WB, and DR conceived and designed the work. CC, CQ, and HQ searched and read the MR, CT, Ultrasound literature, respectively, and drafted the manuscript together. WB, DR, GT, and JD provided the critical revision with insightful and constructive comments to improve the manuscript. All authors read and approved the manuscript.

FUNDING

This work was supported by the SmartHeart EPSRC Programme Grant (EP/P001009/1). HQ was supported by the EPSRC Programme Grant (EP/R005982/1).

ACKNOWLEDGMENTS

We would like to thank our colleagues: Karl Hahn, Qingjie Meng, James Batten, and Jonathan Passerat-Palmbach who provided the insight and expertise that greatly assisted the work, and also constructive and thoughtful comments from Turkay Kart that greatly improved the manuscript.

REFERENCES

1. Petitjean C, Zuluaga MA, Bai W, Dacher JN, Grosgeorge D, Caudron J, et al. Right ventricle segmentation from cardiac MRI: a collation study. *Med Image Anal.* (2015) 19:187–202. doi: 10.1016/j.media.2014.10.004
2. Peng P, Lekadir K, Gooya A, Shao L, Petersen SE, Frangi AF. A review of heart chamber segmentation for structural and functional analysis using cardiac magnetic resonance imaging. *Magn Reson Mater Phys Biol Med.* (2016) 29:155–95. doi: 10.1007/s10334-015-0521-4
3. Tavakoli V, Amini AA. A survey of shaped-based registration and segmentation techniques for cardiac images. *Comput Vis Image Understand.* (2013) 117:966–89. doi: 10.1016/j.cviu.2012.11.017
4. Lesage D, Angelini ED, Bloch I, Funka-Lea G. A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes. *Med Image Anal.* (2009) 13:819–45. doi: 10.1016/j.media.2009.07.011
5. Greenspan H, Van Ginneken B, Summers RM. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans Med Imaging.* (2016) 35:1153–9. doi: 10.1109/TMI.2016.2553401
6. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng.* (2017) 19:221–48. doi: 10.1146/annurev-bioeng-071516-044442
7. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* (2017) 42:60–88. doi: 10.1016/j.media.2017.07.005
8. Gandhi S, Mosleh W, Shen J, Chow CM. Automation, machine learning, and artificial intelligence in echocardiography: a brave new world. *Echocardiography.* (2018) 35:1402–18. doi: 10.1111/echo.14086
9. Mazurowski MA, Buda M, Saha A, Bashir MR. Deep learning in radiology: an overview of the concepts and a survey of the state of the art with focus on MRI. *J Magn Reson Imaging.* (2019) 49:939–54. doi: 10.1002/jmri.26534
10. Andreopoulos A, Tsotsos JK. Efficient and generalizable statistical models of shape and appearance for analysis of cardiac MRI. *Med Image Anal.* (2008) 12:335–57. doi: 10.1016/j.media.2007.12.003. Data source: http://www.cse.yorku.ca/~mridataset/
11. Radau P, Lu Y, Connelly K, Paul G, Dick AJ, Wright GA. Evaluation framework for algorithms segmenting short axis cardiac MRI. *MIDAS J.* (2009) 49. Available online at: http://hdl.handle.net/10380/3070. Data source: http://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/
12. Suinesiaputra A, Cowan BR, Al-Agamy AO, Elattar MA, Ayache N, Fahmy AS, et al. A collaborative resource to build consensus for automated left ventricular segmentation of cardiac MR images. *Med Image Anal.* (2014) 18:50–62. doi: 10.1016/j.media.2013.09.001. Data source: http://www.cardiacatlas.org/challenges/lv-segmentation-challenge/
13. Karim R, Housden RJ, Balasubramaniam M, Chen Z, Perry D, Uddin A, et al. Evaluation of current algorithms for segmentation of scar tissue from late gadolinium enhancement cardiovascular magnetic resonance of the left atrium: an open-access grand challenge. *J Cardiovasc Magn Reson.* (2013) 15:105. doi: 10.1186/1532-429X-15-105. Data source: https://www.doc.ic.ac.uk/~rkarim/la_lv_framework/fibrosis
14. Karim R, Bhagirath P, Claus P, James Housden R, Chen Z, Karimaghaloo Z, et al. Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late gadolinium enhancement MR images. *Med Image Anal.* (2016) 30:95–107. doi: 10.1016/j.media.2016.01.004 Data source: https://www.doc.ic.ac.uk/~rkarim/la_lv_framework/lv_infarct/
15. Tobon-Gomez C, Geers AJ, Peters J, Weese J, Pinto K, Karim R, et al. Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. *IEEE Trans Med Imaging.* (2015) 34:1460–73. doi: 10.1109/TMI.2015.2398818. Data source: www.cardiacatlas.org/challenges/left-atrium-segmentation-challenge/
16. Pace DF, Dalca AV, Geva T, Powell AJ, Moghari MH, Golland P. Interactive whole-heart segmentation in congenital heart disease. *Med Image Comput Comput Assist Interv.* (2015) 9351:80–8. doi: 10.1007/978-3-319-24574-4_10. Data source: http://segchd.csail.mit.edu/
17. Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng PA, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Trans Med Imaging.* (2018) 37:2514–25. doi: 10.1109/TMI.2018.2837502 Data source: https://acdc.creatis.insa-lyon.fr/
18. Zhao J, Xiong Z. *2018 Left Atrial Segmentation Challenge Dataset* (2018). available online at: http://atriaseg2018.cardiacatlas.org/. Data source: http://atriaseg2018.cardiacatlas.org/
19. Zhuang X, Li L, Payer C, Stern D, Urschler M, Heinrich MP, et al. Evaluation of algorithms for multi-modality whole heart segmentation: an open-access grand challenge. *Med Image Anal.* (2019) 58:101537. doi: 10.1016/j.media.2019.101537. Data source: http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/mmwhs17/
20. Schaap M, Metz CT, van Walsum T, van der Giessen AG, Weustink AC, Mollet NR, et al. Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. *Med Image Anal.* (2009) 13:701–14. doi: 10.1016/j.media.2009.06.003. Data source: http://coronary.bigr.nl/centerlines/
21. Kirişli HA, Schaap M, Metz CT, Dharampal AS, Meijboom WB, Papadopoulou SL, et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. *Med Image Anal.* (2013) 17:859–76. doi: 10.1016/j.media.2013.05.007. Data source: http://coronary.bigr.nl/stenoses
22. Bernard O, Bosch JG, Heyde B, Alessandrini M, Barbosa D, Camarasu-Pop S, et al. Standardized evaluation system for left ventricular segmentation algorithms in 3D echocardiography. *IEEE Trans Med Imaging.* (2016) 35:967–77. doi: 10.1109/TMI.2015.2503890 Data source: https://www.creatis.insa-lyon.fr/Challenge/CETUS
23. Leclerc S, Smistad E, Pedrosa J, Ostvik A, Cervenansky F, Espinosa F, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. *IEEE Trans Med Imaging.* (2019) 38:2198–210. doi: 10.1109/TMI.2019.2900516
24. Tran PV. A fully convolutional neural network for cardiac segmentation in short-axis MRI. *arxiv* (2016) abs/1604.00494. Available online at: http://arxiv.org/abs/1604.00494 (accessed September 1, 2019).
25. Baumgartner CF, Koch LM, Pollefeys M, Konukoglu E. An exploration of 2D and 3D deep learning techniques for cardiac MR image segmentation. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *International Workshop on Statistical Atlases and Computational Models of the Heart*, Vo. 10663. Cham: Springer (2017). p. 1–8.
26. Isensee F, Jaeger PF, Full PM, Wolf I, Engelhardt S, Maier-Hein KH. Automatic cardiac disease assessment on cine-MRI via time-series segmentation and domain specific features. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges.* Quebec City, QC: Springer International Publishing (2017). p. 120–9.
27. Zheng Q, Delingette H, Duchateau N, Ayache N. 3-D consistent and robust segmentation of cardiac images by deep learning with spatial propagation. *IEEE Trans Med Imaging.* (2018) 37:2137–48. doi: 10.1109/TMI.2018.2820742

28. Qin C, Bai W, Schlemper J, Petersen SE, Piechnik SK, Neubauer S, et al. Joint learning of motion estimation and segmentation for cardiac MR image sequences. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *Proceedings, Part I of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2018*. Granada: Springer International Publishing (2018). p. 472–80.
29. Khened M, Kollerathu VA, Krishnamurthi G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Med Image Anal.* (2019) 51:21–45. doi: 10.1016/j.media.2018.10.004
30. Yu L, Cheng JZ, Dou Q, Yang X, Chen H, Qin J, et al. Automatic 3D cardiovascular MR segmentation with densely-connected volumetric ConvNets. In: Descoteaux M, Maier-Hein L, Franz AM, Jannin P, Collins DL, Duchesne S, editors. *Proceedings, Part II of the 20th International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2017*. Quebec City, QC: Springer International Publishing (2017). p. 287–95.
31. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J Cardiovasc Magn Reson.* (2018) 20:65. doi: 10.1186/s12968-018-0471-x
32. Bai W, Suzuki H, Qin C, Tarroni G, Oktay O, Matthews PM, et al. Recurrent neural networks for aortic image sequence segmentation with sparse annotations. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2018*. Granada: Springer International Publishing (2018). p. 586–94.
33. Duan J, Bello G, Schlemper J, Bai W, Dawes TJW, Biffi C, et al. Automatic 3D bi-ventricular segmentation of cardiac images by a shape-constrained multi-task deep learning approach. *IEEE Transactions on Medical Imaging.* (2019) PP:1. doi: 10.1109/TMI.2019.2894322
34. Chen C, Bai W, Rueckert D. Multi-task learning for left atrial segmentation on GE-MRI. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction With MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Granada: Springer International Publishing (2018). p. 292–301.
35. Yu L, Wang S, Li X, Fu CW, Heng PA. Uncertainty-aware self-ensembling model for semi-supervised 3D left atrium segmentation. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A, editors. *Proceedings, Part II of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2019.* Shenzhen: Springer International Publishing (2019). p. 605–13.
36. Yang X, Bian C, Yu L, Ni D, Heng PA. Hybrid loss guided convolutional networks for whole heart parsing. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer (2017). p. 215–23.
37. Xu Z, Wu Z, Feng J. CFUN: combining faster R-CNN and U-net network for efficient whole heart segmentation. *arxiv* (2018) abs/1812.04914. Available online at: http://arxiv.org/abs/1812.04914 (accessed September 1, 2019).
38. Merkow J, Marsden A, Kriegman D, Tu Z. Dense volume-to-volume vascular boundary detection. In: Ourselin S, Joskowicz L, Sabuncu MR, Ünal GB, Wells W, editors. *Proceedings, Part III of the 19th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2016*. Athens: Springer (2016). p. 371–9.
39. Dou Q, Ouyang C, Chen C, Chen H, Heng PA. Unsupervised cross-modality domain adaptation of ConvNets for biomedical image segmentations with adversarial loss. In: *International Joint Conferences on Artificial Intelligence* (2018). p. 691–7.
40. Dou Q, Ouyang C, Chen C, Chen H, Glocker B, Zhuang X, et al. PnP-AdaNet: plug-and-play adversarial domain adaptation network at unpaired cross-modality cardiac segmentation. *IEEE Access.* (2019) 7:99065–76. doi: 10.1109/ACCESS.2019.2929258
41. Chen C, Dou Q, Zhou J, Qin J, Heng PA. Synergistic image and feature adaptation: towards cross-modality domain adaptation for medical image segmentation. In: *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*. Honolulu, HI: AAAI Press (2019). p. 865–72.
42. Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L, et al. Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy. *Circulation.* (2018) 138:1623–35. doi: 10.1161/CIRCULATIONAHA.118.034338
43. Goodfellow I. *Deep Learning. Adaptive Computation and Machine Learning*. Cambridge, MA; London: The MIT Press (2016).
44. Szegedy C, Liu W, Jia Y, Sermanet P, Reed SE, Anguelov D, et al. Going deeper with convolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*. Boston, MA (2015). p. 1–9.
45. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: *3rd International Conference on Learning Representations, ICLR 2015*. San Diego, CA (2015). p. 14. Available online at: http://arxiv.org/abs/1409.1556
46. Ciresan DC, Giusti A. Deep neural networks segment neuronal membranes in electron microscopy images. In: *Conference on Neural Information Processing Systems* (2012). p. 2852–60. Available online at: http://papers.nips.cc/paper/4741-deep-neural-networks-segment-neuronal-membranes-in-electron-microscopy-images
47. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal.* (2016) 30:108–19. doi: 10.1016/j.media.2016.01.005
48. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Conference on Computer Vision and Pattern Recognition* (2015). p. 3431–40. doi: 10.1109/CVPR.2015.7298965
49. Ronneberger FP Olaf, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells III WM, Frangi AF, editors. *Proceedings, Part III of the 18th International Conference, Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2015*. Munich: Springer International Publishing (2015). p. 234–41.
50. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* (2017) 39:640–51. doi: 10.1109/TPAMI.2016.2572683
51. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: *19th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2016*. Athens (2016). p. 424–32.
52. Milletari F, Navab N, Ahmadi S. V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision, 3DV*. Stanford, CA: IEEE Computer Society (2016). p. 565–71.
53. Tao Q, Yan W, Wang Y, Paiman EHM, Shamonin DP, Garg P, et al. Deep learning–based method for fully automatic quantification of left ventricle function from cine MR images: a multivendor, multicenter study. *Radiology.* (2019) 290:180513. doi: 10.1148/radiol.2018180513
54. Xia Q, Yao Y, Hu Z, Hao A. Automatic 3D atrial segmentation from GE-MRIs using volumetric fully convolutional networks. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges–9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018*. Granada: Springer International Publishing (2018). p. 211–20.
55. Poudel RPK, Lamata P, Montana G. Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. In: *1st International Workshops on Reconstruction and Analysis of Moving Body Organs, RAMBO 2016 and 1st International Workshops on Whole-Heart and Great Vessel Segmentation from 3D Cardiovascular MRI in Congenital Heart Disease, HVSMR 2016* (2016). p. 83–94. Available online at: http://segchd.csail.mit.edu/ (accessed September 1, 2019). Data source: http://segchd.csail.mit.edu/
56. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* (1997) 9:1735–80.

57. Cho K, van Merrienboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In: *Conference on Empirical Methods in Natural Language Processing*. ACL (2014). p. 1724–34. Available online at: https://www.aclweb.org/anthology/D14-1179/
58. Schlemper J, Oktay O, Bai W, Castro DC, Duan J, Qin C, et al. Cardiac MR segmentation from undersampled k-space using deep latent representation learning. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, ediotrs. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Granada: Springer International Publishing (2018). p. 259–67.
59. Oktay O, Ferrante E, Kamnitsas K, Heinrich M, Bai W, Caballero J, et al. Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. *IEEE Trans Med Imaging*. (2018) 37:384–95. doi: 10.1109/TMI.2017.2743464
60. Biffi C, Oktay O, Tarroni G, Bai W, De Marvao A, Doumou G, et al. Learning interpretable anatomical features through deep generative models: application to cardiac remodeling. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention–MICCAI 2018*. Vol. 11071 LNCS. Granada: Springer International Publishing (2018). p. 464–71.
61. Painchaud N, Skandarani Y, Judge T, Bernard O, Lalande A, Jodoin PM. Cardiac MRI segmentation with strong anatomical guarantees. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A, editors. *Proceedings, Part II of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2019*. Shenzhen: Springer International Publishing (2019). p. 632–40.
62. Oktay O, Bai W, Lee M, Guerrero R, Kamnitsas K, Caballero J, et al. Multi-input cardiac image super-resolution using convolutional neural networks. In: Ourselin S, Joskowicz L, Sabuncu MR, Ünal GB, Wells W, editors. *19th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2016.*. Vol. 9902 LNCS. Athens: Springer International Publishing (2016). p. 246–54.
63. Yue Q, Luo X, Ye Q, Xu L, Zhuang X. Cardiac segmentation from LGE MRI using deep neural network incorporating shape and spatial priors. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A, editors. *Medical Image Computing and Computer Assisted Intervention – MICCAI, 2019*. Cham: Springer International Publishing (2019). p. 559–67.
64. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. In: *Conference on Neural Information Processing Systems*. Curran Associates, Inc. (2014). p. 2672–80. Available online at: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf
65. Luc P, Couprie C, Chintala S, Verbeek J. Semantic segmentation using adversarial networks. In: *NIPS Workshop on Adversarial Training* (2016). p. 1–12.
66. Savioli N, Vieira MS, Lamata P, Montana G. A generative adversarial model for right ventricle segmentation. *arxiv* (2018) abs/1810.03969. Available online at: http://arxiv.org/abs/1810.03969 (accessed September 1, 2019).
67. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*. Las Vegas, NV: IEEE Computer Society (2016). p. 2818–26.
68. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, inception-ResNet and the impact of residual connections on learning. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*. San Francisco, CA (2017). p. 4278–84. Available online at: http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14806
69. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. *Conference on Neural Information Processing Systems*. Curran Associates, Inc. (2017). p. 5998–6008.
70. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: learning where to look for the pancreas. In: *Medical Imaging With Deep Learning* (2018). p. 1804.03999.
71. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016*. IEEE Computer Society: Las Vegas, NV (2016). p. 770–8.
72. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. In: *International Conference on Learning Representations* (2016). p. 1–13.
73. Lee CY, Xie S, Gallagher P, Zhang Z, Tu Z. Deeply-supervised nets. In: Lebanon G, Vishwanathan SVN, editors. *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research*, Vol. 38. San Diego, CA: PMLR (2015). p. 562–70. Available online at: http://proceedings.mlr.press/v38/lee15a.html
74. Chen LC, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. *arxiv* (2017) abs/1706.05587. Available online at: http://arxiv.org/abs/1706.05587 (accessed September 1, 2019).
75. He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell*. (2015) 37:1904–16. doi: 10.1109/TPAMI.2015.2389824
76. Jetley S, Lord NA, Lee N, Torr PHS. Learn to pay attention. In: *6th International Conference on Learning Representations, ICLR 2018, Conference Track Proceedings*. Vancouver, BC (2018).
77. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2018*, Salt Lake City, UT: IEEE Computer Society (2018). p. 7132–41.
78. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: *Conference on Computer Vision and Pattern Recognition* (2017). p. 2261–9.
79. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. (1986) 323:533–6.
80. Jang Y, Hong Y, Ha S, Kim S, Chang HJ. Automatic segmentation of LV and RV in cardiac MRI. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *International Workshop on Statistical Atlases and Computational Models of the Heart*. Springer (2017). p. 161–9.
81. Yang X, Bian C, Yu L, Ni D, Heng PA. Class-balanced deep neural network for automatic ventricular structure segmentation. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer International Publishing (2017). p. 152–60.
82. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res*. (2014) 15:1929–58. doi: 10.5555/2627435.2670313
83. Kamnitsas K, Bai W, Ferrante E, Mcdonagh S, Sinclair M. Ensembles of multiple models and architectures for robust brain tumour segmentation. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries–Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017*. Quebec City, QC (2017). p. 450–62.
84. Zheng H, Zhang Y, Yang L, Liang P, Zhao Z, Wang C, et al. A new ensemble learning framework for 3D biomedical image segmentation. In: *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, Honolulu, HI: AAAI Press (2019). p. 5909–16.
85. Chen S, Ma K, Zheng Y. Med3D: transfer learning for 3D medical image analysis. *arxiv* (2019) abs/1904.00625. Available online at: http://arxiv.org/abs/1904.00625 (accessed September 1, 2019).
86. Xue W, Brahm G, Pandey S, Leung S, Li S. Full left ventricle quantification via deep multitask relationships learning. *Med Image Anal*. (2018) 43:54–65. doi: 10.1016/j.media.2017.09.005
87. Zheng Q, Delingette H, Ayache N. Explainable cardiac pathology classification on cine MRI with motion characterization by semi-supervised learning of apparent flow. *Med Image Anal*. (2019) 56:80–95. doi: 10.1016/j.media.2019.06.001
88. Bello GA, Dawes TJW, Duan J, Biffi C, de Marvao A, Howard LSGE, et al. Deep learning cardiac motion analysis for human survival prediction. *Nat Mach Intell*. (2019) 1:95–104. doi: 10.1038/s42256-019-0019-2
89. Van Der Geest RJ, Reiber JH. Quantification in cardiac MRI. *J Magn Reson Imaging*. (1999) 10:602–8.
90. Lieman-Sifry J, Le M, Lau F, Sall S, Golden D. FastVentricle: cardiac segmentation with ENet. In: Pop M, Wright GA, editors. *Functional Imaging*

and Modelling of the Heart. Vol. 10263 LNCS of Lecture Notes in Computer Science. Cham: Springer International Publishing (2017). p. 127–38.

91. Fahmy AS, El-Rewaidy H, Nezafat M, Nakamori S, Nezafat R. Automated analysis of cardiovascular magnetic resonance myocardial native T1 mapping images using fully convolutional neural networks. *J Cardiovasc Magn Reson.* (2019) 21:1–12. doi: 10.1186/s12968-018-0516-1
92. Patravali J, Jain S, Chilamkurthy S. 2D–3D fully convolutional neural networks for cardiac MR segmentation. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer (2017). p. 130–9.
93. Wolterink JM, Leiner T, Viergever MA, Išgum I. Automatic segmentation and disease classification using cardiac cine MR images. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges—8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017*. Vol. 10663 LNCS of Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Quebec City, QC: Springer International Publishing (2017). p. 101–10.
94. Tan LK, Liew YM, Lim E, McLaughlin RA. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. *Med Image Anal.* (2017) 39:78–86. doi: 10.1016/j.media.2017.04.002
95. Vigneault DM, Xie W, Ho CY, Bluemke DA, Noble JA. Ω-Net (Omega-Net): fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks. *Med Image Anal.* (2018) 48:95–106. doi: 10.1016/j.media.2018.05.008
96. Avendi MR, Kheradvar A, Jafarkhani H. Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach. *Magn Reson Med.* (2017) 78:2439–48. doi: 10.1002/mrm.26631
97. Yang H, Sun J, Li H, Wang L, Xu Z. Deep fusion net for multi-atlas segmentation: application to cardiac MR images. In: Ourselin S, Joskowicz L, Sabuncu MR, Ünal GB, Wells W, editors. *19th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2016*. Athens: Springer International Publishing (2016). p. 521–8. doi: 10.1007/978-3-319-46723-8_60
98. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal.* (2017) 35:159–71. doi: 10.1016/j.media.2016.05.009
99. Mortazi A, Karim R, Rhode K, Burt J, Bagci U. CardiacNET: segmentation of left atrium and proximal pulmonary veins from MRI using multi-view CNN. In: Descoteaux M, Maier-Hein L, Franz AM, Jannin P, Collins DL, Duchesne S, editors. *20th International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2017*. Quebec City, QC: Springer (2017). p. 377–85.
100. Xiong Z, Fedorov VV, Fu X, Cheng E, Macleod R, Zhao J. Fully automatic left atrium segmentation from late gadolinium enhanced magnetic resonance imaging using a dual fully convolutional neural network. *IEEE Trans Med Imaging.* (2019) 38:515–24. doi: 10.1109/TMI.2018.2866845
101. Yang G, Zhuang X, Khan H, Haldar S, Nyktari E, Li L, et al. Fully automatic segmentation and objective assessment of atrial scars for long-standing persistent atrial fibrillation patients using late gadolinium-enhanced MRI. *Med Phys.* (2018) 45:1562–76. doi: 10.1002/mp.12832
102. Chen J, Yang G, Gao Z, Ni H, Firmin D, others. Multiview two-task recursive attention model for left atrium and atrial scars segmentation. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*, Vol. 11071. Granada: Springer (2018). p. 455–63.
103. Zabihollahy F, White JA, Ukwatta E. Myocardial scar segmentation from magnetic resonance images using convolutional neural network. In: *Medical Imaging 2018: Computer-Aided Diagnosis*. Vol. 10575. International Society for Optics and Photonics (2018). p. 105752Z.
104. Moccia S, Banali R, Martini C, Muscogiuri G, Pontone G, Pepi M, et al. Development and testing of a deep learning-based strategy for scar segmentation on CMR-LGE images. *Magn Reson Mater Phys Biol Med.* (2019) 32:187–95. doi: 10.1007/s10334-018-0718-4
105. Xu C, Xu L, Gao Z, Zhao S, Zhang H, Zhang Y, et al. Direct delineation of myocardial infarction without contrast agents using a joint motion feature learning architecture. *Med Image Anal.* (2018) 50:82–94. doi: 10.1016/j.media.2018.09.001
106. Li J, Zhang R, Shi L, Wang D. Automatic whole-heart segmentation in congenital heart disease using deeply-supervised 3D FCN. In: Zuluaga MA, Bhatia KK, Kainz B, Moghari MH, Pace DF, editors. *Proceedings of the First International Workshops, RAMBO 2016 and HVSMR 2016, Held in Conjunction with MICCAI 2016, Reconstruction, Segmentation, and Analysis of Medical Images*. Athens: Springer International Publishing (2017). p. 111–8.
107. Wolterink JM, Leiner T, Viergever MA, Išgum I. Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease. In: Zuluaga MA, Bhatia KK, Kainz B, Moghari MH, Pace DF, editors. *Proceedings of the First International Workshops, RAMBO 2016 and HVSMR 2016, Held in Conjunction with MICCAI 2016, Reconstruction, Segmentation, and Analysis of Medical Images*. Athens: Springer International Publishing (2017). p. 95–102. doi: 10.1007/978-3-319-52280-7
108. Li C, Tong Q, Liao X, Si W, Chen S, Wang Q, et al. APCP-NET: aggregated parallel cross-scale pyramid network for CMR segmentation. In: *16th IEEE International Symposium on Biomedical Imaging, ISBI*, Venice (2019). p. 784–8.
109. Zotti C, Luo Z, Lalande A, Jodoin PM. Convolutional neural Network with shape prior applied to cardiac MRI segmentation. *IEEE J Biomed Health Inform.* (2019) 23:1119–28. doi: 10.1109/JBHI.2018.2865450
110. Zotti C, Luo Z, Lalande A, Humbert O, Jodoin PM. GridNet with automatic shape prior registration for automatic MRI cardiac segmentation. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Vol. 10663. Quebec City, QC: Springer (2017). p. 73–81.
111. Rohé MM, Sermesant M, Pennec X. Automatic multi-atlas segmentation of myocardium with SVF-Net. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017 Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*, Vol. 10663. Quebec City, QC: Springer International Publishing (2017). p. 170–7.
112. Tziritas G, Grinias E. Fast fully-automatic localization of left ventricle and myocardium in MRI using MRF model optimization, substructures tracking and B-spline smoothing. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer International Publishing (2017). p. 91–100.
113. Zreik M, Leiner T, De Vos BD, Van Hamersvelt RW, Viergever MA, Isgum I. Automatic segmentation of the left ventricle in cardiac CT angiography using convolutional neural networks. In: *International Symposium on Biomedical Imaging* (2016). p. 40–3.
114. Payer C, Štern D, Bischof H, Urschler M. Multi-label whole heart segmentation using CNNs and anatomical label configurations. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer International Publishing (2018). p. 190–8.
115. Tong Q, Ning M, Si W, Liao X, Qin J. 3D deeply-supervised U-net based whole heart segmentation. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer (2017). p. 224–32.

116. Wang C, MacGillivray T, Macnaught G, Yang G, Newby D. A two-stage 3D Unet framework for multi-class segmentation on full resolution image. *arxiv* (2018) 1–10. Available online at: http://arxiv.org/abs/1804.04341 (accessed September 1, 2019).
117. Wang C, Smedby Ö. Automatic whole heart segmentation using deep learning and shape context. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges—8th International Workshop, STACOM 2017, Held in Conjunction With MICCAI 2017*. Quebec City, QC: Springer (2017). p. 242–9.
118. Mortazi A, Burt J, Bagci U. Multi-planar deep segmentation networks for cardiac substructures from MRI and CT. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer International Publishing (2017). p. 199–206.
119. Ye C, Wang W, Zhang S, Wang K. Multi-depth fusion network for whole-heart CT image segmentation. *IEEE Access.* (2019) 7:23421–9. doi: 10.1109/ACCESS.2019.2899635
120. Zreik M, Lessmann N, van Hamersvelt RW, Wolterink JM, Voskuil M, Viergever MA, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.* (2018) 44:72–85. doi: 10.1016/j.media.2017.11.008
121. Joyce T, Chartsias A, Tsaftaris SA. Deep multi-class segmentation without ground-truth labels. In: *Medical Imaging With Deep Learning* (2018). p. 1–9.
122. Moeskops P, Wolterink JM, van der Velden BH, Gilhuijs KG, Leiner T, Viergever MA, et al. Deep learning for multi-task medical image segmentation in multiple modalities. In: Ourselin S, Joskowicz L, Sabuncu MR, Ünal GB, Wells W, editors. *19th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2016*. Athens: Springer (2016). p. 478–86.
123. Lee MCH, Petersen K, Pawlowski N, Glocker B, Schaap M. TETRIS: template transformer networks for image segmentation with shape priors. *IEEE Trans Med Imaging.* (2019) 38:2596–606. doi: 10.1109/TMI.2019.2905990
124. Gülsün MA, Funka-Lea G, Sharma P, Rapaka S, Zheng Y. Coronary centerline extraction via optimal flow paths and CNN path pruning. In: Ourselin S, Joskowicz L, Sabuncu MR, Ünal GB, Wells W, editors. *19th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI, 2016*. Athens: Springer (2016). p. 317–25.
125. Guo Z, Bai J, Lu Y, Wang X, Cao K, Song Q, et al. DeepCenterline: a multi-task fully convolutional network for centerline extraction. In: Chung ACS, Gee JC, Yushkevich PA, Bao S, editors. *Proceedings of the 26th International Conference, Information Processing in Medical Imaging—IPMI, 2019*. Hong Kong: Springer (2019). p. 441–53.
126. Shen Y, Fang Z, Gao Y, Xiong N, Zhong C, Tang X. Coronary arteries segmentation based on 3D FCN with attention gate and level set function. *IEEE Access.* (2019) 7:42826–35. doi: 10.1109/ACCESS.2019.2908039
127. Wolterink JM, van Hamersvelt RW, Viergever MA, Leiner T, Išgum I. Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier. *Med Image Anal.* (2019) 51:46–60. doi: 10.1016/j.media.2018.10.005
128. Wolterink JM, Leiner T, Išgum I. Graph convolutional networks for coronary artery segmentation in cardiac CT angiography. *arxiv* (2019) abs/1908.05343. Available online at: http://arxiv.org/abs/1908.05343 (accessed September 1, 2019).
129. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal.* (2016) 34:123–36. doi: 10.1016/j.media.2016.04.004
130. Lessmann N, Išgum I, Setio AA, de Vos BD, Ciompi F, de Jong PA, et al. Deep convolutional neural networks for automatic coronary calcium scoring in a screening study with low-dose chest CT. In: Tourassi GD, Armato SG III, editors. *Medical Imaging 2016: Computer-Aided Diagnosis*. Vol. 9785. San Diego, CA: International Society for Optics and Photonics; SPIE (2016). p. 978511.
131. Lessmann N, van Ginneken B, Zreik M, de Jong PA, de Vos BD, Viergever MA, et al. Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions. *IEEE Trans Med Imaging.* (2017) 37:615–25. doi: 10.1109/TMI.2017.2769839
132. Liu J, Jin C, Feng J, Du Y, Lu J, Zhou J. A vessel-focused 3D convolutional network for automatic segmentation and classification of coronary artery plaques in cardiac CTA. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Granada: Springer (2018). p. 131–41.
133. Santini G, Della Latta D, Martini N, Valvano G, Gori A, Ripoli A, et al. An automatic deep learning approach for coronary artery calcium segmentation. In: *International Federation for Medical and Biological Engineering*. Vol. 65 (2017). p. 374–7.
134. Shadmi R, Mazo V, Bregman-Amitai O, Elnekave E. Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest CT. In: *15th IEEE International Symposium on Biomedical Imaging, ISBI, 2018*. Washington, DC: IEEE (2018). p. 24–8.
135. Zhang W, Zhang J, Du X, Zhang Y, Li S. An end-to-end joint learning framework of artery-specific coronary calcium scoring in non-contrast cardiac CT. *Computing.* (2019) 101:667–78. doi: 10.1007/s00607-018-0678-6
136. Ma J, Zhang R. Automatic calcium scoring in cardiac and chest CT using DenseRAUnet. *arxiv* (2019) abs/1907.11392. Available online at: http://arxiv.org/abs/1907.11392 (accessed September 1, 2019).
137. Yang X, Bian C, Yu L, Ni D, Heng PA. 3D convolutional networks for fully automatic fine-grained whole heart partition. In: Pop M, Sermesant M, Jodoin P-M, Lalande A, Zhuang X, Yang G, Young AA, Bernard O, editors. *Proceedings of the 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges*. Quebec City, QC: Springer (2017). p. 181–9.
138. Carneiro G, Nascimento J, Freitas A. Robust left ventricle segmentation from ultrasound data using deep neural networks and efficient search methods. In: *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2010). p. 1085–8.
139. Carneiro G, Nascimento JC, Freitas A. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Trans Image Process.* (2012) 21:968–82. doi: 10.1109/TIP.2011.2169273
140. Nascimento JC, Carneiro G. Deep learning on sparse manifolds for faster object segmentation. *IEEE Trans Image Process.* (2017) 26:4978–90. doi: 10.1109/TIP.2017.2725582
141. Nascimento JC, Carneiro G. Non-rigid segmentation using sparse low dimensional manifolds and deep belief networks. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition—CVPR, 2014*. Columbus, OH: IEEE Computer Society (2014). p. 288–95. doi: 10.1109/CVPR.2014.44
142. Nascimento JC, Carneiro G. One shot segmentation: unifying rigid detection and non-rigid segmentation using elastic regularization. *IEEE Trans Pattern Anal Mach Intell.* (2019). doi: 10.1109/TPAMI.2019.2922959
143. Veni G, Moradi M, Bulu H, Narayan G, Syeda-Mahmood T. Echocardiography segmentation based on a shape-guided deformable model driven by a fully convolutional network prior. In: *15th IEEE International Symposium on Biomedical Imaging—ISBI, 2018*. Washington, DC: IEEE (2018). p. 898–902. doi: 10.1109/ISBI.2018.8363716
144. Carneiro G, Nascimento JC. Multiple dynamic models for tracking the left ventricle of the heart from ultrasound data using particle filters and deep learning architectures. In: *Conference on Computer Vision and Pattern Recognition*. IEEE (2010). p. 2815–22. doi: 10.1109/CVPR.2010.5540013
145. Carneiro G, Nascimento JC. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. *IEEE Trans Pattern Anal Mach Intell.* (2013) 35:2592–607. doi: 10.1109/TPAMI.2013.96
146. Jafari MH, Girgis H, Liao Z, Behnami D, Abdi A, Vaseli H, et al. A unified framework integrating recurrent fully-convolutional networks and optical flow for segmentation of the left ventricle in echocardiography data. In: *Deep*

Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Springer International Publishing (2018). p. 29–37.

147. Carneiro G, Nascimento JC. Incremental on-line semi-supervised learning for segmenting the left ventricle of the heart from ultrasound data. In: *2011 International Conference on Computer Vision*. Barcelona: IEEE (2011). p. 1700–7. doi: 10.1109/ICCV.2011.6126433
148. Carneiro G, Nascimento JC. The use of on-line co-training to reduce the training set size in pattern recognition methods: application to left ventricle segmentation in ultrasound. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. Providence, RI: IEEE (2012). p. 948–55.
149. Smistad E, Ostvik A, Haugen BO, Lovstakken L. 2D left ventricle segmentation using deep learning. In: *2017 IEEE International Ultrasonics Symposium (IUS)*. IEEE (2017). p. 1–4.
150. Yu L, Guo Y, Wang Y, Yu J, Chen P. Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks. *IEEE Trans Biomed Eng.* (2017) 64:1886–95. doi: 10.1109/TBME.2016.2628401
151. Jafari MH, Girgis H, Abdi AH, Liao Z, Pesteie M, Rohling R, et al. Semi-supervised learning for cardiac left ventricle segmentation using conditional deep generative models as prior. In: *16th IEEE International Symposium on Biomedical Imaging ISBI 2019*. Venice: IEEE (2019). p. 649–52.
152. Girdhar R, Fouhey DF, Rodriguez M, Gupta A. Learning a predictable and generative vector representation for objects. In: Leibe B, Matas J, Sebe N, Welling M, editors. *Proceedings, Part VI of the 14th European Conference, Computer Vision—ECCV 2016*. Amsterdam: Springer International Publishing (2016). p. 484–99.
153. Chen H, Zheng Y, Park JH, Heng PA, Zhou SK. Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images. In: Ourselin S, Leo Joskowicz, Mert R. Sabuncu, Ünal GB, William Wells, editors. *Proceedings, Part II of the 19th International Conference, Medical Image Computing and Computer-Assisted Intervention—MICCAI 2016*. Athens: Springer International Publishing (2016). p. 487–95.
154. Smistad E, Østvik A, Mjal Salte I, Leclerc S, Bernard O, Lovstakken L. Fully automatic real-time ejection fraction and MAPSE measurements in 2D echocardiography using deep neural networks. In: *2018 IEEE International Ultrasonics Symposium (IUS)*. IEEE (2018). p. 1–4.
155. Leclerc S, Smistad E, Grenier T, Lartizien C, Ostvik A, Espinosa F, et al. Deep learning applied to multi-structure segmentation in 2D echocardiography: a preliminary investigation of the required database size. In: *2018 IEEE International Ultrasonics Symposium (IUS)*. IEEE (2018). p. 1–4.
156. Jafari MH, Girgis H, Van Woudenberg N, Liao Z, Rohling R, Gin K, et al. Automatic biplane left ventricular ejection fraction estimation with mobile point-of-care ultrasound using multi-task learning and adversarial training. *Int J Comput Assist Radiol Surg.* (2019) 14:1027–37. doi: 10.1007/s11548-019-01954-w
157. Dong S, Luo G, Wang K, Cao S, Li Q, Zhang H. A combined fully convolutional networks and deformable model for automatic left ventricle segmentation based on 3D echocardiography. *Biomed Res Int.* (2018) 2018:5682365. doi: 10.1155/2018/5682365
158. Dong S, Luo G, Wang K, Cao S, Mercado A, Shmuilovich O, et al. VoxelAtlasGAN: 3D left ventricle segmentation on echocardiography with atlas guided generation and voxel-to-voxel discrimination. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Granada: Springer International Publishing (2018). p. 622–9.
159. Ghesu FC, Krubasik E, Georgescu B, Singh V, Yefeng Zheng, Hornegger J, et al. Marginal space deep learning: efficient architecture for volumetric image parsing. *IEEE Trans Med Imaging.* (2016) 35:1217–28. doi: 10.1109/TMI.2016.2538802
160. Degel MA, Navab N, Albarqouni S. Domain and geometry agnostic CNNs for left atrium segmentation in 3D ultrasound. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Springer International Publishing (2018). p. 630–7.
161. Karim R, Blake LE, Inoue J, Tao Q, Jia S, James Housden R, et al. Algorithms for left atrial wall segmentation and thickness–evaluation on an open-source CT and MRI image database. *Med Image Anal.* (2018) 50:36–53. doi: 10.1016/j.media.2018.08.004
162. Li J, Yu Z, Gu Z, Liu H, Li Y. Dilated-inception net: multi-scale feature aggregation for cardiac right ventricle segmentation. *IEEE Trans Biomed Eng.* (2019) 66:3499–508. doi: 10.1109/TBME.2019.2906667
163. Zhou XY, Yang GZ. Normalization in training U-Net for 2D biomedical semantic segmentation. *IEEE Robot Autom Lett.* (2019) 4:1792–9. doi: 10.1109/LRA.2019.2896518
164. Zhang J, Du J, Liu H, Hou X, Zhao Y, Ding M. LU-NET: an Improved U-Net for ventricular segmentation. *IEEE Access.* (2019) 7:92539–46. doi: 10.1109/ACCESS.2019.2925060
165. Cong C, Zhang H. Invert-U-Net DNN segmentation model for MRI cardiac left ventricle segmentation. *J Eng.* (2018) 2018:1463–7. doi: 10.1049/joe.2018.8302
166. Sander J, de Vos BD, Wolterink JM, Išgum I. Towards increased trustworthiness of deep learning segmentation methods on cardiac MRI. In: *Medical Imaging 2019: Image Processing*. Vol. 10949. International Society for Optics and Photonics (2019). p. 1094919.
167. Chen M, Fang L, Liu H. FR-NET: focal Loss constrained deep residual networks for segmentation of cardiac MRI. In: *16th IEEE International Symposium on Biomedical Imaging, ISBI, 2019*. Venice: IEEE (2019). p. 764–7.
168. Chen C, Biffi C, Tarroni G, Petersen S, Bai W, Rueckert D. Learning shape priors for robust cardiac MR segmentation from multi-view images. In: *Medical Image Computing and Computer Assisted Intervention* (2019). p. 523–31.
169. Du X, Yin S, Tang R, Zhang Y, Li S. Cardiac-DeepIED: automatic pixel-level deep segmentation for cardiac bi-ventricle using improved end-to-end encoder-decoder network. *IEEE J Transl Eng Health Med.* (2019) 7:1–10. doi: 10.1109/JTEHM.2019.2900628
170. Yan W, Wang Y, Li Z, van der Geest RJ, Tao Q. Left ventricle segmentation via optical-flow-net from short-axis cine MRI: preserving the temporal coherence of cardiac motion. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2018*. Vol. 11073 LNCS. Granada: Springer International Publishing (2018). p. 613–21.
171. Savioli N, Vieira MS, Lamata P, Montana G. Automated segmentation on the entire cardiac cycle using a deep learning work—flow. In: *2018 Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS)* (2018). p. 153–8.
172. Clough JR, Oksuz I, Byrne N, Schnabel JA, King AP. Explicit topological priors for deep-learning based image segmentation using persistent homology. In: Chung ACS, Gee JC, Yushkevich PA, Bao S, editors. *Proceedings of the 26th International Conference,IPMI 2019, Information Processing in Medical Imaging*. Vol. 11492 LNCS. Hong Kong (2019). p. 16–28.
173. Chen X, Williams BM, Vallabhaneni SR, Czanner G, Williams R, Zheng Y. Learning active contour models for medical image segmentation. In: *Conference on Computer Vision and Pattern Recognition* (2019). p. 11632–40.
174. Qin C, Bai W, Schlemper J, Petersen SE, Piechnik SK, Neubauer S, et al. Joint motion estimation and segmentation from undersampled cardiac MR image. In: Knoll F, Maier AK, Rueckert D, editors. *Machine Learning for Medical Image Reconstruction - First International Workshop, MLMIR 2018, Held in Conjunction with MICCAI 2018*. Granada: Springer International Publishing (2018). p. 55–63.
175. Dangi S, Yaniv Z, Linte CA. Left ventricle segmentation and quantification from cardiac cine MR images via multi-task learning. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges—9th International Workshop, STACOM 2018, Held in Conjunction With MICCAI 2018*. Granada: Springer (2018). p. 21–31.
176. Zhang L, Karanikolas GV, Akçakaya M, Giannakis GB. Fully automatic segmentation of the right ventricle via multi-task deep neural networks. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2018*. Calgary, AB: IEEE (2018). p. 6677–81.

177. Chartsias A, Joyce T, Papanastasiou G, Semple S, Williams M, Newby D, et al. Factorised spatial representation learning: application in semi-supervised myocardial segmentation. In: *Proceedings, Part II of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2018*. Vol. 11071 LNCS. Granada: Springer International Publishing (2018). p. 490–8.
178. Chartsias A, Joyce T, Papanastasiou G, Semple S, Williams M, Newby DE, et al. Disentangled representation learning in cardiac image analysis. *Med Image Anal.* (2019) 58:101535. doi: 10.1016/j.media.2019.101535
179. Huang Q, Yang D, Yi J, Axel L, Metaxas D. FR-Net: joint reconstruction and segmentation in compressed sensing cardiac MRI. In: Coudière Y, Ozenne V, Vigmond EJ, Zemzemi N, editors. *10th International Conference on Functional Imaging and Modeling of the Heart—FIMH 2019*. Bordeaux: Springer International Publishing (2019). p. 352–60.
180. Liao F, Chen X, Hu X, Song S. Estimation of the volume of the left ventricle from MRI images using deep neural networks. *IEEE Trans Cybernet.* (2019) 49:495–504. doi: 10.1109/TCYB.2017.2778799
181. Duan J, Schlemper J, Bai W, Dawes TJW, Bello G, Doumou G, et al. Deep nested level sets: fully automated segmentation of cardiac MR images in patients with pulmonary hypertension. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2018*. Granada: Springer International Publishing (2018). p. 595–603.
182. Medley DO, Santiago C, Nascimento JC. Segmenting the left ventricle in cardiac in cardiac MRI: from handcrafted to deep region based descriptors. In: *16th IEEE International Symposium on Biomedical Imaging ISBI, 2019*. Venice: IEEE (2019). p. 644–8.
183. Lu X, Chen X, Li W, Qiao Y. Graph cut segmentation of the right ventricle in cardiac MRI using multi-scale feature learning. In: Wang Y, Chang C-C, editors. *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy—ICCSP 2019*. Kuala Lumpur: ACM (2019). p. 231–5.
184. Karim R, Mohiaddin R, Rueckert D. Left atrium segmentation for atrial fibrillation ablation. In: *Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling*. San Diego, CA: SPIE (2008). p. 69182U.
185. Tao Q, Shahzad R, Ipek EG, Berendsen FF, Nazarian S, van der Geest RJ. Fully automatic segmentation of left atrium and pulmonary veins in late gadolinium-enhanced MRI: towards objective atrial scar assessment. *J Magn Reson Imaging.* (2016) 44:346–54. doi: 10.1002/jmri.25148
186. Zhuang X, Rhode KS, Razavi RS, Hawkes DJ, Ourselin S. A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *IEEE Trans Medical Imaging.* (2010) 29:1612–25. doi: 10.1109/TMI.2010.2047112
187. Preetha CJ, Haridasan S, Abdi V, Engelhardt S. Segmentation of the left atrium from 3D gadolinium-enhanced MR images with convolutional neural networks. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Granada: Springer International Publishing (2018). p. 265–72.
188. Bian C, Yang X, Ma J, Zheng S, Liu YA, Nezafat R, et al. Pyramid network with online hard example mining for accurate left atrium segmentation. In: *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges 2018*. Granada: Springer International Publishing (2018). p. 237–45.
189. Savioli N, Montana G, Lamata P. V-FCNN: volumetric fully convolution neural network for automatic atrial segmentation. In: *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Granada: Springer International Publishing (2018). p. 273–81.
190. Jia S, Despinasse A, Wang Z, Delingette H, Pennec X, Jaïs P, et al. Automatically segmenting the left atrium from cardiac images using successive 3D U-Nets and a contour loss. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Granada: Springer (2018). p. 221–9.
191. Vesal S, Ravikumar N, Maier A. Dilated convolutions in neural networks for left atrial segmentation in 3D gadolinium enhanced-MRI. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, ediotrs. *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges—9th International Workshop, STACOM 2018, Held in Conjunction With MICCAI 2018*. Granada: Springer (2018). p. 319–28.
192. Li C, Tong Q, Liao X, Si W, Sun Y, Wang Q, et al. Attention based hierarchical aggregation network for 3D left atrial segmentation: 9th international workshop, STACOM 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Revised Selected Papers. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young A, et al., editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges. Vol. 11395 of Lecture Notes in Computer Science*. Cham; Granada: Springer International Publishing (2018). p. 255–64.
193. Yang G, Chen J, Gao Z, Zhang H, Ni H, Angelini E, et al. Multiview sequential learning and dilated residual learning for a fully automatic delineation of the left atrium and pulmonary veins from late gadolinium-enhanced cardiac MRI images. In: *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2018*. Honolulu, HI: IEEE (2018) 2018:1123–7. doi: 10.1109/EMBC.2018.8512550
194. Kim RJ, Fieno DS, Parrish TB, Harris K, Chen EL, Simonetti O, et al. Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function. *Circulation.* (1999) 100:1992–2002.
195. Carminati MC, Boniotti C, Fusini L, Andreini D, Pontone G, Pepi M, et al. Comparison of image processing techniques for nonviable tissue quantification in late gadolinium enhancement cardiac magnetic resonance images. *J Thorac Imaging.* (2016) 31:168–76. doi: 10.1097/RTI.0000000000000206
196. Yang G, Zhuang X, Khan H, Haldar S, Nyktari E, Ye X, et al. Segmenting atrial fibrosis from late gadolinium-enhanced cardiac MRI by deep-learned features with stacked sparse auto-encoders. In: Hernández MCV, González-Castro V, editors. *Proceedingsof the 21st Annual Conference on Medical Image Understanding and Analysis—MIUA*. Edinburgh: Springer International Publishing (2017). p. 195–206.
197. Fahmy AS, Rausch J, Neisius U, Chan RH, Maron MS, Appelbaum E, et al. Automated cardiac MR scar quantification in hypertrophic cardiomyopathy using deep convolutional neural networks. *JACC Cardiovasc Imaging.* (2018) 11:1917–8. doi: 10.1016/j.jcmg.2018.04.030
198. Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging.* (1999) 18:712–21.
199. Herment A, Kachenoura N, Lefort M, Bensalah M, Dogui A, Frouin F, et al. Automated segmentation of the aorta from phase contrast MR images: validation against expert tracing in healthy volunteers and in patients with a dilated aorta. *J Magn Reson Imaging.* (2010) 31:881–8. doi: 10.1002/jmri.22124
200. Shi Z, Zeng G, Zhang L, Zhuang X, Li L, Yang G, et al. Bayesian VoxDRN: a probabilistic deep voxelwise dilated residual network for whole heart segmentation from 3D MR images. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2018*. Vol. 49. Granada: Springer International Publishing (2018). p. 569–77. Available online at: http://hdl.handle.net/10380/3070 (accessed September 1, 2019).
201. Kang DW, Woo J, Slomka PJ, Dey D, Germano G, Kuo J. Heart chambers and whole heart segmentation techniques: review. *J Electron Imaging.* (2012) 21:010901. doi: 10.1117/1.JEI.21.1.010901
202. Dormer JD, Ma L, Halicek M, Reilly CM, Schreibmann E, Fei B. Heart chamber segmentation from CT using convolutional neural networks. In: Fei B, Webster RJ, editors. *Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging*. Houston, TX: SPIE (2018). p. 105782S.

203. de Vos BD, Wolterink JM, De Jong PA, Leiner T, Viergever MA, Isgum I. ConvNet-based localization of anatomical structures in 3-D medical images. *IEEE Trans Med Imaging.* (2017) 36:1470–81. doi: 10.1109/TMI.2017.2673121
204. Zhang DP. *Coronary Artery Segmentation and Motion Modelling*. London: Imperial College (2010).
205. Huang W, Huang L, Lin Z, Huang S, Chi Y, Zhou J, et al. Coronary artery segmentation by deep learning neural networks on computed tomographic coronary angiographic images. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC, 2018*. Honolulu, HI: IEEE (2018). p. 608–11.
206. Chen YC, Lin YC, Wang CP, Lee CY, Lee WJ, Wang TD, et al. Coronary artery segmentation in cardiac CT angiography using 3D multi-channel U-net. In: Jorge Cardoso M, Feragen A, Glocker B, Konukoglu E, Oguz I, Unal GB, Vercauteren T, editors. *International Conference on Medical Imaging with Deep Learning, MIDL, 2019*. London (2019). p. 1907.12246.
207. Duan Y, Feng J, Lu J, Zhou J. Context aware 3D fully convolutional networks for coronary artery segmentation. In: Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, Rhode KS, Mansi T, editors. *Proceedings of the 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges*. Granada: Springer (2018). p. 85–93.
208. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol.* (1990) 15:827–32.
209. de Vos BD, Wolterink JM, Leiner T, de Jong PA, Lessmann N, Isgum I. Direct automatic coronary calcium scoring in cardiac and chest CT. *IEEE Trans Med Imaging.* (2019) 38:2127–38. doi: 10.1109/TMI.2019.2899534
210. Cano-Espinosa C, González G, Washko GR, Cazorla M, Estépar RSJ. Automated Agatston score computation in non-ECG gated CT scans using deep learning. In: Angelini ED, Landman BA, editors. *Medical Imaging 2018: Image Processing*. Vol. 10574. Houston, TX: International Society for Optics and Photonics; SPIE (2018). p. 105742K.
211. Zreik M, van Hamersvelt RW, Wolterink JM, Leiner T, Viergever MA, Išgum I. A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. *IEEE Trans Med Imaging.* (2018) 38:1588–98. doi: 10.1109/TMI.2018.2883807
212. Noble JA, Boukerroui D. Ultrasound image segmentation: a survey. *IEEE Trans Med Imaging.* (2006) 25:987–1010. doi: 10.1109/TMI.2006.877092
213. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science.* (2006) 313:504–7. doi: 10.1126/science.1127647
214. Smistad E, Lindseth F. Real-time tracking of the left ventricle in 3D ultrasound using Kalman filter and mean value coordinates. In: *Medical Image Segmentation for Improved Surgical Navigation* (2014). p. 189.
215. Georgescu B, Zhou XS, Comaniciu D, Gupta A. Database-guided segmentation of anatomical structures with complex appearance. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 2. San Diego, CA: IEEE (2005). p. 429–36.
216. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. *Int J Comput Vis.* (1988) 1:321–31. doi: 10.1007/BF00133570
217. Kamnitsas K, Baumgartner C, Ledig C, Newcombe V, Simpson J, Kane A, et al. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In: Niethammer M, Styner M, Aylward SR, Zhu H, Oguz I, Yap P-T, Shen D, editors. *Proceedings of the 25th International Conference on Information Processing in Medical Imaging—IPMI, 2017*. Boone, NC: Springer International Publishing (2017). p. 597–609. doi: 10.1007/978-3-319-59050-9_47
218. Zheng Y, Barbu A, Georgescu B, Scheuering M, Comaniciu D. Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. *IEEE Trans Med Imaging.* (2008) 27:1668–81. doi: 10.1109/TMI.2008.2004421
219. Kingma DP, Welling M. Auto-encoding variational bayes. In: Bengio Y, LeCun Y, editors. *2nd International Conference on Learning Representations, ICLR 2014*. Banff, AB (2013). p. 1–14.
220. Cubuk ED, Zoph B, Mane D, Vasudevan V, Le QV. AutoAugment: learning augmentation policies from data. In: *Conference on Computer Vision and Pattern Recognition* (2019). p. 113–23.
221. Volpi R, Namkoong H, Sener O, Duchi JC, Murino V, Savarese S. Generalizing to unseen domains via adversarial data augmentation. In: Bengio S, Wallach HM, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R, editors. *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018*. Montréal, QC (2018). p. 5339–49.
222. Zhao A, Balakrishnan G, Durand F, Guttag JV, Dalca AV. Data augmentation using learned transformations for one-shot medical image segmentation. In: *Conference on Computer Vision and Pattern Recognition* (2019). p. 8543–53.
223. Chaitanya K, Karani N, Baumgartner C, Donati O, Becker A, Konukoglu E. Semi-supervised and task-driven data augmentation. In: Chung ACS, Gee JC, Yushkevich PA, Bao S, editors. *Proceedings of the 26th International Conference—IPMI, 2019, Information Processing in Medical Imaging*. Hong Kong (2019). p. 29–41.
224. Bai W, Oktay O, Sinclair M, Suzuki H, Rajchl M, Tarroni G, et al. Semi-supervised learning for network-based cardiac MR image segmentation. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S, editors. *Proceedings, Part II of the 20th International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2017*. Vol. 10434. Quebec City, QC: Springer International Publishing (2017). p. 253–60.
225. Can YB, Chaitanya K, Mustafa B, Koch LM, Konukoglu E, Baumgartner CF. Learning to segment medical images with scribble-supervision alone. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer International Publishing (2018). p. 236–44.
226. Kervadec H, Dolz J, Tang M, Granger E, Boykov Y, Ben Ayed I. Constrained-CNN losses for weakly supervised segmentation. *Med Image Anal.* (2019) 54:88–99. doi: 10.1016/j.media.2019.02.009
227. Bai W, Chen C, Tarroni G, Duan J, Guitton F, Petersen SE, et al. Self-supervised learning for cardiac MR image segmentation by anatomical position prediction. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A, editors. *Proceedings, Part II of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI, 2019*. Shenzhen: Springer International Publishing (2019). p. 541–9.
228. Dou Q, Liu Q, Heng PA, Glocker B. Unpaired multi-modal segmentation via knowledge distillation. *arxiv* (2020) abs/2001.0311. Available online at: https://arxiv.org/abs/2001.03111
229. Ouyang C, Kamnitsas K, Biffi C, Duan J, Rueckert D. Data efficient unsupervised domain adaptation for cross-modality image segmentation. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P-T, Khan A, editors. *22nd International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*. Shenzhen: Springer International Publishing (2019). p. 669–77.
230. Chen J, Zhang H, Zhang Y, Zhao S, Mohiaddin R, Wong T, et al. Discriminative consistent domain generation for semi-supervised learning. In: Shen D, Liu T, Peters TM, Staib LH, Essert C, Zhou S, Yap P, Khan A, editors. *22nd International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2019, Vol. 11765*. Shenzhen: Springer (2019). p. 595–604.
231. Dou Q, Castro DC, Kamnitsas K, Glocker B. Domain generalization via model-agnostic learning of semantic features. In: *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019*. Vancouver, BC (2019). p. 6447–58. Available online at: http://arxiv.org/abs/1910.13580
232. Chen C, Bai W, Davies RH, Bhuva AN, Manisty C, Moon JC, et al. Improving the generalizability of convolutional neural network-based segmentation on CMR images. *arxiv* (2019) abs/1907.01268. Available online at: http://arxiv.org/abs/1907.01268 (accessed September 1, 2019).
233. Zhang L, Wang X, Yang D, Sanford T, Harmon S, Turkbey B, et al. When unseen domain generalization is unnecessary? Rethinking data augmentation. *arxiv* (2019) abs/1906.03347. Available online at: http://arxiv.org/abs/1906.03347 (accessed September 1, 2019).
234. Mahapatra D, Bozorgtabar B, Thiran JP, Reyes M. Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st*

International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2018. Granada: Springer International Publishing (2018). p. 580–8.
235. Castro FM, Marín-Jiménez MJ, Guil N, Schmid C, Alahari K. End-to-end incremental learning. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, editors. *Computer Vision—ECCV 2018 - 15th European Conference*. Munich: Springer (2018). p. 241–57.
236. Szegedy C, Zaremba W, Sutskever I, Bruna J, Erhan D, Goodfellow I, et al. Intriguing properties of neural networks. In: *2nd International Conference on Learning Representations, ICLR 2014, Conference Track Proceedings*. Banff, AB (2014). p. 1–10.
237. Kurakin A, Goodfellow I, Bengio S. Adversarial examples in the physical world. In: *5th International Conference on Learning Representations, ICLR 2017, Workshop Track Proceedings*. Toulon (2017). p. 1–14.
238. Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. In: *International Conference on Learning Representations* (2015). p. 43405. Available online at: http://arxiv.org/abs/1412.6572
239. Finlayson SG, Bowers JD, Ito J, Zittrain JL, Beam AL, Kohane IS. Adversarial attacks on medical machine learning. *Science*. (2019) 363:1287–9. doi: 10.1126/science.aaw4399
240. Robinson R, Valindria VV, Bai W, Oktay O, Kainz B, Suzuki H, et al. Automated quality control in image segmentation: application to the UK biobank cardiovascular magnetic resonance imaging study. *J Cardiovasc Magn Reson*. (2019) 21:18. doi: 10.1186/s12968-019-0523-x
241. Heo J, Lee HB, Kim S, Lee J, Kim KJ, Yang E, et al. Uncertainty-aware attention for reliable interpretation and prediction. In: Bengio S, Wallach H, Larochelle H, Grauman K, Cesa-Bianchi N, Garnett R, editors. *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018*. Montréal, QC: Curran Associates Inc. (2018). p. 909–18.
242. Ruijsink B, Puyol-Antón E, Oksuz I, Sinclair M, Bai W, Schnabel JA, et al. Fully automated, quality-controlled cardiac analysis from CMR: validation and large-scale application to characterize cardiac function. *J Am Coll Cardiol*. (2019). doi: 10.1016/j.jcmg.2019.05.030
243. Tarroni G, Oktay O, Bai W, Schuh A, Suzuki H, Passerat-Palmbach J, et al. Learning-based quality control for cardiac MR images. *IEEE Trans Med Imaging*. (2019) 38:1127–38. doi: 10.1109/TMI.2018.2878509
244. Peng B, Zhang L. Evaluation of image segmentation quality by adaptive ground truth composition. In: *European Conference on Computer Vision*. Berlin; Heidelberg: Springer Berlin Heidelberg (2012). p. 287–300.
245. Zhou L, Deng W, Wu X. Robust image segmentation quality assessment without ground truth. *arxiv* (2019) abs/1903.08773. Available online at: http://arxiv.org/abs/1903.08773 (accessed September 1, 2019).
246. Alansary A, Folgoc LL, Vaillant G, Oktay O, Li Y, Bai W, et al. Automatic view planning with multi-scale deep reinforcement learning agents. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI, 2018*. Springer (2018). p. 277–85.
247. Dangi S, Linte CA, Yaniv Z. Cine cardiac MRI slice misalignment correction towards full 3D left ventricle segmentation. In: Fei B, Webster RJ. *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*. Houston, TX: SPIE (2018). p. 1057607.
248. Tarroni G, Oktay O, Sinclair M, Bai W, Schuh A, Suzuki H, et al. A comprehensive approach for learning-based fully-automated inter-slice motion correction for short-axis cine cardiac MR image stacks. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G, editors. *21st International Conference on Medical Image Computing and Computer Assisted Intervention—MICCAI 2018*. Granada (2018). p. 268–76.
249. Oksuz I, Clough J, Bai W, Ruijsink B, Puyol-Antón E, Cruz G, et al. High-quality segmentation of low quality cardiac MR images using k-space artefact correction. In: Cardoso MJ, Feragen A, Glocker B, Konukoglu E, Oguz I, Unal G, et al., editors. *International Conference on Medical Imaging with Deep Learning, MIDL, 2019*. London: PMLR (2019). p. 380–9.
250. Meng Q, Sinclair M, Zimmer V, Hou B, Rajchl M, Toussaint N, et al. Weakly supervised estimation of shadow confidence maps in fetal ultrasound imaging. *IEEE Trans Med Imaging*. (2019) 38:2755–67. doi: 10.1109/TMI.2019.2913311
251. Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative adversarial networks for noise reduction in low-dose CT. *IEEE Trans Med Imaging*. (2017) 36:2536–45. doi: 10.1109/TMI.2017.2708987
252. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, et al. CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. In: *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*. Honolulu, HI: AAAI Press (2019). p. 590–7. doi: 10.1609/aaai.v33i01.3301590
253. Dwork C, Roth A. The algorithmic foundations of differential privacy. *Foundat Trends Theor Comput Sci*. (2014) 9:211–407. doi: 10.1561/0400000042
254. Abadi M, Chu A, Goodfellow IJ, McMahan HB, Mironov I, Talwar K, et al. Deep learning with differential privacy. In: *The 2016 ACM SIGSAC Conference on Computer and Communications Security* Vienna (2016). p. 308–18.
255. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, et al. Practical secure aggregation for privacy-preserving machine learning. In: *The 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017*. Dallas, TX (2017). p. 1175–91.
256. Ryffel T, Trask A, Dahl M, Wagner B, Mancuso J, Rueckert D, et al. A generic framework for privacy preserving deep learning. In: *Privacy Preserving Machine Learning* (2018). p. 1–8.
257. Papernot N. A Marauder's map of security and privacy in machine learning: an overview of current and future research directions for making machine learning secure and private. In: *The 11th ACM Workshop on Artificial Intelligence and Security, CCS 2018*. Toronto, ON (2018). p. 1.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Chen, Qin, Qiu, Tarroni, Duan, Bai and Rueckert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.