



Mini Review: Deep Learning for Atrial Segmentation From Late Gadolinium-Enhanced MRIs

Kevin Jamart¹, Zhaohan Xiong¹, Gonzalo D. Maso Talou¹, Martin K. Stiles² and Jichao Zhao^{1*}

¹ Auckland Bioengineering Institute, The University of Auckland, Auckland, New Zealand, ² Waikato Clinical School, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand

OPEN ACCESS

Edited by:

Steffen Erhard Petersen,
Queen Mary University of London,
United Kingdom

Reviewed by:

Wenjia Bai,
Imperial College London,
United Kingdom
Joao Bicho Augusto,
Barts Heart Centre, United Kingdom
Redha Boubertakh,
Singapore Bioimaging Consortium
(A*STAR), Singapore

*Correspondence:

Jichao Zhao
j.zhao@auckland.ac.nz

Specialty section:

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

Received: 08 January 2020

Accepted: 21 April 2020

Published: 27 May 2020

Citation:

Jamart K, Xiong Z, Maso Talou GD, Stiles MK and Zhao J (2020) Mini Review: Deep Learning for Atrial Segmentation From Late Gadolinium-Enhanced MRIs. *Front. Cardiovasc. Med.* 7:86. doi: 10.3389/fcvm.2020.00086

Segmentation and 3D reconstruction of the human atria is of crucial importance for precise diagnosis and treatment of atrial fibrillation, the most common cardiac arrhythmia. However, the current manual segmentation of the atria from medical images is a time-consuming, labor-intensive, and error-prone process. The recent emergence of artificial intelligence, particularly deep learning, provides an alternative solution to the traditional methods that fail to accurately segment atrial structures from clinical images. This has been illustrated during the recent 2018 Atrial Segmentation Challenge for which most of the challengers developed deep learning approaches for atrial segmentation, reaching high accuracy (>90% Dice score). However, as significant discrepancies exist between the approaches developed, many important questions remain unanswered, such as which deep learning architectures and methods to ensure reliability while achieving the best performance. In this paper, we conduct an in-depth review of the current state-of-the-art of deep learning approaches for atrial segmentation from late gadolinium-enhanced MRIs, and provide critical insights for overcoming the main hindrances faced in this task.

Keywords: atrial fibrillation, left atrium, machine learning, image segmentation, convolutional neural network, LGE-MRI

INTRODUCTION

The ability to perform body imaging has been described as one of the most important revolutions in medicine of the past 1,000 years for its contribution to medical prevention, diagnosis, and prognosis (1). Since then, medical imaging has never ceased to improve, allowing cardiologists, and researchers to assess heart size using chest x-rays (2), to evaluate heart mechanical work with echocardiography imaging (3–5) and to accurately determine the heart's dimensions using cardiac magnetic resonance imaging (MRI) (6). Due to its good image quality, excellent soft-tissue contrast, and absence of ionizing radiation, MRI has become the gold standard modality to precisely identify patients' cardiac structures and etiology, guiding diagnosis and therapy decisions (7).

Improvements of MRI techniques, particularly with the aid of contrast agents such as gadolinium, led to the development of late gadolinium-enhanced MRI (LGE-MRI), allowing for the detection of scar tissue located within the myocardium. This technique has been extensively employed for clinical studies at Utah University (8–10) to analyze and understand the role of fibrosis and underlying structures that sustain atrial fibrillation (AF), the most common cardiac arrhythmia predicted to become a new epidemic in the coming decades (11, 12). They notably

demonstrated the correlation between an increased amount of fibrosis present in the left atrial (LA) wall and a poor outcome of AF ablation (10). Over time, LGE-MRIs have become a widely accepted technique of choice allowing the detection and quantification of scar tissues located in the atrial wall.

The currently widely used clinical practice, including those conducted at Utah University, to analyze atrial structures and determine and quantify fibrosis distribution is by performing manual segmentation of the LA chamber from LGE-MRIs. However, the LA cavity represents a small volume ($73 \pm 14.9 \text{ cm}^3$), constrained by a thin atrial wall (2–3 mm) and comprised of complex anatomy (13–15). Moreover, the anatomical structures surrounding the atria display similar intensities that can mislead some segmentation algorithms (16) (**Figure 1**). As a consequence, manual segmentation of the atrium is a time-consuming, labor-intensive, and error-prone process (8, 17, 18).

Before the advent of deep learning, researchers tried to develop and improve automated approaches to alleviate the burden of manual segmentation (19, 20). Earlier algorithms proposed would require important manual tunings such as thresholding methods or region growing approaches (21, 22). Other methods were later developed to provide a higher degree of automation using classifiers or clustering approaches such as *k*-nearest-neighbor (23) or *k*-means clustering (24), respectively. More recent methods, using statistical classifiers like support vector machine (25), active shape model (26), or multi-atlases (27) approaches, gained increasing interest for medical image analysis and cardiac segmentation. Though many of these approaches showed promising results, none presented enough consistency to be implemented widely in clinical practice.

In recent years, the development of more powerful computational hardware and the growth of clinical databases enabled deep learning, a subset of artificial intelligence (AI) (28–32) capable of automatic feature extraction and learning, to achieve tremendous advances notably in image classification and segmentation (33, 34). When applied to clinical images, deep learning even surpassed human-level accuracies for the detection of cancer on cervical images (35). Certain architectures employed for deep learning have also been proven to be very effective when applied to cardiac imaging. For example, Avendi et al. (36, 37) used a three-stage approach combining convolutional neural network (CNN), stacked encoder, and deformable models to segment the left ventricle (and later the right ventricle) on a small MRI dataset of 45 patients. On the other hand, Bai et al. (38) used a large MRI dataset provided by the UK Biobank database to develop their CNN for ventricular chamber assessment (volume, mass, ejection fraction) and segmentation, obtaining accuracy scores competing with human-level precision.

This increasing interest around deep learning can also be seen in the number of participants using deep learning approaches for the various challenges designed to promote the development of more robust methods for cardiac image segmentation (39–41). Atrial segmentation is becoming a matter of greater importance and can highly benefit from the development of deep learning. As an example, during the 2018 Atrial Segmentation Challenge, 15 of the 17 published approaches used deep learning to segment the

LA cavity from LGE-MRI images, yielding high accuracy results and outperforming conventional segmentation approaches (42). The number is in sharp contrast with the previous atrial segmentation challenge held in 2013, during which only one approach used a learning algorithm (16). Thus, this growing interest for deep learning in research challenges illustrates the shift occurring in atrial segmentation and more broadly in clinical imaging development, moving more and more toward deep learning-based approaches that will revolutionize clinical practice in the coming years.

In this paper, we aim to provide an analysis of the current deep learning technique used for atrial segmentation on LGE-MRIs. Firstly, we will describe some of the fundamental concepts employed in deep learning for medical image segmentation. Subsequently, we will detail the various deep learning approaches addressing the main obstacles faced performing automated atrial segmentation. Finally, we will conclude our review with an outline of future developments for atrial segmentation using deep learning and more broadly the future of AI in clinical practice.

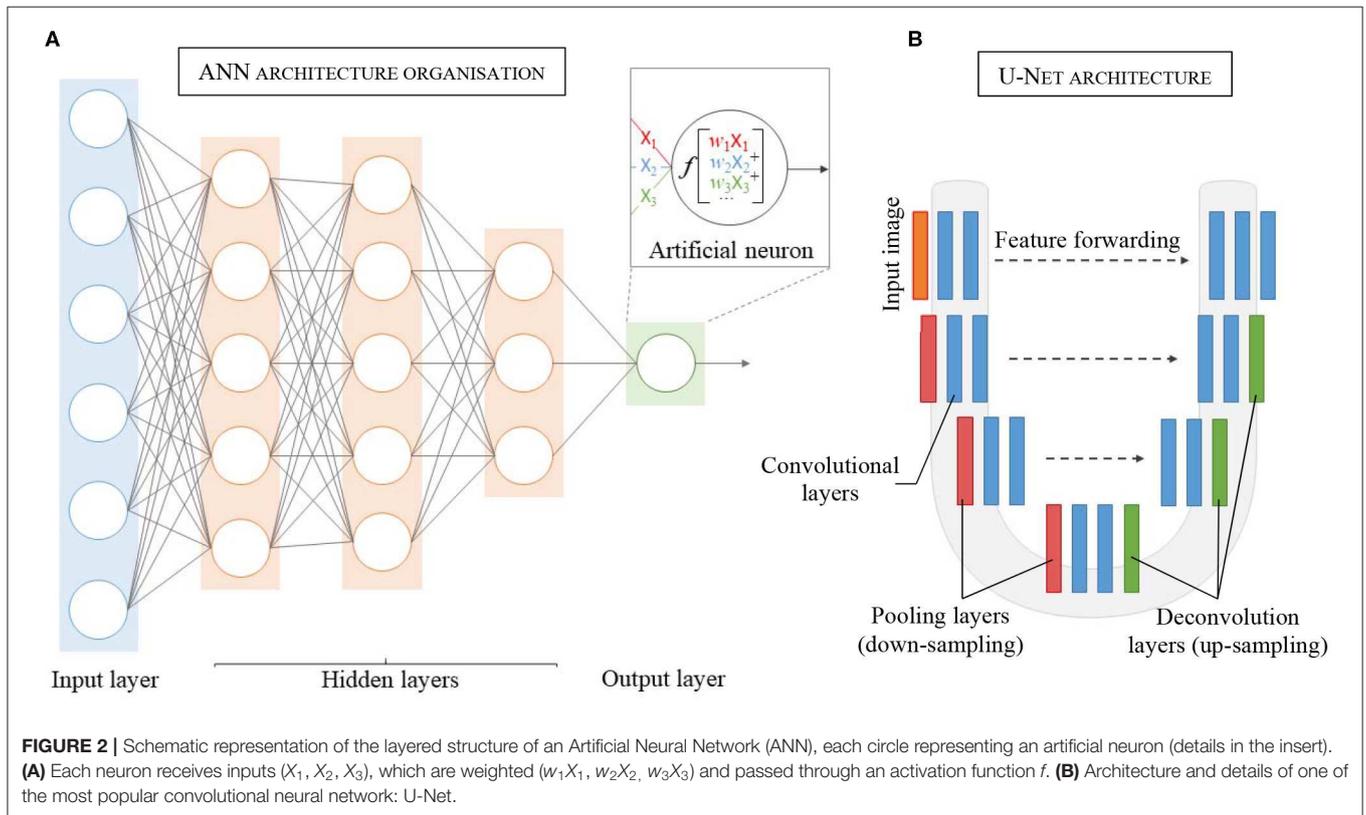
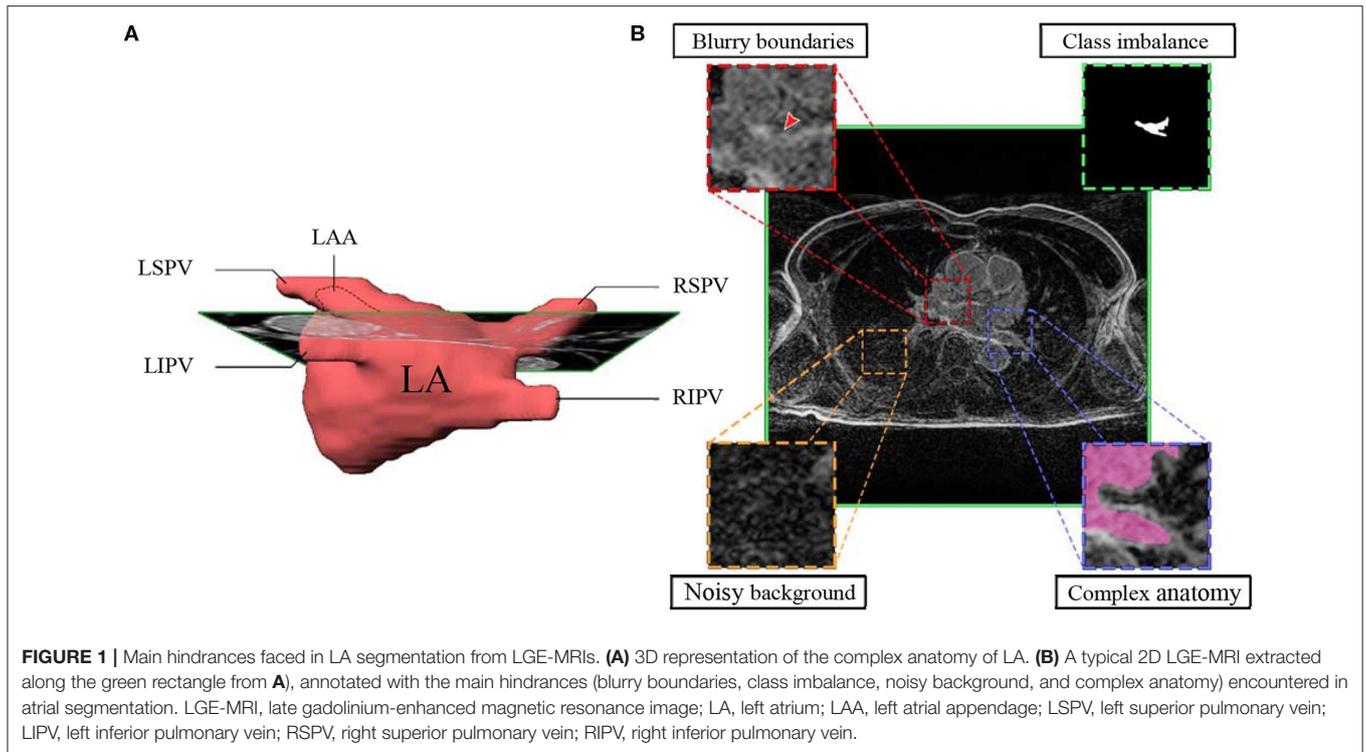
CORE CONCEPTS OF DEEP LEARNING

Since Alan Turing published his article “Computing Machinery and Intelligence” asking “Can machines think?” researchers have thrived to comprehend, develop, and achieve AI (43, 44) although today, after over 60 years, general AI is still not within reach. Nevertheless, in recent years, the growth of computer processing power and technologies has allowed researchers to develop algorithms capable of learning proficiently through deep learning using artificial neural networks (ANNs). As ANNs represent the most popular structure to perform deep learning, this section will describe the core concepts of ANNs and their various practical use in medical imaging.

Artificial Neural Networks

Inspired by the biological neural networks found in the human brain (45), an ANN represents a collection of connected and tunable computational units, called artificial neurons, organized in a layered structure comprising a network (**Figure 2A**). Each neuron is a processing unit that can take multiple inputs. Each input is multiplied by an adjustable parameter called weight. All weighted inputs are summed together and passed through a non-linear function to yield a single output (30). Neural networks can address complex, highly non-linear problems due to the layered and connected structure of ANNs. In particular, the introduction of more advanced feature learning tools such as convolutional layers, the improvement of large datasets and better activation functions, e.g., ReLU, greatly helped the development of deep learning for segmentation tasks.

The key attribute of an ANN lies in its ability to learn the unique traits of a dataset by adjusting its weights accordingly during a training process. Typically, the weights are randomly initialized at the start of training. The training process can then be described in three consecutive phases: (1) forward propagation, (2) error calculation, and (3) back-propagation. In the forward propagation stage, the input data (e.g., LGE-MRI image) is fed to the network and flows through the different layers that extract



the characteristic traits of the data, to ultimately yield a prediction (e.g., desired segmented image). The prediction is then compared to a reference data (e.g., manually segmented image by experts),

called labeled data, and error is calculated using a dedicated function (called loss function). Finally, the weights are modified to minimize the estimated error, improving prediction accuracy.

These three phases are repeated several times until the error converges to a significant minimized value.

Different Tasks, Different Networks

Medical imaging encompasses a wide field of applications, and different tasks can represent different aspects of a diagnosis. Examples include the detection of an abnormal ECG signal, its classification as AF (46), or even atrial segmentation for planning for AF ablation (47). Therefore, each task requires a specific ANN architecture to properly model the desired operator, as the inputs and output can be drastically different depending on the nature of the task to be performed.

The number, types, and connections of layers in an ANN defines the network architecture. The CNN model is one of the most widely employed architectures in image analysis. CNN is a specific ANN architecture in which its hidden layers comprise one or more convolutional layers. The convolutional layers act as feature extractors from the input image, applying different convolution kernels to the initial image to generate feature maps containing meaningful information. Moreover, in convolutional layers, each artificial neuron receives their inputs from multiple neighboring neurons from the previous layer, sharing their weights and keeping the most spatially relevant information. This feature also allows a reduction in the number of parameters to adjust and therefore lowers the computational processing cost. Generally inserted in between sets of successive convolutional layers are pooling layers that are used to reduce the dimensionality of each generated feature map while retaining the relevant information. This down-sampling of the feature maps, typically by a factor of two, allows reduction of the computational cost while enlarging the field of view for the later convolutional layers.

For CNNs dedicated to image classification or detection, the architectures usually incorporate a fully connected layer as an end layer to summarize all information contained in the feature maps into a unique final prediction (output). Furthermore, CNNs can also be adapted for segmentation tasks by discarding the final fully connected layer and incorporating up-convolution layers in the network (35). These networks are called fully convolutional networks (FCNs). Up-convolution layers allow up-sampling of the feature maps to produce, *in fine*, output with the same size as the original input size (48). Thus, FCNs using up-convolution layers can perform pixel-wise prediction and therefore image segmentation.

First proposed by Long et al. (33) for semantic segmentation, the FCN architecture has been adapted and further extended for medical imaging notably with U-Net, a U-shape architecture (Figure 2B) developed for segmentation of histological images (48). By using skip-connections between down-sampled feature maps and up-sampled feature maps, the U-Net architecture allows features forwarding between the encoding part and the decoding part of the network, preventing singularities and achieving higher accuracy (49–51). After winning the ISBI cell tracking challenge in 2015, U-Net became the principal FCN architecture for medical imaging segmentation. Other studies further developed the U-shape architecture to use 3D images as

input to render the spatial resolution of anatomical structures more accurately (52, 53).

ATRIAL SEGMENTATION USING DEEP LEARNING

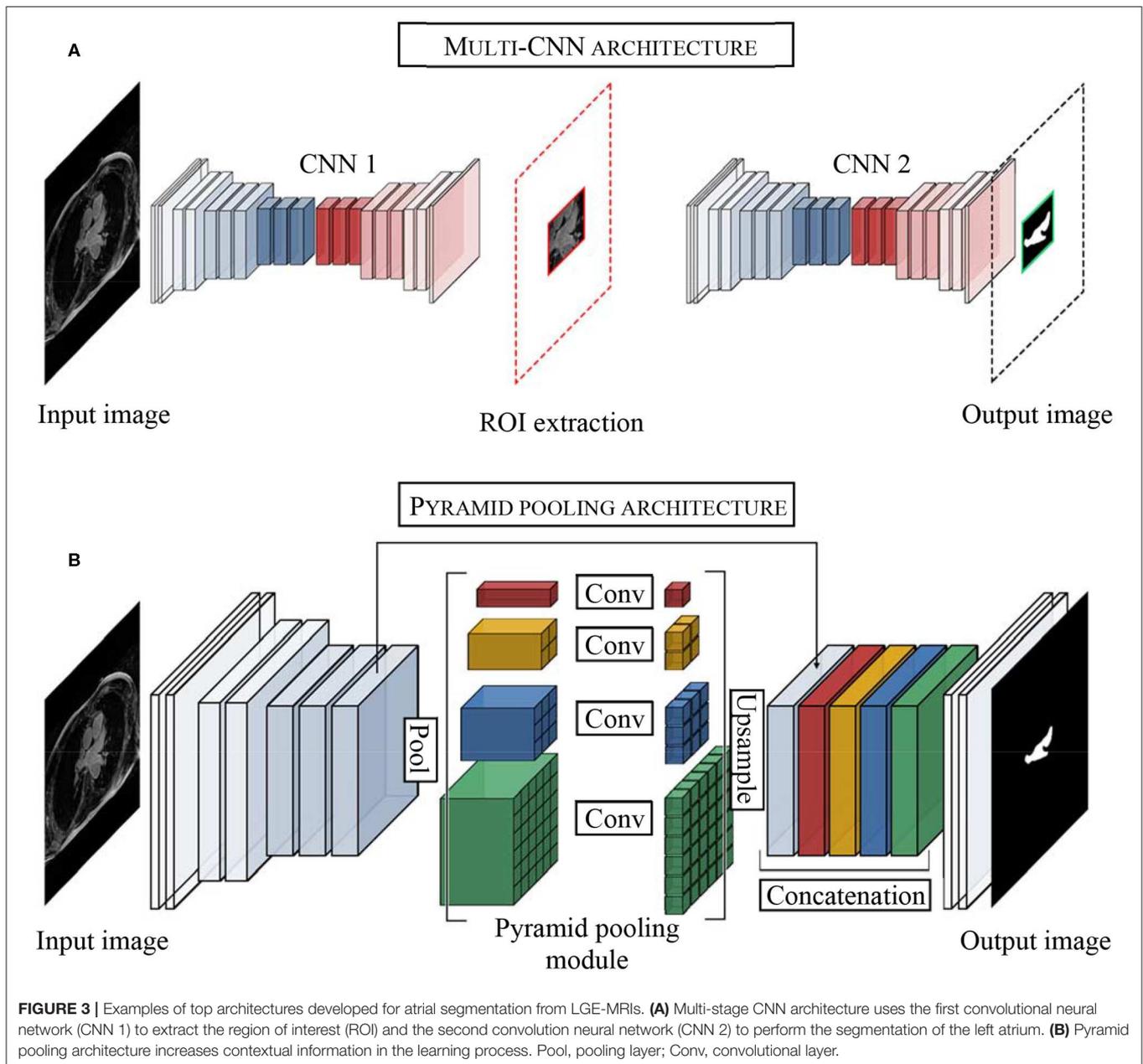
In this section, we provide a summary of the main difficulties encountered in atrial segmentation and the state-of-the-art deep learning approaches developed from LGE-MRIs to address them. To this regard, many of the methods reviewed were proposed for the MICCAI 2018 Atrial Segmentation Challenge which represented a cornerstone for the development of deep learning approaches for atrial segmentation from LGE-MRIs. Firstly, we will analyze the main methods employed to address class imbalance issues, a recurrent problem in segmentation of small structures such as the LA. Secondly, we will review the approaches developed to exploit image context providing more information for semantic segmentation of the LA using multi-scale strategies. Next, we will analyze the impact of loss function selection regarding either volumetric segmentation or surface segmentation. Finally, we will discuss the influence of the input dimensionality (2D/3D) for atrial segmentation when dataset size represents a significant shortcoming.

Multi-Stage CNN and Class Imbalance

One of the difficulties of atrial segmentation is that the atrial cavity represents only a small fraction of the image volume (~0.7%) and therefore creates a severe class imbalance between the over-represented background and the under-represented atrial structures, impairing the learning process. To address this issue, Vesal et al. (54) proposed to crop the input images from the center of the image, using fixed coordinates, to substantially remove the predominant background surrounding the LA. As a result, the learning process was entirely focused on a smaller region of interest (ROI), allowing better representation of the LA features. Based on a similar principle, other researchers (55–57), pushed this idea a step further by using a multi-CNN approach for atrial segmentation (Figure 3A). In their approaches, two consecutive networks were employed instead. The first CNN was specially trained to localize the LA on each input, allowing to subsequently crop out the unwanted background around the LA, as a prior step to segmentation. Then, the second network was dedicated to the segmentation task itself, focusing entirely on a small patch of each image.

Despite following a similar idea, it is important to distinguish these two methods. As the LA can show different positions on LGE-MRIs, using fixed coordinates from the center of the image to crop may result in unwanted cropping of relevant LA pixels. On the other hand, by dynamically centering the ROI on the LA for each input, multi-CNN approaches ensured the conservation of the atrial structures, cropping exclusively superfluous background pixels, and consequently optimizing background isotropy for the learning process.

To quantify the impact of each cropping approach, our recent study has investigated the importance of cropping the input patch



to the CNN either from the center of the image (image-centered) or from the center of the LA (center of mass/centroid of the atrium) using different patch sizes (ranging from 240×240 to 576×576) (58). When using center cropping of the image, we did not observe any significant influence of the patch size on the Dice score (92.03 vs. 91.95% Dice score for 240×240 and 512×512 image size, respectively). On the other hand, cropping the images from the centroid of the LA using dynamic cropping, we noticed a significant increase in the accuracy when using small patches (240×240) compared to large patches (576×576) (Dice score 92.86 vs. 92.26%, $p < 0.01$). The utilization of LA centroid-centered patches allows the CNN to process a more condensed region of the large LGE-MRI scan as the exact location of the LA

is known, reducing the class imbalance of each patch processed by the network.

Multi-Scale Approaches and Context Learning

Another problem that decreases segmentation performance and limits the extraction of relevant cues during the training phase is the inconsistency in the sizes of the LA anatomical structures such as the pulmonary veins or the left atrial appendage seen in LGE-MRIs from different patients.

He et al. (59) initially developed a pyramid pooling module, a multi-scale pooling, intended to prevent object misclassification by using image context information. By incorporating multi-scale

pooling, the CNN could associate contextual features, delivering more accurate classification. Based on this idea, Zhao et al. (60) proposed PSPNet, a neural network with pyramid pooling which incorporates object and image context to the learning process. These two approaches were developed using large miscellaneous datasets such as ImageNet (61), PASCAL VOC 2012 (62), or ADE20K dataset (63), and the pyramid pooling exploiting the context variability of the dataset, allowed to alleviate object misclassification or segmentation errors.

Inspired by He et al. (59) and the PSPNet developed by Zhao et al. (60), Bian et al. (64) proposed a multi-scale 2D CNN using spatial pyramid pooling to extract different scale features of the training dataset (**Figure 3B**). Thus, by means of different pooling kernel sizes and their combination, they proposed a CNN able to learn different cue size and improve network robustness against high shape variability usually encountered in clinical datasets.

However, the dataset employed for this approach (154 3D LGE-MRIs of the chest cavity) does not provide as much contextual variability as the large image database aforementioned, but rather displays the same object (the LA) in the same anatomical context (the thoracic cavity), providing only a few contextual variations to train on. Thus, arguably using pyramid pooling module for LA segmentation in the chest cavity might only show limited benefits from context learning.

Pyramid pooling also grants the ability to generate a fixed-length vector on a fully connected layer for classification tasks. This was illustrated by Chen et al. (65) using the pyramid pool module to extract more information from the dataset and classify the images between pre-/post-surgery, as they used a deeper U-Net to segment the LA simultaneously.

Based on the similar idea of incorporating multi-scale cues during the learning process, Vesal et al. (54) employed dilated convolution layers (also called atrous convolution layers) at the deepest level of their network. These convolution layers use dilatation rates to enlarge their receptive fields, allowing the network to learn different scale features (66). However, at each convolution the receptive field of each neuron is increased, therefore if not used wisely, receptive fields can become larger than the input image, resulting in a waste of memory while not improving the learning process.

These approaches ensure the incorporation of shallow features (spatial cues) and deep features (semantic cues) during the learning process. Therefore, combining effective class imbalance management with contextual cues could potentially improve even more the current methods. However, cropping to the smallest ROI possible using a first CNN of a two-stage approach, like Xia et al. drastically reduces the image context shown to the network. Therefore, the pyramid pooling module might not be able to provide contextual cues from the LA surrounding structures to improve the learning process. Moreover, during the cropping process, the input image size is significantly reduced. Thus, the use of dilated convolution for segmentation in the second network of this strategy becomes almost obsolete as the receptive fields would quickly grow larger than the input image during the learning process. Thus, fusing these strategies, although interesting, needs to be considered wisely.

Loss Function

The current main evaluation metrics employed in segmentation task using deep learning is the Dice score, for which a higher accuracy reflects almost exclusively a volume of pixel accurately annotated rather than well-defined anatomy. Hence, most of the deep learning approaches for segmentation employ pixel-wise segmentation relying either on cross-entropy loss function or dice loss function. However, these loss functions weigh more volume over contours, which can impair the learning of accurate boundaries in favor of a correct volume.

To improve boundary accuracy, several teams have developed contour-oriented loss functions. For example, Jia et al. (67) proposed a contour loss function (based on the pixel Euclidean distance) that decreases when the contour gets nearer to the reference contours of the label images during training, providing spatial distance information to the learning process. In their approach, they associated the dice function loss to obtain pixel-wise information, and their contour loss function for spatial information, achieving good shape consistency. In another strategy, Yang et al. (57) also defined a composite loss function, combining the overlap loss function (to reduce intersection between foreground and background) and a novel loss function called “focal positive loss” to guide the learning of voxel specific threshold and emphasize the foreground, improving, *in fine*, classification sensitivity. By recognizing ambiguous boundary location and enforcing positive prediction, this novel loss function improved the learning process and consequently the final atrial segmentation. However, these approaches did not obtain a better score than other approaches using more conventional loss function (e.g., dice loss, cross-entropy loss).

Therefore, it would be interesting to investigate the impact of a combined loss function allowing the network to learn from the volume (cross-entropy loss function or dice loss function) and from the contours of the LA. As segmentation tasks not only rely on minimizing volume error but also relies on boundaries accuracy (particularly for small structures). it is crucial to consider these two major aspects to ensure the reliability of the approach employed.

Spatial Context (2D vs. 3D)

Even if clinical datasets are becoming bigger and better with the creation of centralized databases, for example, the UK Biobank (with more than 90000 3D MRI scans) (68), most of the current clinical databases available remain of humble size, making it difficult for a CNN to provide robust generalized solutions for segmentation. As an example, the current largest LGE-MRI dataset with only 154 3D LGE-MRIs (which represent nearly 9,000 2D images for training) appears relatively small when compared to the hundreds of thousands of images used for the major classification challenges for which the proposed approaches reach outstanding accuracy (59, 69, 70).

Thus, in this race of performance, it is important to consider how to make the best of the dataset employed. To this regard, the choice of the image dimensions employed (2D and 3D) approaches must be considered wisely. As 2D approaches need considerably fewer trainable parameters to yield good results, they are less gluttonous regarding memory consumption, and

therefore require less time during the training process. Moreover, 2D approaches allow the processing of bigger batches of images compared to 3D approaches, as they require less memory to be processed. Therefore, 2D methods, using bigger batch size, help reduce gradient fluctuation and lead to faster convergence during the learning process. Additionally, 2D approaches can exploit more efficiently small datasets, reducing the risk of overfitting as the neural networks are fed with more images for the learning.

On the other hand, 3D approaches provide better spatial representation, fully exploiting data dimensionality as well as inter-slice continuity during training. This allows the network to learn major spatial features to render a more accurate 3D anatomy and yield, *in fine*, higher accuracy. Moreover, with the ever improvement of GPU technology, the current memory limitations will become of less importance in the near future; therefore, 3D approaches will become easier to use. Furthermore, as datasets are growing better and bigger, 3D approaches will be able to rely on more data and become more and more prominent in clinical imaging deep learning.

Nevertheless, relying on 2D images, Puybareau et al. (71) tried to improve the spatial representation of their dataset using a method called “pseudo-3D.” Their method employed the generation of color images from the 2D grayscale images, each slice being color expanded into the R, G, B space using slice $n-1$, slice n and slice $n+1$, to generate a three-channel image. This approach allows an improved spatial representation and alleviates low contrast intensity between atrial tissues and background and enrich the dataset. However, even if this approach does not provide the expected spatial representation, it can be a method of choice if resources are limited.

Following the multi-view approach developed by Mortazi et al. (72), Chen et al. investigated the possibility to combine 2D images and 3D representation (73). In their study, Chen et al. extracted the 2D images for each anatomical view (axial, coronal, and sagittal) from 100 3D LGE-MRIs. Then, they combined a first encoder-decoder network using long short term memory convolutional layers to preserve inter-slice correlation using the axial view, and a second network to learn complementary information from the sagittal and coronal views. Finally, the outputs for each view of the network were fused to yield LA and PV segmentation simultaneously. Using their approach, they obtained 90.83% Dice score accuracy for PV and atrial segmentation. Employing the same method, Yang et al. studied the influence of dilated convolution to counter image resolution variability encountered using a multi-view approach (74). Using 100 3D LGE-MRIs, they achieved 89.7% Dice score accuracy underlining the necessity to investigate systematic parameters tuning to obtain optimal performances on a task-specific basis.

In the present context, it is important to consider the trade-off using either a 2D approach requiring less memory and profiting more from the dataset (8,800 images rather than 154 3D LGE-MRIs) a 3D approach allowing more accurate spatial representation at the cost of longer and more difficult training. However, at the current stage, it is difficult to assess which method yields systematically better results. For example, during the 2018 Atrial Segmentation Challenge, the performances of 2D and 3D approaches remained very close (Table 1). Another

possibility is to use a multi-view approach combining 2D images from different views to improve the spatial representation. These methods require training each view separately before combining the different output for the final prediction. While interesting, these methods still need improvement to reach the current state-of-the-art for atrial segmentation. Therefore, further improvements need to be sought regarding the size of the dataset, the number of approaches compared and the metrics employed to be able to draw a better conclusion.

Evaluation Metrics

Another crucial point is to use metrics that provide a reliable evaluation of the final output using deep learning. One of the main scores employed is called Dice score and gauges the pixel-wise similarity between the predicted segmentation and the reference data. Dice score provides a good representation of the specificity and the sensitivity of the model. However, Dice score metric has some limitations as it only evaluates a percentage of pixel accurately annotated neglecting contours and shapes of organs that can be a critical part of diagnosis in clinical practice. Other metrics providing distance measurements, such as mean surface distance and Hausdorff maximum distance, are usually employed to provide an alternative evaluation. Mean surface distance estimates the average error (in mm) between the outer surfaces of the reference data and the predicted segmentation. Given the size and structure of LA, mean surface distance is a meaningful tool to reliably assess the anatomical boundaries of the predicted segmentation compared to the reference data. Hausdorff maximum distance (in mm) represents the maximum error between the surface of the predicted segmentation and the surface of the reference data. Therefore, Hausdorff distance indicates solely the distance at the worst part of the segmentation, providing only partial information of the correctness of the predicted segmentation. By combining mean surface distance and Hausdorff distance, it is possible to evaluate the fidelity of the boundaries of the segmented structures reliably. Finally, a more clinical aspect of the predictions can be examined to express the reliability of the approach by calculating volume error or anteroposterior atrial diameter error when comparing the segmented prediction with the reference image.

Atrial Wall and Scar Segmentation

While the deep learning methods for atrial cavity segmentation on LGE-MRIs are effective, the more clinically relevant tasks, such as LA wall and fibrosis (scar) segmentation, remain challenging. For LA wall segmentation, several approaches have been developed using traditional strategies such as multi-atlas segmentation or graph-cuts method (83, 84). However, currently no deep learning approaches have been proposed for direct LA wall segmentation from LGE-MRI. Yang et al. (85) proposed a hybrid approach combining multi-atlases and an unsupervised sparse auto-encoders for LA scar segmentation. A multi-atlas algorithm was used to segment the LA blood pool from the LGE-MRIs. Then, this initial LA cavity segmentation was dilated uniformly by 3 mm to include the LA wall. Next, they used a sparse auto-encoder to delineate and segment the fibrosis from the atrial wall. They achieved $90 \pm 0.12\%$ Dice score

TABLE 1 | Summary of deep learning approaches developed for the 2018 atrial segmentation challenge.

First Author	Summary	DC	Architecture	Pros/Cons
Xia et al. (56)	2 stage network (LA localization, LA segmentation), Dice loss	93.2	2x 3D U-Net	Good class imbalance management, highest performance/computationally expensive
Bian et al. (64)	LA segmentation using, ResNet101, atrous convolutional layers and pyramid pooling, online hard negative example mining (objective function)	92.6	2D Pyramid Network	Multi-scale representation/competitive training can reinforce overfitting
Vesal et al. (54)	LA segmentation using manual cropping, dilated convolution at the deepest level of U-Net, combination of Dice loss and cross-entropy loss function	92.6	3D U-Net	Class imbalance management, new loss function/Risk of loss of information using center cropping
Li et al. (55)	2 stage network: 3D U-Net for detection, Hierarchical Aggregation network (HAANet) for LA segmentation, Dice loss	92.3	3D U-Net + HAANet	Class imbalance management/Slow and small benefits from Hierarchical mechanism (0.4%)
Puybareau et al. (71)	Assembly of three 2D gray-scale images to create RGB 2D color image, transfer learning (VGG), multinomial loss function for LA segmentation	92.3	VGG-Net	Fast to train, pseudo-spatial representation/pseudo spatial representation not multi-view or 3D
Yang et al. (57)	2 stage approach: LA detection (Faster-RCNN), LA segmentation (U-Net). Deep supervision, transfer learning. Composite loss function: Overlap loss and Focal Positive loss	92.3	Faster-RCNN/3D U-Net	Good ROI detection, composite loss function /Recursive training with risk of overfitting
Chen et al. (73)	LA segmentation and classification (pre/post-ablation) of images, using cross-entropy and sigmoid loss function, respectively	92.1	2D U-Net	Fast to train (2D), interesting data augmentation
Jia et al. (67)	2 stage network (LA localization, LA segmentation), contour loss	90.7	3D U-Net	Contour loss/computationally expensive
Liu et al. (75)	Manual center cropping, evaluation of 2 different networks U-Net and FCN for LA segmentation, Dice loss	90.3	2D U-Net and FCN	Quick (2D)/Native U-Net/FCN
Borra et al. (76)	Otsu's algorithm for cropping, LA and pulmonary veins joined segmentation, Dice loss	89.8	3D U-Net	Otsu's for cropping/computationally expensive
de Vente et al. (77)	U-net for LA segmentation, Dice loss	89.7	2D U-Net	Fast (2D)/Native U-Net
Preetha et al. (78)	Deep supervision (79) and U-Net for LA segmentation	88.8	2D U-Net	Deep supervision, Fast (2D)/Native U-Net
Qiao et al. (80)	Multi-atlas selection and registration for LA segmentation	86.2	Multi-atlas	Groupwise registration/Slow prediction(multi-atlas)
Nuñez-Garcia et al. (81)	Multi-atlas whole heart labeling and shape-based atlas selection	85.9	Multi-atlas	Registration using global-atlases, shape based clustering/Difficulties do manage high variability in small dataset
Savioli et al. (82)	LA segmentation using V-Net and combination of mean squared error and Dice loss	85.1	3D V-Net	Composite loss function/computationally expensive

DC, Dice Score; LA, Left Atrium.

for blood pool segmentation and $78 \pm 0.08\%$ Dice score for fibrosis segmentation. In their subsequent study (86), by fine-tuning the sparse auto-encoder parameters, the accuracy was improved to $82 \pm 0.05\%$ Dice score for fibrosis segmentation. While showing promising results, with these methods being only developed and tested on 20 3D LGE MRIs, they remain untested on larger datasets to assess their reliability against a broader range of anatomical variabilities regarding LA structures and fibrosis. Chen et al. (73) developed a CNN with an attention mechanism (87) to highlight salient features (in this case, the enhanced pixels of the scar tissues on LGE MRIs) and to force the model to focus on the scars locations. With this approach, Chen et al. obtained 77.64% Dice score for atrial scar segmentation using 100 3D LGE MRIs. This lower score (compared to that obtained from LA cavity segmentation) is potentially due to the scarcity of the LA scar pixels, which are small patches of inhomogeneous enhanced pixels within the atrial wall, impairing the extraction of meaningful features for fibrosis identification during the learning process of the CNN.

While these methods require atrial wall segmentation to be performed before fibrosis detection, Li et al. proposed a hybrid approach using a graph-cuts framework combined with a multi-scale CNN approach for direct scar identification (88). In their approach, the LA and PV were initially delineated using a multi-atlas segmentation method. Then fibrosis was segmented and quantified using a graph-cut network in which two neural networks were dedicated to predicting edge weights. The first network was dedicated to predicting the probabilities of a node belonging to scar or normal tissue, while the second network was devoted to evaluate the connection between two nodes, yielding, in fine, the fibrosis segmentation. By embedding the CNN networks in the graph-cut framework, Li et al. obtained a mean Dice score of 70.2% for scar tissue segmentation, showing the possibility of effectively assessing LA fibrosis without the need for prior wall segmentation. Thus, even if the two networks employed did not directly perform the fibrosis segmentation task, the CNNs contributed to the optimisation process refining the graph-cut approach used in this study. However, these

methods tended to find fibrotic tissue out of the atrial wall boundaries regions, resulting in a drastic decrease in the final scores. Hence, the current models remain insufficient to provide anatomically accurate assessments allowing reliable fibrosis quantification due to the low Dice scores obtained. Thus, these approaches still require improvements to reach reliability and clinical applicability.

DISCUSSION AND CONCLUSION

In this paper, we provided an in-depth analysis of the main automatic approaches using deep learning for atrial cavity segmentation from LGE-MRIs. Most of the proposed deep learning approaches for atrial segmentation used FCNs, most notably the very popular U-Net architecture. While U-Net is widely used for medical image segmentation in many disciplines (38, 89, 90), the discrepancy in the accuracy obtained between different studies still presents inherent issues involved in the generalized implementation of such architectures. By presenting a normalized survey of U-Net for the task of atrial segmentation, we showed the importance of proper class imbalance management, appropriate features extraction process, and meaningful loss function selection to yield precise and accurate atrial segmentation.

The current leading approach for LA segmentation from LGE-MRIs dataset involved a two-stage 3D CNN method which reached a remarkable Dice accuracy of 93.2%, currently the best-benchmarked performance using 100 3D LGE MRIs (42). In this approach, the first network reduces class imbalance effectively while optimizing background isotropy using dynamic cropping, providing the second network with a targeted region for more localized segmentation. Additionally, they employed extensive data augmentation to enhance the generalization capability of their approach. Finally, they employed a 3D approach reinforcing the features' spatial representation, allowing them to obtain the current highest score for LA segmentation using machine learning.

Small training datasets represent one of the main limitations of clinical datasets as annotation and data gathering remains difficult. For example, the current largest LGE-MRIs dataset only contains 154 cases and therefore cannot effectively represent human anatomical variability. In fact, in order to improve performance, most of the developed approaches rely heavily on data augmentation such as affine transformations, cropping and scaling to virtually enlarge the dataset, also taking the risk of introducing more artifacts in the dataset. Moreover, the annotation process of anatomical structures is a complex and tedious process, which can be seen in the inter/intra-observer variability reported in several studies (38, 91). For example, atrial structures such as the mitral valve are difficult to segment due to the lack of clear anatomical border between LA and left ventricle. Moreover, the PVs are a very thin structure and represent a challenge for experts to distinguish from other structures on poorly contrasted images, and current protocols for defining the degree of extension of the PVs from the LA wall still remains subjective. Thus, this labeling uncertainty leads to some label variability in the dataset used, impairing the training process and potentially misleading the deep learning algorithm for the

prediction process. However, despite all these difficulties the study shows the success of deep learning approaches reaching a high Dice score accuracy (>90% Dice score), showing the importance of careful parameter selection and architecture design for achieving the best performance (38).

In this study, we showed the potential of applying deep learning to perform automatic segmentation of the LA directly from clinical imaging data. The current accuracy of the various approaches presented is promising for future clinical implementation by providing highly accurate anatomical maps of the LA. Additionally, multiple teams already proposed auspicious solutions for fibrosis assessment using deep learning, providing particularly valuable information for AF ablation strategies that could highly benefit initial patient stratification, diagnosis, prognosis, and potential guidance for an optimized ablation strategy. Moreover, the ability to generate high fidelity segmentations such as the LA opens the way for further applications of deep learning to segment other anatomical structures. For instance, high accuracy left atrial appendage segmentation would provide crucial information for atrial thrombosis risk assessment (92). Thus, practitioners would be able to provide adapted treatment strategies on time, potentially reducing the number of stroke accidents caused by migrating atrial thrombus. Additionally, LA segmentation approaches could also be applied to the RA, providing a better understanding of the role of fibrotic extents spread through the RA myocardium notably in sinoatrial diseases (93).

Finally, it is important to underline the limitation of the current metrics employed. As most of the segmentation tasks rely on pixel-wise classification, Dice score proposes an efficient way to determine the correctness of the overlapping prediction. However, Dice score can be defined as a volumetric metric as it weighs more generously toward an accurate volume over precise anatomical delimitations. In clinical practice, Dice score and volume accuracy are important for assessing LA dilatation, but becomes irrelevant when assessing boundaries of fine structures such as LA. Therefore, other metrics such as mean surface distance representing the distance between the labeled surface and the predicted surface should be considered to produce better anatomical accuracy evaluation. The Hausdorff distance, representing the maximum distance between two surfaces, can also be used to evaluate the maximum error between prediction and label, potentially guiding algorithms to minimize their maximum error. Moreover, other limitations such as variations in image quality and resolution or the introduction of image artifacts intrinsic to scanner manufacturer have to be taken to account for future clinical deployment. At the current stage, no study has investigated the influence of LGE-MRI image quality on the Dice score but empirically, the best image quality tends to yield higher accuracy scores. However, in clinical practice image quality can vary tremendously as cardiac motion, body fat, and chest breathing motion, amongst others, can generate artifacts to various degrees on the final images. Therefore, to provide good generalization capacity, deep learning models have to be able to extract meaningful features regardless of the quality of the image. Similarly to the image quality issue, to obtain good generalization capacity, a network should be trained with many images from many different scanners. Thus, large multi-center datasets need

to be built to ensure satisfying scanner variability and image quality variability representation for the learning process. Finally, it is crucial to promote deep models with efficient inherent generalization capabilities, as different image resolutions can represent a major difficulty for deep learning models using large scale datasets. However, promising results were demonstrated using pyramid pooling architecture ensuring extraction of multi-scale features. Thus, at the current stage efforts remain to be made to develop a deep learning model satisfying these criteria for further clinical deployment.

With the development of computational hardware and the general effort to enrich medical image databases, the effectiveness of deep learning will only improve with time. Arguably, the current trend would lead to improve all fields of clinical practices as AI technologies become more widely developed and implemented. Furthermore, the current flourishing of the deep learning approaches in all areas of medical practice has already breached out research. Despite initial professional reluctance, AI technologies will become of major importance in the near future.

REFERENCES

- Angell M, Kassirer J, Relman AS. Looking back on the millennium in medicine [editorial]. *N Engl J Med.* (2000) 342:42–9. doi: 10.1056/NEJM200001063420108
- Joo HS, Wong J, Naik VN, Savodelli GL. The value of screening preoperative chest X-rays: a systematic review. *Can J Anaesthesia.* (2005) 52:568–74. doi: 10.1007/BF03015764
- Sahn DJ, DeMaria A, Kisslo J, Weyman A. Recommendations regarding quantitation in M-mode echocardiography: results of a survey of echocardiographic measurements. *Circulation.* (1978) 58:1072–83. doi: 10.1161/01.CIR.58.6.1072
- Nagueh SF, Appleton CP, Gillebert TC, Marino PN, Oh JK, Smiseth OA, et al. Recommendations for the evaluation of left ventricular diastolic function by echocardiography. *J Am Soc Echocardiogr.* (2009) 10:165–93. doi: 10.1016/j.echo.2008.11.023
- Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of cardiovascular imaging. *J Am Soc Echocardiogr.* (2015) 16:233–71. doi: 10.1016/j.echo.2014.10.003
- La AG, Claessen G, de Bruaene Van A, Pattyn N, Van JC, Gewillig M, et al. Cardiac MRI: a new gold standard for ventricular volume quantification during high-intensity exercise. *Circul Cardiovasc Imaging.* (2013) 6:329–38. doi: 10.1161/CIRCIMAGING.112.980037
- Karamitsos TD, Francis JM, Myerson S, Selvanayagam JB, Neubauer S. The role of cardiovascular magnetic resonance imaging in heart failure. *J Am College Cardiol.* (2009) 54:1407–24. doi: 10.1016/j.jacc.2009.04.094
- Oakes RS, Badger TJ, Kholmovski EG, Akoum N, Burgon NS, Fish EN, et al. Detection and quantification of left atrial structural remodeling using delayed enhancement MRI in patients with atrial fibrillation. *Circulation.* (2009) 119:1758. doi: 10.1161/CIRCULATIONAHA.108.811877
- Akoum N, Fernandez G, Wilson B, McGann C, Kholmovski E, Marrouche N. Association of atrial fibrosis quantified using LGE-MRI with atrial appendage thrombus and spontaneous contrast on transesophageal echocardiography in patients with atrial fibrillation. *J Cardiovascul Electrophysiol.* (2013) 24:1104–9. doi: 10.1111/jce.12199
- McGann C, Akoum N, Patel A, Kholmovski E, Revelo P, Damal K, et al. Atrial fibrillation ablation outcome is predicted by left atrial

AUTHOR CONTRIBUTIONS

KJ and JZ conceived and designed the work. KJ searched and read the literature and drafted the manuscript. ZX, GM, MS, and JZ provided guidelines, critical revision, and insightful comments to improve the manuscript. All authors read and approved the manuscript.

FUNDING

This work was supported by the Health Research Council of New Zealand.

ACKNOWLEDGMENTS

We would like to thank our colleagues: Dr. Nawshin Dastagir, Joseph Ashby, and Christopher Walker for their precious comments and insights that greatly helped to improve the manuscript. We would also like to thank Vincent Guichot who provided great assistance for the creation of the figures.

- remodeling on MRI. *Circul Arrhythmia Electrophysiol.* (2014) 7:23–30. doi: 10.1161/CIRCEP.113.000689
- Chugh SS, Blackshear JL, Shen W-K, Hammill SC, Gersh BJ. Epidemiology and natural history of atrial fibrillation: clinical implications. *J Am College Cardiol.* (2001) 37:371–8. doi: 10.1016/S0735-1097(00)01107-4
- Anter E, Jessup M, Callans DJ. Atrial fibrillation and heart failure: treatment considerations for a dual epidemic. *Circulation.* (2009) 119:2516–25. doi: 10.1161/CIRCULATIONAHA.108.821306
- Maceira AM, Cosin-Sales J, Roughton M, Prasad SK, Pennell DJ. Reference left atrial dimensions and volumes by steady state free precession cardiovascular magnetic resonance. *J Cardiovascul Magnet Resonance.* (2010) 12:65. doi: 10.1186/1532-429X-12-65
- Zhao J, Hansen BJ, Wang Y, Csepe TA, Sul LV, Tang A, et al. Three-dimensional integrated functional, structural, and computational mapping to define the structural ‘fingerprints’ of heart-specific atrial fibrillation drivers in human heart *ex vivo*. *J Am Heart Assoc.* (2017) 6:e005922. doi: 10.1161/JAHA.117.005922
- Wang Y, Xiong Z, Nalar A, Hansen BJ, Kharche S, Seemann G, et al. A robust computational framework for estimating 3D Bi-Atrial chamber wall thickness. *Comput Biol Med.* (2019) 114:103444. doi: 10.1016/j.combiomed.2019.103444
- 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. *Proc IEEE Med Imaging.* (2015) 34:1460–73. doi: 10.1109/TMI.2015.2398818
- Petitjean C, Dacher J-NJM. A review of segmentation methods in short axis cardiac MR images. *Med Image Analysis.* (2011) 15:169–84. doi: 10.1016/j.media.2010.12.004
- Caudron J, Fares J, Lefebvre V, Vivier P-H, Petitjean C, Dacher JN. Cardiac MRI assessment of right ventricular function in acquired heart disease: factors of variability. *Acad Radiol.* (2012) 19:991–1002. doi: 10.1016/j.acra.2012.03.022
- 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
- Litjens G, Ciompi F, Wolterink JM, de Vos BD, Leiner T, Teuwen J, et al. State-of-the-art deep learning in cardiovascular image analysis. *JACC Cardiovascul Imaging.* (2019) 12:1549–65. doi: 10.1016/j.jcmg.2019.06.009
- Metz C, Schaap M, Weustink A, Mollet N, van Walsum T, Niessen WJM, et al. Coronary centerline extraction from CT coronary angiography images

- using a minimum cost path approach. *Med Phys.* (2009) 36:5568–79. doi: 10.1118/1.3254077
22. Feng C, Zhang S, Zhao D, Li C. Simultaneous extraction of endocardial and epicardial contours of the left ventricle by distance regularized level sets. *Med Phys.* (2016) 43:2741–55. doi: 10.1118/1.4947126
 23. Bezdek JC, Hall L, Clarke LP. Review of MR image segmentation techniques using pattern recognition. *Med Phys.* (1993) 20:1033–48. doi: 10.1118/1.597000
 24. Kaus MR, Von Berg J, Weese J, Niessen W, Pekar V. Automated segmentation of the left ventricle in cardiac MRI. *Med Image Anal.* (2004) 8:245–54. doi: 10.1016/j.media.2004.06.015
 25. Lao Z, Shen D, Liu D, Jawad AF, Melhem ER, Launer LJ, et al. Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine. *Acad Radiol.* (2008) 15:300–13. doi: 10.1016/j.acra.2007.10.012
 26. Van Assen HC, Danilouchkine MG, Behloul F, Lamb HJ, van der Geest RJ, Reiber JH, et al. Cardiac LV segmentation using a 3D active shape model driven by fuzzy inference. *Int Conf Med Image Comput Comput Assisted Intervent.* (2003) 2878:533–40. doi: 10.1007/978-3-540-39899-8_66
 27. Isgum I, Staring M, Rutten A, Prokop M, Viergever MA, Van Ginneken B. Multi-atlas-based segmentation with local decision fusion-application to cardiac and aortic segmentation in CT scans. *Proc IEEE Med Imaging.* (2009) 28:1000–10. doi: 10.1109/TMI.2008.2011480
 28. Deng L, Yu D. Deep learning: methods and applications. *Found Trends Signal Proces.* (2014) 7:197–387. doi: 10.1561/2000000039
 29. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* (2015) 521:436–44. doi: 10.1038/nature14539
 30. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA: MIT Press (2016).
 31. Garcia-Garcia A, Orts-Escolano S, Oprea S, Villena-Martinez V, Garcia-Rodriguez J. A review on deep learning techniques applied to semantic segmentation. *arXiv.* (2017) 17, 41–65. doi: 10.1016/j.asoc.2018.05.018
 32. Ker J, Wang L, Rao J, Lim T. Deep learning applications in medical image analysis. *IEEE Access.* (2017) 6:9375–89. doi: 10.1109/ACCESS.2017.2788044
 33. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA (2016). p. 3431–40. doi: 10.1109/CVPR.2015.7298965
 34. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV (2016). p. 770–8. doi: 10.1109/CVPR.2016.90
 35. Hu L, Bell D, Antani S, Xue Z, Yu K, Horning MP, et al. An observational study of deep learning and automated evaluation of cervical images for cancer screening. *J National Cancer Institue.* (2019) 74:343–4. doi: 10.1097/OGX.0000000000000687
 36. Avendi M, 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
 37. Avendi MR, Kheradvar A, Jafarkhani H. Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach. *Magnetic Resonance Med.* (2017) 78:2439–48. doi: 10.1002/mrm.26631
 38. 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 Cardiovascul Magnetic Resonance.* (2018) 20:65. doi: 10.1186/s12968-018-0471-x
 39. Suinesiaputra A, Cowan BR, Finn JP, Fonseca CG, Kadish AH, Lee DC, et al. Left ventricular segmentation challenge from cardiac MRI: a collation study. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 88–97. doi: 10.1007/978-3-642-28326-0_9
 40. Ho SY, Cabrera JA, Sanchez-Quintana D. Left atrial anatomy revisited. *Circul Arrhythmia Electrophysiol.* (2012) 5:220–8. doi: 10.1161/CIRCEP.111.962720
 41. Zuluaga MA, Bhatia K, Kainz B, Moghari MH, Pace DF. Reconstruction, segmentation, and analysis of medical images: first international workshops. In: *RAMBO 2016 and HVSMR 2016, Held in Conjunction with MICCAI 2016*. Athens, Greece (2017).
 42. Pop M, Sermesant M, Zhao J, Li S, McLeod K, Young AA, et al. Statistical atlases and computational models of the heart: atrial segmentation and LV quantification challenges: 9th international workshop. In: *STACOM 2018, Held in Conjunction With MICCAI 2018*, Granada, Spain, (2019).
 43. Turing AM. I.— Computing machinery and intelligence. *Mind LIX.* (1950) LIX:433–60. doi: 10.1093/mind/LIX.236.433
 44. Saygin AP, Cicekli I, Akman V. Turing test: 50 years later. *Minds Mach.* (2000) 10:463–518. doi: 10.1023/A:1011288000451
 45. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Mathemat Biol.* (1943) 5:115–33. doi: 10.1007/BF02478259
 46. Xiong Z, Nash MP, Cheng E, Fedorov VV, Stiles MK, Zhao J. ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiol Measurement.* (2018) 39:094006. doi: 10.1088/1361-6579/aad9ed
 47. 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 on Medical Imaging.* (2018) 38:515–24. doi: 10.1109/TMI.2018.2866845
 48. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. (2015). p. 234–41. doi: 10.1007/978-3-319-24574-4_28
 49. Drozdal M, Vorontsov E, Chartrand G, Kadoury S, Pal C. The importance of skip connections in biomedical image segmentation. *Deep Learn Data Label Med Appl.* (2016) 10008:179–87. doi: 10.1007/978-3-319-46976-8_19
 50. Mao X, Shen C, Yang YB. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. *Adv Neural Inform Proc Syst.* (2016) arXiv:1603.09056:2802–10.
 51. Orhan AE, Pitkow X. Skip Connections Eliminate Singularities. In: *Sixth International Conference on Learning Representations*. Toulon (2017).
 52. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. *Int Conf Med Image Comput Comput Assisted Interv.* (2016) 9901:424–32. doi: 10.1007/978-3-319-46723-8_49
 53. Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV)*: IEEE (2016). p. 565–71. doi: 10.1109/3DV.2016.79
 54. Vesal S, Ravikumar N, Maier A. Dilated convolutions in neural networks for left atrial segmentation in 3d gadolinium enhanced-MRI. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 319–28. doi: 10.1007/978-3-030-12029-0_35
 55. Li C, Tong Q, Liao X, Si W, Sun Y, Wang Q, et al. Attention based hierarchical aggregation network for 3D left atrial segmentation. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 255–64. doi: 10.1007/978-3-030-12029-0_28
 56. Xia Q, Yao Y, Hu Z, Hao A. Automatic 3D Atrial Segmentation from GE-MRIs Using Volumetric Fully Convolutional Networks. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 211–20. doi: 10.1007/978-3-030-12029-0_23
 57. Yang X, Wang N, Wang Y, Wang X, Nezafat R, Ni D, et al. Combating uncertainty with novel losses for automatic left atrium segmentation. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 246–54. doi: 10.1007/978-3-030-12029-0_27
 58. Jamart K, Xiong Z, Talou GM, Stiles MK, Zhao J. Two-stage 2D CNN for automatic atrial segmentation from LGE-MRIs. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2019). p. 81–9. doi: 10.1007/978-3-030-39074-7_9
 59. He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. In: *Proceedings of the IEEE on Pattern Analysis and Machine Intelligence*. (2014) 37:1904–16. doi: 10.1007/978-3-319-10578-9_23
 60. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2017). p. 2881–90. doi: 10.1109/CVPR.2017.660
 61. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. Imagenet large scale visual recognition challenge. *Int J Comput Vision.* (2015) 115:211–52. doi: 10.1007/s11263-015-0816-y
 62. Everingham M, Van Gool L, Williams C, Winn J, Zisserman A. *The Pascal Visual Object Classes Challenge 2012 (voc2012) Results*. (2012).

- Available online at: URL <http://www.pascal-network.org/challenges/VOC/voc2011/workshop/index.html>.
63. Zhou B, Zhao H, Puig X, Fidler S, Barriuso A, Torralba A. Scene parsing through ade20k dataset. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2017). p. 633–41. doi: 10.1109/CVPR.2017.544
 64. 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: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 237–45. doi: 10.1007/978-3-030-12029-0_26
 65. Chen C, Bai W, Rueckert D. Multi-task learning for left atrial segmentation on GE-MRI. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 292–301. doi: 10.1007/978-3-030-12029-0_32
 66. Yu F, Koltun V, Funkhouser T. Dilated residual networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2017). p. 472–80. doi: 10.1109/CVPR.2017.75
 67. Jia S, Despinasse A, Wang Z, Delingette H, Pennec X, Jais P, et al. Automatically segmenting the left atrium from cardiac images using successive 3D U-nets and a contour loss. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 221–9. doi: 10.1007/978-3-030-12029-0_24
 68. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. (2015) 12:e1001779. doi: 10.1371/journal.pmed.1001779
 69. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, et al. Microsoft coco: common objects in context. In: *European Conference on Computer Vision*. (2015). p. 740–55.
 70. Wu B, Chen W, Fan Y, Zhang Y, Hou J, Huang J, et al. Tencent ML-images: a large-scale multi-label image database for visual representation learning. In: *IEEE Access*. (2019). p. 172683–93. doi: 10.1109/ACCESS.2019.2956775
 71. Puybareau É, Zhao Z, Khouli Y, Carlinet E, Xu Y, Lacotte J, et al. Left atrial segmentation in a few seconds using fully convolutional network and transfer learning. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. (2018). p. 339–47. doi: 10.1007/978-3-030-12029-0_37
 72. 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: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. (2017). p. 377–85. doi: 10.1007/978-3-319-66185-8_43
 73. Chen J, Yang G, Gao Z, Ni H, Angelini E, Mohiaddin R, et al. Multiview two-task recursive attention model for left atrium and atrial scars segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer (2018). p. 455–63. doi: 10.1007/978-3-030-00934-2_51
 74. 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: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. (2018). p. 1123–7. doi: 10.1109/EMBC.2018.8512550
 75. Liu Y, Dai Y, Yan C, Wang K. Deep learning based method for left atrial segmentation in GE-MRI. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 311–8. doi: 10.1007/978-3-030-12029-0_34
 76. Borra D, Masci A, Esposito L, Andalò A, Fabbri C, Corsi C. A semantic-wise convolutional neural network approach for 3-D left atrium segmentation from late gadolinium enhanced magnetic resonance imaging. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 329–38. doi: 10.1007/978-3-030-12029-0_36
 77. de Vente C, Veta M, Razeghi O, Niederer S, Pluim J, Rhode K, et al. Convolutional neural networks for segmentation of the left atrium from gadolinium-enhancement MRI images. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 348–56. doi: 10.1007/978-3-030-12029-0_38
 78. Preetha CJ, Haridasan S, Abdi V, Engelhardt S. Segmentation of the left atrium from 3D gadolinium-enhanced MR images with convolutional neural networks. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 265–72. doi: 10.1007/978-3-030-12029-0_29
 79. Yu F, Wang D, Shelhamer E, Darrell T. Deep layer aggregation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2018). p. 2403–12. doi: 10.1109/CVPR.2018.00255
 80. Qiao M, Wang Y, van der Geest RJ, Tao Q. Fully automated left atrium cavity segmentation from 3D GE-MRI by multi-atlas selection and registration. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 230–6. doi: 10.1007/978-3-030-12029-0_25
 81. Nuñez-García M, Zhuang X, Sanroma G, Li L, Xu L, Butakoff C, et al. Left atrial segmentation combining multi-atlas whole heart labeling and shape-based atlas selection. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 302–10. doi: 10.1007/978-3-030-12029-0_33
 82. Savioli N, Montana G, Lamata P. V-FCNN: volumetric fully convolution neural network for automatic atrial segmentation. In: *International Workshop on Statistical Atlases and Computational Models of the Heart*. Cham: Springer (2018). p. 273–81. doi: 10.1007/978-3-030-12029-0_30
 83. Veni G, Fu Z, Awate SP, Whitaker RT. Bayesian segmentation of atrium wall using globally-optimal graph cuts on 3D meshes. In: *International Conference on Information Processing in Medical Imaging*. (2013). p. 656–67. doi: 10.1007/978-3-642-38868-2_55
 84. Tao Q, Ipek EG, Shahzad R, 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 Magnet Resonance Imaging*. (2016) 44:346–54. doi: 10.1002/jmri.25148
 85. Yang G, Zhuang X, Khan H, Haldar S, Nyktari E, Ye X, et al. A fully automatic deep learning method for atrial scarring segmentation from late gadolinium-enhanced MRI images. In: *2017 IEEE 14th International Symposium on Biomedical Imaging*. (2017). p. 844–8. doi: 10.1109/ISBI.2017.7950649
 86. 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: M. Valdés Hernández and V. González-Castro, *Medical Image Understanding and Analysis*. Cham: Springer International Publishing (2017). p. 195–206. doi: 10.1007/978-3-319-60964-5_17
 87. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. *arXiv [preprint]*. arXiv:1409.0473 (2014)
 88. Li L, Wu F, Yang G, Xu L, Wong T, Mohiaddin R, et al. Atrial scar quantification via multi-scale CNN in the graph-cuts framework. *Med Image Anal.* (2020) 60:101595. doi: 10.1016/j.media.2019.101595
 89. Dong H, Yang G, Liu F, Mo Y, Guo Y. Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. In: *Annual Conference on Medical Image Understanding and Analysis*. (2017). p. 506–17. doi: 10.1007/978-3-319-60964-5_44
 90. Li X, Chen H, Qi X, Dou Q, Fu C-W, Heng PA. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *Proc IEEE Med Imaging*. (2018) 37:2663–74. doi: 10.1109/TMI.2018.2845918
 91. Bernard O, Lalonde A, Zotti C, Cervenansky F, Yang X, Heng P-A, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *Proc IEEE Med Imag*. (2018) 37:2514–25. doi: 10.1109/TMI.2018.2837502
 92. Jin C, Feng J, Wang L, Yu H, Liu J, Lu J, et al. Left atrial appendage segmentation using fully convolutional neural networks and modified three-dimensional conditional random fields. *IEEE J Biomed Health Inform*. (2018) 22:1906–16. doi: 10.1109/JBHI.2018.2794552
 93. Csepe TA, Kalyanasundaram A, Hansen BJ, Zhao J, Fedorov VV. Fibrosis: a structural modulator of sinoatrial node physiology and dysfunction. *Front Physiol*. (2015) 6:37. doi: 10.3389/fphys.2015.00037

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 Jamart, Xiong, Maso Talou, Stiles and Zhao. 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.