

Review

# Recent Progress in Epicardial and Pericardial Adipose Tissue Segmentation and Quantification Based on Deep Learning: A Systematic Review

Marin Benčević <sup>1,\*</sup>, Irena Galić <sup>1</sup>, Marija Habijan <sup>1</sup> and Aleksandra Pižurica <sup>2</sup>

<sup>1</sup> Faculty of Electrical Engineering, Computer Science and Information Technology, Josip Juraj Strossmayer University of Osijek, 31000 Osijek, Croatia; irena.galic@ferit.hr (I.G.); marija.habijan@ferit.hr (M.H.)

<sup>2</sup> TELIN-GAIM, Faculty of Engineering and Architecture, Ghent University, 9000 Ghent, Belgium; aleksandra.pizurica@ugent.be

\* Correspondence: marin.bencevic@ferit.hr

**Abstract:** Epicardial and pericardial adipose tissues (EAT and PAT), which are located around the heart, have been linked to coronary atherosclerosis, cardiomyopathy, coronary artery disease, and other cardiovascular diseases. Additionally, the volume and thickness of EAT are good predictors of CVD risk levels. Manual quantification of these tissues is a tedious and error-prone process. This paper presents a comprehensive and critical overview of research on the epicardial and pericardial adipose tissue segmentation and quantification methods, evaluates their effectiveness in terms of segmentation time and accuracy, provides a critical comparison of the methods, and presents ongoing and future challenges in the field. Described methods are classified into pericardial adipose tissue segmentation, direct epicardial adipose tissue segmentation, and epicardial adipose tissue segmentation via pericardium delineation. A comprehensive categorization of the underlying methods is conducted with insights into their evolution from traditional image processing methods to recent deep learning-based methods. The paper also provides an overview of the research on the clinical significance of epicardial and pericardial adipose tissues as well as the terminology and definitions used in the medical literature.

**Keywords:** epicardial adipose tissue; medical imaging; pericardial adipose tissue; segmentation; quantification



**Citation:** Benčević, M.; Galić, I.; Habijan, M.; Pižurica, A. Recent Progress in Epicardial and Pericardial Adipose Tissue Segmentation and Quantification Based on Deep Learning: A Systematic Review. *Appl. Sci.* **2022**, *12*, 5217. <https://doi.org/10.3390/app12105217>

Academic Editor: Giancarlo Mauri

Received: 22 April 2022

Accepted: 20 May 2022

Published: 21 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cardiovascular diseases (CVDs) are one of the major causes of death in the world, leading to almost a third of all global deaths [1]. There is a large volume of research linking cardiac adipose tissue (body fat) volume and thickness to several CVDs and general CVD risk. Clinically, the most significant cardiac adipose tissues are epicardial adipose tissue (EAT) and pericardial adipose tissue (PAT) [2]. EAT is a fatty tissue located inside the pericardial sac (pericardium), a thin protective layer around the heart. PAT is the union of all the adipose tissue around the pericardium, which includes EAT.

EAT and PAT have been shown to be markers of CVDs [3,4] but are not widely used in clinical practice due to how difficult it is to accurately measure their volume. Hence, computer-assisted segmentation and quantification of EAT and PAT can be used to improve diagnostics and CVD risk assessment. Segmentation refers to manual, semi-automatic, or fully-automatic identification of EAT and PAT regions on medical images, while quantification refers to measuring EAT and PAT's total volume and density around the heart. Therefore, quantification relies on accurate segmentation. In this paper, we review papers that segment whole EAT or PAT volumes from their surrounding tissues, and not papers that aim to segment different regions inside EAT or PAT, such as left atrial

EAT vs. left ventricular EAT. While there is evidence that the distribution of EAT around the heart is important [5], segmenting the whole volume is a prerequisite for the analysis of its distribution and is, therefore, the focus of this review.

Manual segmentation and quantification of EAT and PAT are tedious and time-consuming processes often prone to inter-observer variability due to their complex distribution around the heart [6]. The development of computer-aided medical software built upon advanced image processing algorithms reduces the load on radiologists and physicians and provides more input into treatment and disease prevention. Specifically, efficient methods to segment epicardial fat from CT scans are exceedingly beneficial when compared to other imaging modalities, namely because of coronary artery calcium (CAC) scoring. CAC scoring is a procedure done by a CT scan of the heart. CT scans obtained for CAC scoring are also appropriate for measuring EAT volume since it covers the heart and neighborhood. Due to how commonly CAC scoring is done, many medical institutions have existing datasets of CAC scoring CT scans [7], and CAC scoring is often performed as part of large clinical trials [8]. The dataset availability makes it easier to perform future studies of EAT, given an accurate and efficient enough method of quantification. Additionally, segmenting EAT from MRI is more difficult than from CT since the pericardium is often blurred due to partial volume effects [9]. The vast majority of existing EAT and PAT segmentation methods focus on quantifying EAT from CT. Segmenting and quantifying EAT is particularly challenging when compared to other cardiovascular segmentation tasks. EAT is distributed unevenly around the heart, giving it a conspicuous shape that is hard to detect automatically. A compounding issue is that epicardial and paracardial adipose tissues have similar intensities and are delineated by the pericardium, which is often too thin to be fully seen in CT or MRI images.

Although various segmentation methods for EAT and PAT have been reported, to the best of our knowledge there is no comprehensive review of these methods in the literature. We hope this review will be useful as a guideline and an indicator for further research directions and algorithm development.

This work's primary motivation is to provide a comprehensive and critical overview of automatic EAT and PAT segmentation and quantification methods. The objectives of this study are:

- to a comprehensive and critical overview of the field and thus increase its visibility and enable future research.
- highlight the advantages and shortcomings of current methods in the hopes of improving the quality of future research.
- to provide a discussion of existing challenges in the field and ways to overcome them.

The remainder of the paper is organized as follows. Section 2 gives a brief clinical background and explains differences in used terminology regarding EAT and PAT. Section 3 gives a theoretical overview of commonly used algorithms and methods for segmentation and quantification of EAT and PAT. Section 4 presents methods for EAT segmentation via pericardium delineation. In Section 5 we present direct EAT segmentation methods. Methods used for PAT segmentation are described in Section 6. Discussion about significant findings, remaining challenges, and future developments in the field are given in Section 7. Finally, the conclusion is given in Section 8.

### *Review Methodology*

The papers included in this review were collected in January of 2022 using various online databases including Science Citation Index Expanded, Emerging Sources Citation Index, Conference Proceedings Citation Index, BIOSIS Citation Index, Data Citation Index, MEDLINE, Derwent Innovations Index, KCI-Korean Journal Database, SciELO Citation Index, and the Russian Science Citation Index. We searched the databases for titles matching the query: (cardiac OR pericardial OR pericardiac OR epicardial OR heart) AND (adipose tissue OR fat) AND (segmentation OR quantification). The search resulted in 70 total results consisting of papers that mention EAT segmentation and quantification. We only include

the papers for those that describe methods for automatic EAT or PAT segmentation or quantification (as opposed to papers that measure EAT manually e.g., for clinical trials and similar), resulting in a total of 23 papers. As this is a clear inclusion criterium, there were no disagreements between the authors about which studies to include. The PRISMA flow diagram is shown in Figure 1.

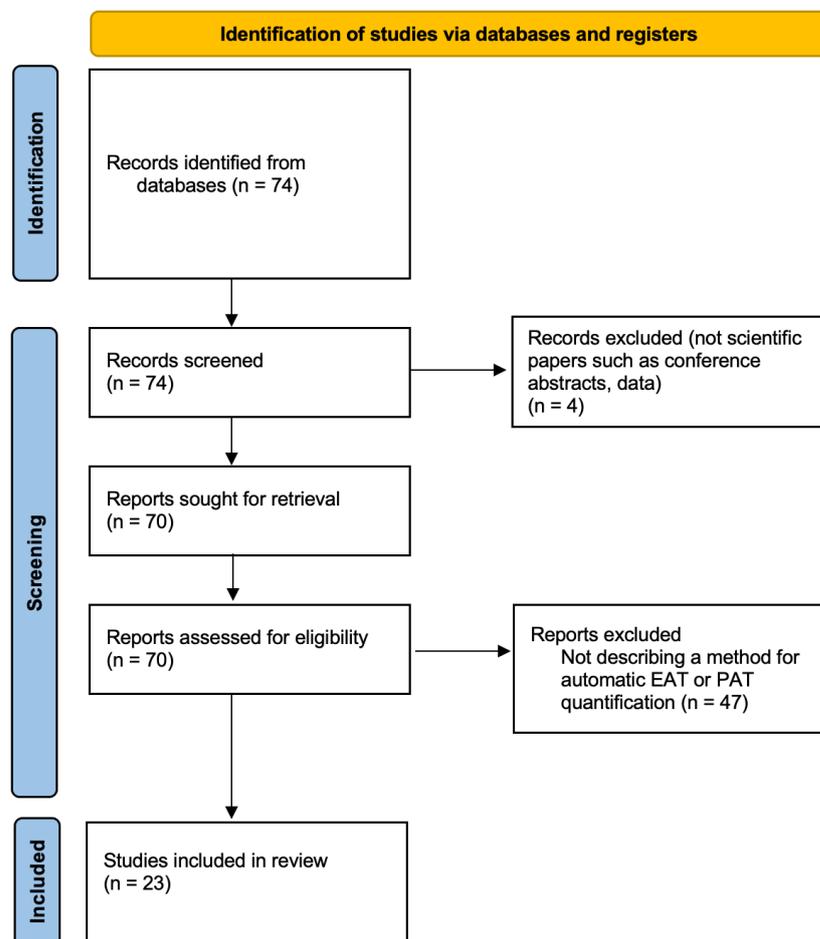
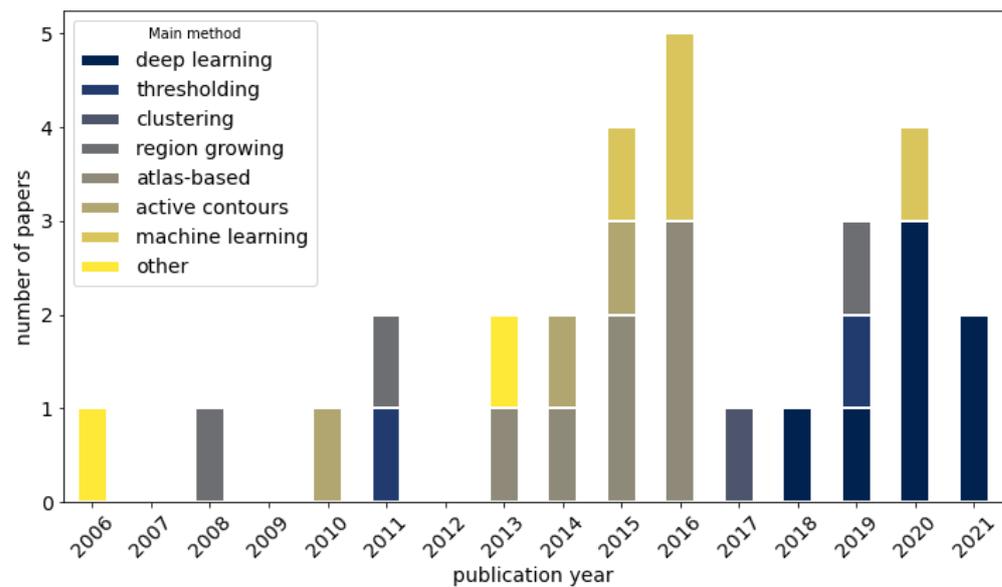


Figure 1. PRISMA flow diagram of our review methodology.

We review semi- and fully-automatic methods for direct EAT segmentation and quantification, EAT segmentation and quantification via pericardium delineation, and PAT segmentation and quantification. We further classify each category based on used image processing approaches into four groups: traditional image processing methods, atlas-based methods, machine learning-based methods, and deep learning-based methods.

The first computer-assisted methods for EAT and PAT segmentation start to appear around 2006. Until 2010 most methods rely on traditional image processing techniques like active contours and region growing. From 2010 onwards, applications of learned model parameters including atlas-based segmentation emerged as a new standard. From 2015 onwards, machine learning is commonly used in combination with traditional image processing methods. In recent years, deep-learning-based approaches started to appear, first by extracting pre-determined features from the images, and later directly using CT images as input. A relationship between the number of published papers and used methods over time is presented in Figure 2.



**Figure 2.** A summary of paper frequency and used methods.

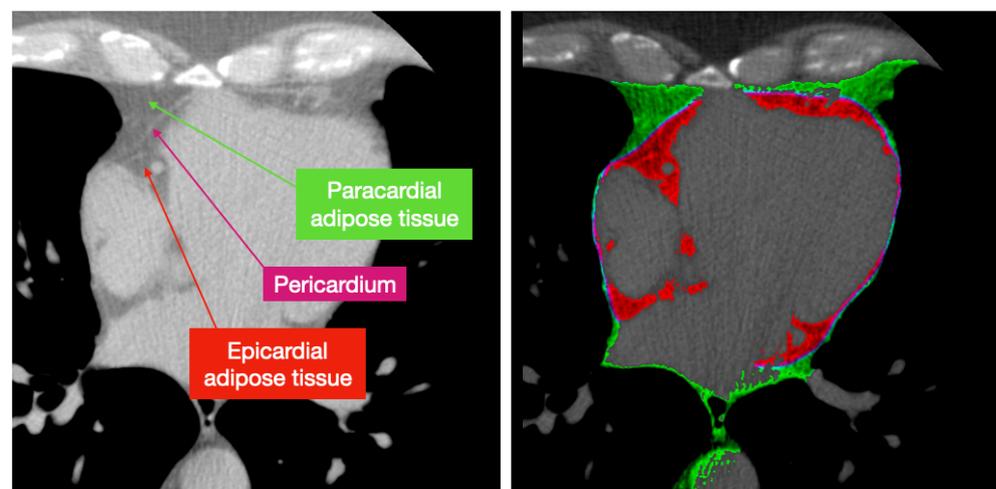
## 2. Clinical Background

In this section, we give a brief clinical overview of the heart wall structure and its surrounding adipose tissue. EAT, PAT, and other relevant anatomical structures are defined. We also provide a brief overview of how EAT and PAT interact with various CVDs.

Measuring EAT and PAT is important for understanding, assessing, and preventing various cardiovascular diseases. EAT is a metabolically active organ that shares the same micro-circulation with the myocardium, and there are potential interactions between them [2]. There is evidence EAT plays a direct role in causing coronary atherosclerosis and cardiomyopathy [6,10]. EAT thickness has been shown to correlate with metabolic syndrome [11] and coronary artery disease independently of obesity [2]. It relates in general with the progression of coronary artery calcification [12,13]. Additionally, the volume and density of EAT have been linked to major adverse cardiac events in asymptomatic subjects [14]. PAT volume is associated with atrial fibrillation [15]. EAT volume is also considered to be a good marker for cardiovascular diseases [3,4]. Additionally, EAT plays a role in visceral adiposity, heart morphology, liver enzymes, insulin resistance, fasting glucose, and is an accurate therapeutic target [16]. There is evidence that computer-assisted methods for quantifying EAT and PAT, especially on large datasets, could be used to predict major adverse cardiovascular events [17]. While there are still unanswered questions about EAT, there is evidence that EAT quantification could be used for risk stratification and risk management for atrial fibrillation as well as novel therapies for atrial arrhythmogenesis [18]. Epicardial fat volume can also be measured and provide diagnostic insights during PET/CT scans such as when assessing myocardial perfusion [19].

### 2.1. Heart Wall Structure and Terminology

The human heart is surrounded and protected by a sac of tissue called the pericardium. The pericardium consists of a thin outer layer of connective tissue called the fibrous pericardium. The fibrous pericardium envelops the serous pericardium, a tissue consisting of two layers. The outer layer of the serous pericardium, called the parietal pericardium, is fused to the fibrous pericardium. On the other hand, the inner-most layer of the pericardium called the visceral pericardium or epicardium [10] immediately surrounds the heart muscle (myocardium). These tissues can be seen in Figure 3.



**Figure 3.** The anatomy of the pericardium, EAT and paracardial adipose tissue on a CT scan [20].

Epicardial adipose tissue is most loosely defined as any adipose tissue within the pericardial sac. The fat immediately outside the pericardial sac is called paracardial fat, and together they form pericardial fat [6,21]. Epicardial fat is mostly concentrated in the atrioventricular and interventricular grooves and along the major branches of the coronary arteries and around the atria Sacks and Fain [10]. Paracardial fat is concentrated on the outside of the parietal pericardium within the mediastinum.

It should be noted that there are disagreements in the medical literature on the terminology and anatomical definitions of cardiac adipose tissues. Namely, pericardial fat is sometimes referred to as intrathoracic fat [6] and EAT and PAT are sometimes separated as two disjoint tissues and not a subset of each other [16,21]. It can be difficult to compare methods of EAT and PAT quantification since the used definitions can vary the quantified volume. In this review, we took that into account and analyzed the definitions used by each mentioned paper to make sure a valid comparison is made. In the rest of the paper we adopt the following definitions, according to the definitions most commonly used in the analyzed literature:

- Epicardial adipose tissue is all adipose tissue enclosed by the pericardium.
- Paracardial adipose tissue lies just outside the pericardium.
- The union of these tissues is pericardial adipose tissue.

Moreover, note that in this review, the acronym PAT is used to denote pericardial adipose tissue, not to be confused with paracardial adipose tissue.

## 2.2. Measuring EAT and PAT

The most common way to quantify EAT is to measure EAT thickness from echocardiography. EAT thickness is usually measured as the thickness of EAT perpendicular to the free wall of the right ventricle, usually end-systole (at the end of a heartbeat). EAT thickness is used as a middle-ground that provides better accuracy than surrogate markers like waist circumference but is more practical than CT or MRI in a clinical setting [22]. Still, estimating EAT volume from thickness at a single point can produce inaccurate results.

CT and MRI have greater sensitivity and measure 3D volumes, meaning they can more accurately measure EAT directly [22]. However, manual accurate quantification of EAT and PAT from CT or MRI is time-consuming and prone to inter-observer variability because of how distributed these tissues are around the heart. In literature, inter-observer variability for EAT volume measurement ranged in percentage from 7% to 14% [6,23,24] and in Dice scores from 0.89 to 0.91 [7,25,26]. Manually segmenting EAT from CT scans can last from a few minutes to an hour, depending on the number of slices in a scan and used software [9]. As another surrogate marker, Rodrigues et al. [27] shows that it is possible to use machine learning to predict the volume of paracardial adipose tissue volume from EAT

volume and vice-versa. Their method achieves good correlation ( $r = 0.9683$  for predicting EAT) but still requires measuring one of the tissues directly. However, this might reduce the need to measure both tissues, thus saving either time to manually segment the tissues or computing power in the case of automatic methods.

### 3. Image Processing Approaches Used for EAT and PAT Segmentation

This section presents a brief introductory overview of the image processing techniques commonly used to segment EAT and PAT and aims to give background for readers not familiar with computer-assisted medical image segmentation. We classify the papers analyzed in this review into four groups based on their primary approach: traditional image processing methods, atlas-based methods, machine learning, and deep learning. Traditional image processing methods include techniques like thresholding, region growing, and active contours.

#### 3.1. Traditional Image Processing Techniques

For this review, we consider techniques like thresholding, region growing, active contours, morphological operations, and other image processing techniques that don't include a learning component as traditional image processing methods. These techniques are often combined with using different heuristics to narrow down the search region, perform a rough initialization or guide more complex methods. Heuristics are "rules of thumb", relatively simple assumptions based on known examples and experience. For instance, a heuristic is that the femur can be found on an X-ray image as the longest component in the bone intensity range.

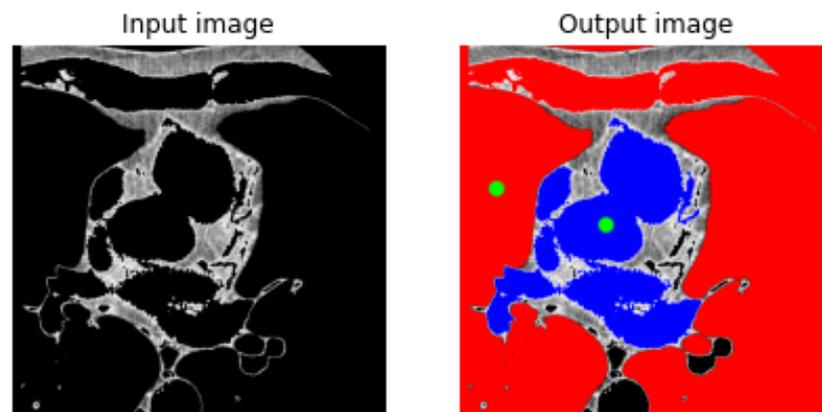
##### 3.1.1. Thresholding CT Images

The intensity values of CT scan's pixels are usually on a scale called the Hounsfield scale which is proportional to the attenuation of X-ray radiation of the tissue on that pixel [28]. Adipose tissue pixels have values between  $-250$  HU and  $-30$  HU [6]. This enables segmenting adipose tissue by thresholding the image, i.e., removing all pixels outside of a given range. Most papers mentioned in this review use thresholding at least in parts of their methods for segmenting EAT.

Thresholding, however, can segment all adipose tissue but fails to segment only EAT. Most papers mentioned in this review try to segment the pericardium in some way, and then use thresholding or other means to segment adipose tissue contained by the pericardium, which they deem EAT. As discussed in Section 2, this corresponds to a commonly used definition of EAT, as well as the definition used by this review.

##### 3.1.2. Region Growing

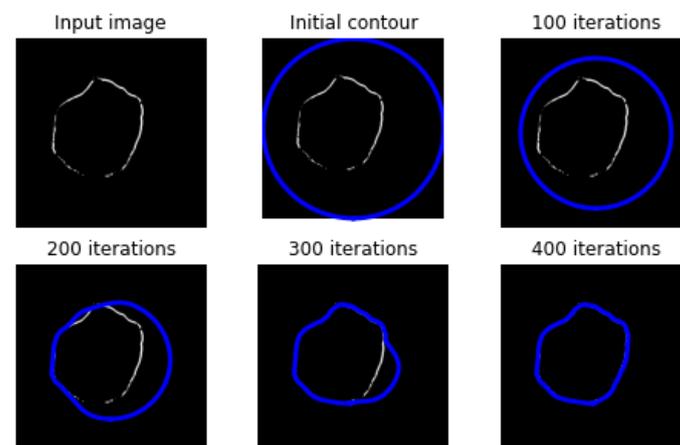
Region growing is a type of pixel-wise image segmentation method [29]. Region growing begins with a selected pixel called the seed point. At each iteration, the neighbors of the selected pixel are analyzed, and if they satisfy a given criterion, the region is enlarged to include them. The criterion is usually based on pixel intensities, textures, colors, or similar features. When compared to other methods presented in this review, region growing is a simple image segmentation method. It can achieve clear segmentation edges and good results. However, it is highly dependent on selecting a precise seed point, which can often be hard to do automatically in medical imaging tasks. Region growing works well when the regions have relatively sharp transitions but often struggles when the differences between regions are subtle, like the difference between similarly dense tissues. Furthermore, region growth is very sensitive to noise. These problems can be alleviated by heavily pre-processing and thresholding the images before the region grows, to increase contrast, accentuate edges, and remove irrelevant portions. An example of region growth is shown in Figure 4.



**Figure 4.** An example of using region growing to segment the inside and outside regions of the heart on a CT slice of the heart, thresholded to the adipose tissue range. The image on the (left) is the input image, while the image on the (right) shows the segmentation results of pixels outside the heart in red and inside the heart in blue. The green points on the image are two seed points manually selected to start the region growing process.

### 3.1.3. Active Contours

Active contours or “snakes” are an image segmentation method that uses curves to segment part of an image [30]. A wide and rough curve is first initialized around the target object. Then, during each iteration, the active contours algorithm tightens the curve around the object until the curve follows the object’s shape. The algorithm does this by minimizing a cost function, usually called the energy function. The energy function assigns a numerical value to each curve. The smoother the curve and the closer an edge, the lower its energy. By minimizing this function, the algorithm finds a curve that smoothly follows the edges of an object. An example of using active contours is shown in Figure 5.

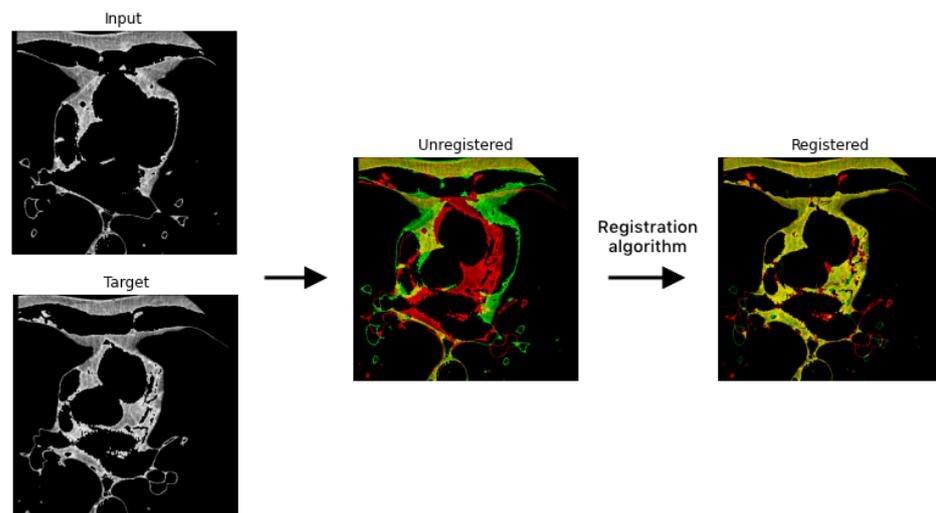


**Figure 5.** An example of using active contours to fill missing portions of a pericardium line obtained by thresholding a CT scan, shown in white. The contour is shown in blue, initially a perfect circle around the image. In each iteration, the contour is deformed to better fit the image.

Active contours often require user interaction to initialize the curve or prior knowledge about the shape they are segmenting. They can track dynamic objects and follow contours that change from image to image. However, they often ignore very fine features of the contour and are sensitive to local minima that might draw the contour away from the target object.

### 3.2. Atlas-Based Methods

While active contours rely on geometrical constraints to segment shapes, atlas-based segmentation methods take into account the spatial relationship between several structures [31]. This is done by forming an atlas, a map of structures found in an image. Atlases are usually formed by an expert who manually segments an image, labeling different structures on the image. In medical imaging, the specific locations and shapes of anatomical structures vary from patient to patient, so multiple atlas images are usually collected and fused into one atlas image. Once an atlas is acquired, it can be used to segment new input images that were not manually segmented. This is achieved using a registration algorithm. In general, registration deforms and transforms an image so that the structures within the image map onto the same structures on a different image. Atlas-based registration treats the atlas as an image and maps the yet unseen image onto the atlas. The correct way to map the image is determined by a similarity measure that measures how well the shapes and features of one image match the other. Registration is often done in a non-rigid way with the ability to deform the image to account for shape irregularities in different input images. Once an image is mapped onto the atlas, structures can easily be segmented from the image by reading the atlas value for the corresponding mapped image pixel. An example of using a registration algorithm to deform an image is shown in Figure 6.



**Figure 6.** A diagram of the registration process. First, input and target images are selected. On the Unregistered image, both the target and input images are superimposed, with the input image in green and the target image in red. During registration, the input image is deformed to match the target image, while the target stays fixed. The Registered image shows the two images superimposed, but this time it shows the deformed input image as a result of the registration algorithm.

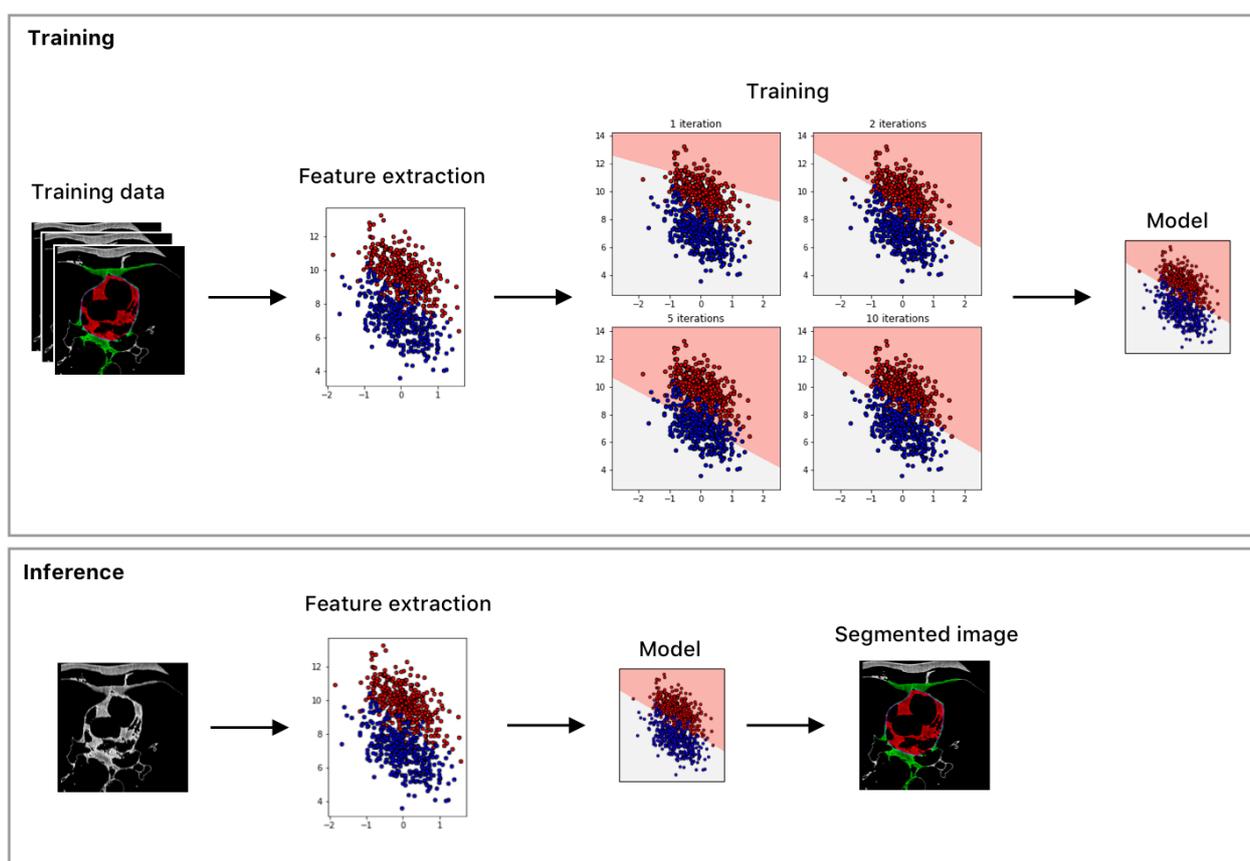
Atlas-based methods offer robust segmentation and consider the spatial relationships of different structures, making them a good fit for medical imaging [32]. On a medical image, the same tissue usually has the same grey levels. However, it can be located in different body areas, so a segmentation method that can be sensitive to location and grey levels can provide good results. The downside of atlas-based methods is that they require selecting and acquiring accurate atlas images via manual image segmentation. The collected atlas can be an inadequate representation of real-world data, as images can differ based on patient backgrounds, genetics, medical conditions, and used equipment. Additionally, the registration process can fail due to noise or significant differences in the input image and the atlas and thus entirely prevent further segmentation.

### 3.3. Machine Learning

Some methods for image classification rely on extracting pre-determined features from images. This vector of image features is fed into a machine learning model that

classifies each pixel as part of a class. In the case of EAT segmentation, the pixels are classified as either EAT or background. The machine learning model is a statistical model that maps the input features into the output class, but the exact parameters of the model are not determined by hand [33]. Instead, the classifier parameters are learned through a process called machine learning. The model learns how to correctly classify the inputs either by learning inherent statistical relationships in the inputs (unsupervised learning) or by being exposed to manually classified examples and extrapolating based on them (supervised learning).

In the case of supervised image segmentation, manually-determined features are extracted from the images and fed into the model. To train the model, inputs for which a manual classification is already determined are used. The model's output is then compared to the known ground truth, and the model tunes its parameters in such a way as to output the expected ground-truth value for that input. This is then repeated for all of the pre-classified inputs. The process of tuning the model's parameters in this way is called training, and the data with already classified ground truth is called training data. An example of a supervised machine learning process for segmenting images is shown in Figure 7.



**Figure 7.** A diagram of the machine learning process using a supervised linear classifier. The top portion of the diagram shows the training process. The features are colored based on their known class from the training data, in red and blue. The parameter of the line separating the areas of the two classes (shown in light red and grey) is learned during the training process. The bottom portion of the figure shows the inference process, where the same features are extracted and the trained model is used to classify each pixel of the image. In the segmented image, the red area is EAT while the green area is paracardial adipose tissue.

The specific features that are used as inputs for machine learning models are selected by the model designer based on how well they describe different aspects of the segmented structures. In the case of images, these are usually the pixel values and locations, mean

pixel values in a neighborhood, geometric moments, and other higher-level descriptions of image features. The model can then determine, for instance, that having a CT pixel around the center of the image with a value of  $-150$  HU has a high likelihood of being EAT, while a pixel of  $30$  HU and on the edge of the scan is less likely to be EAT.

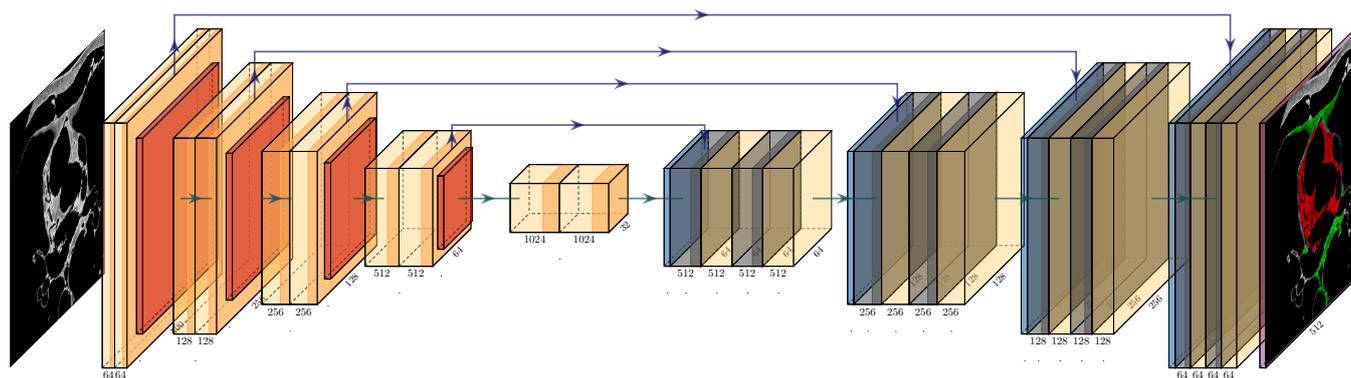
There are also different methods of parameter optimization via machine learning. One such method used for pericardium delineation is genetic algorithms. In general, genetic algorithms are a method for parameter optimization that simulates an evolutionary process. Initially, parameters are randomly determined. Then, the parameters are randomly changed (i.e., mutated) and evaluated using a fitness function, that evaluates the parameter's performance on some task. Mutations that lead to an increase in fitness are carried over to the next iteration when the parameters are mutated again. This process is repeated until the fitness function reaches an optimal value.

### 3.4. Deep Learning

Deep learning extends the concepts of machine learning. In machine learning, there is a known statistical model that learns its parameters from the training data. In deep learning, the models themselves are learned from the training data. To achieve this, deep learning methods use artificial neural networks (ANNs) to perform tasks [34]. ANNs are a type of computing system that consists of multiple connected artificial neurons. Each neuron is a mathematical function that transforms one or more input matrices into an output matrix by summing all the input matrices, multiplied by their respective weights, and then feeding the sum into an activation function, whose job is usually to add an element of non-linearity. Deep learning ANNs have a layered structure, starting from the input layer, which receives the input data. Results from the input layer's neurons are then passed through a series of one or more hidden layers. Finally, the results are passed to an output layer, which offers the prediction for the task at hand. The quality and shape of the output are determined by the network architecture, i.e., the number of neurons, how they are connected, and which activation functions are used.

The most common types of ANNs used for medical image processing are convolutional neural networks (CNNs). CNNs are ANNs that use the mathematical operation of convolution instead of summation in one or more of their neurons. Usually, they contain several layers of this type of neuron called convolutional layers. CNNs are particularly good at working with image data. In CNNs, the first layers of the network generally learn to recognize rudimentary features like edges and shapes. In later layers, the earlier features are combined to form more complex features, while in the deepest layers, the network can often detect very specific shapes like the wheels of a car or eyes on a face.

A common CNN architecture used in medical imaging is U-Net [35], shown to be a state-of-the-art method in a variety of tasks. Variations of U-Net are commonly used for the task of EAT segmentation [36–38]. The U-Net architecture consists of a cascade of layers that progressively downsample the image, called the contracting path, followed by another cascade of layers called the expanding path that upsamples the image back into the output image that contains the segmented portions. Outputs of layers of the contracting path are connected directly to the matching layers of the expanding path. This allows the network to concurrently use the global context from the contracting path with complex local features of the expanding path to achieve good segmentation accuracy. The U-Net architecture is shown in Figure 8.



**Figure 8.** A diagram of the U-Net CNN architecture [35]. The left side of the diagram shows the contracting path, while the right side of the diagram shows the expanding path. The blue arrows represent crop and concatenation operations, while the gray arrows represent convolutional and pooling operations. In the final prediction, the red area is EAT while the green area is paracardial adipose tissue.

In terms of segmentation time, neural networks are usually an order of magnitude faster than more traditional image processing methods and can sometimes achieve much better accuracy [39]. However, deep neural networks require a large amount of training data, which can be challenging, especially in medical imaging tasks. Additionally, neural networks are hard to interpret. Since their models' parameters are learned from the data and not programmed by hand, the way they work and their specific limitations are not known ahead of time and are hard to test without access to a large amount of testing data.

#### 4. EAT Segmentation Methods via Pericardium Delineation

Depending on the method, delineating the pericardium can be an easier task than segmenting EAT outright. EAT and PAT are both unevenly distributed around the heart, leading to wide variation from slice to slice and patient to patient in CT and MRI images. They also have an uneven shape that is hard to easily describe in an algorithm. However, on CT and MRI images the pericardium is a smooth, thin, almost elliptical contour. Methods like active contours or ellipse fitting are naturally suited for segmenting this kind of shape. Once the pericardium is segmented, EAT can be more easily segmented as all adipose tissue within the pericardium.

The biggest challenge with segmenting the epicardium is how thin it is. When collecting CT scans for CAC scoring, slice thickness is usually set to around 2–3 mm [40]. The pericardium is normally less than 2 mm thick, so the pericardium can often be blurred or faint on CT images due to partial volume averaging, especially because the heart is a constantly moving organ.

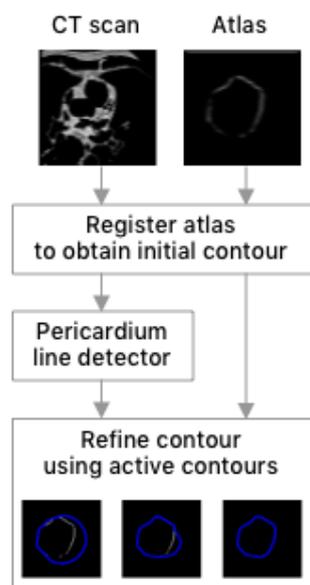
There are several methods proposed for pericardium delineation. Some are methods presented purely as pericardium delineation methods, while others are part of methods to segment and quantify EAT. In this section, we present papers that describe pericardium delineation or are methods to segment EAT by first segmenting the pericardium.

##### 4.1. Traditional Image Processing Methods

Methods in this category rely on various heuristics and thresholding to identify a volume of interest (VoI), i.e., the parts of the CT scan that are likely to contain epicardial fat. Discarding known irrelevant parts of the image not only speeds up the segmentation process but also lowers segmentation errors since other parts of the image can't interfere with the segmentation process. Once the VoI is determined, active contours are used to delineate the pericardium line. EAT is then segmented and quantified by thresholding out adipose tissue inside the contour.

An example of such an approach is Coppini [41], one of the first papers to describe a semi-automatic method for EAT quantification. Their method first discards the lungs, and then an expert observer manually selects control points on the pericardium on each relevant slice, which is then used to define the VoI. The next step of their method is to threshold the VoI to the Hounsfield scale interval corresponding to adipose tissue. The exact values of the threshold can be influenced manually by the observer. The thresholded image is sparse and has oscillations in its contours. To further refine the shape, the authors use geodesic active contours, an active contours variant to find a smooth contour of the sparse thresholded shape on each slice of the CT image.

This approach can also be fully automatic as exemplified by Ding et al. [42]. A summary of their method is shown in Figure 9. While Coppini [41] initializes the pericardium contour manually using an expert observer, Ding et al. [42] initializes the contour automatically using an atlas-based method for pericardium detection. This initialized contour is then refined using geodesic active contours to more accurately detect the pericardium. The authors report a Dice coefficient between 0.88 and 0.95 on 15 patients and a time of 60 s to perform a segmentation of a scan. This method is further validated in Ding et al. [24] on 50 patients, with a reported Dice coefficient of 0.92.



**Figure 9.** A summary of the method presented in Ding et al. [42]. A CT scan is first registered to obtain an initial counter and a line detector is used to find the pericardium line. The detected line is refined using active contours.

#### 4.2. Atlas-Based Methods

Atlas-based methods can be beneficial for the task of pericardium delineation since the pericardium line can be sparse and fuzzy, in some cases leading to complete failures of methods such as active contours. Atlas-based segmentation can be more robust to these cases since they rely on more complex features than active contours. A representative fully-automatic method for EAT classification, presented in Shahzad et al. [25], employs multi-atlas segmentation to segment the pericardium on non-contrast CT images. It is much easier to see the pericardium on contrast-CT images, so the atlases were collected by manually segmenting the contrast-CT scans of 8 patients. These atlases were then registered onto the non-contrast input images, and the registrations are then fused to produce a final pericardium segmentation. Once the pericardium is segmented, the total EAT segmentation and quantification are done by thresholding the segmented portion of the image followed by a connected component analysis step to remove noise from the EAT volume. The authors

reported Dice coefficients of 89.1 and 89.2 on the remaining scans compared to manual segmentations of two experts on 98 patients' scans.

#### 4.3. Machine Learning-Based Methods

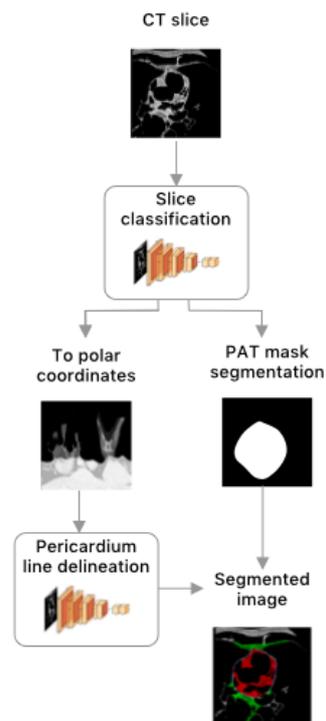
Machine learning can make the problem of finding the pericardium more complex and increase the data requirements but can lead to better and much faster results in terms of processing time. However, these results come at the cost of having to engineer and select appropriate features for training a model. For instance, Norlén et al. [26] uses the mean values, variance, and derivatives of the pixel intensities of a region around a voxel as features to train a random forest classifier. The classifier classifies image pixels into one of three classes: just inside the pericardium, just outside the pericardium, and background. Another important drawback of machine learning is that it requires heavy preprocessing of the input images. Norlén et al. [26] use feature-based multi-atlas registration to initialize a pericardium localization before the image is processed by the machine learning model.

Similarly, Rodrigues et al. [43] segment the paracardial and pericardial adipose tissues using a different machine learning algorithm [20] as a preprocessing step. Then, to segment an elliptical pericardium, the authors use a genetic algorithm to fit an ellipse to the edges between paracardial and epicardial tissue. The algorithm can successfully find an ellipse that delineates EAT and paracardial adipose tissue and that closely follows the pericardium contour. Note that the method requires an already segmented image, and is mainly aimed to be used as a performance optimization for other EAT segmentation methods that delineate the pericardium. By fitting the pericardium ellipse on a few slices of the CT scan, they can interpret the pericardium contour on the rest of the slices, thus saving processing time.

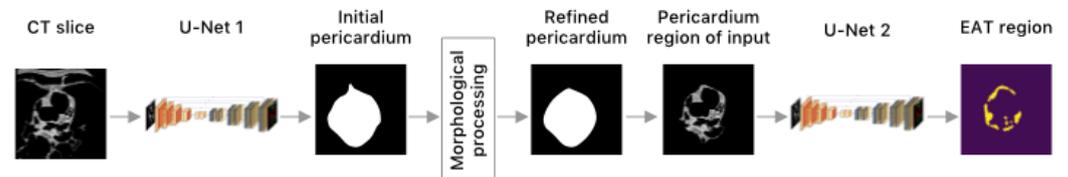
#### 4.4. Deep Learning-Based Methods

Deep learning removes the need to select appropriate features for training, and with CNN-based architectures, the network can utilize very complex high- and low-level features to segment the pericardium. One of the first applications of CNNs to EAT segmentation is proposed in Commandeur et al. [7]. A summary of their method is shown in Figure 10. The authors developed a multi-task framework consisting of two convolutional neural networks. The first CNN performs three separate tasks. The first task is classifying whether or not the given slice contains the heart, which enables stopping the process early and saves performance costs. The second task segments the union of epicardial and paracardial structures, while the third task segments epicardial and paracardial areas separately. The epicardial mask from the first network is used to re-sample the input CT scan, which is then fed into the second CNN, which delineates the pericardium line given a CT scan in radial coordinates. Total EAT volume is obtained by thresholding the segmented epicardial area to the range of adipose tissue. The authors report a Dice coefficient of 0.81 for EAT segmentation, and a time of under 20 s to process one CT scan. Their method is validated on a cohort of 250 patients. The authors also report an inter-observer Dice coefficient of 0.89 on 30 randomly selected patients. This indicates that their EAT segmentation does not perform within the limits of inter-observer variability, however, it is an order of magnitude faster.

More recent approaches rely on comparatively simpler architectures. Instead of using multi-task learning, the methods most commonly rely on training a U-Net-based model to segment the pericardium from a slightly processed input image. For instance, Zhang et al. [36] uses a U-Net-based structure by connecting two cascading U-Nets. A summary of their method is presented in Figure 11. Their first network detects the pericardium from an input CT slice. The output pericardium segmentation is coarse and noisy, so the authors process the output in a morphological processing layer using morphological reconstruction, opening, and closing. This segmented shape is then used as a mask for the inputs to the second U-Net network that performs EAT segmentation. The authors report a mean Dice coefficient of 0.91 on 10 patients taken from Rodrigues et al. [20].



**Figure 10.** A summary of the method presented in Commandeur et al. [7]. A CNN is first used to obtain a PAT mask, which is used to resample the input CT slice. The slice, converted to radial coordinates, is fed into a second CNN which delineates the pericardium line. The pericardium line and PAT mask are used to segment the image into EAT and PAT.



**Figure 11.** A summary of the method presented in Zhang et al. [36]. The CT slice is fed into a U-Net-based CNN which segments the pericardium mask. The mask is refined using morphological processing and used to mask the input for a second U-Net based CNN which segments the EAT region.

Even simpler similar approaches exist where the final EAT segmentation is obtained by thresholding the pericardium region and not using a deep learning model, such as Benčević et al. [44]. The input to the network has two channels, the original CT slice and another channel where each pixel has the same value of the slice depth of that slice. This gives the convolutional neural network access to some depth information of each slice, thus improving performance without requiring a fully 3D neural network. The authors report a mean Dice coefficient of 0.86 using 2-fold cross-validation on 20 patients from Rodrigues et al. [20].

## 5. Direct EAT Segmentation Methods

EAT is usually distributed around the heart and has a rough, discontinuous shape, especially near the myocardium. This makes direct segmentation more challenging than segmentation via pericardium delineation. This section provides a summary and commentary of proposed methods that segment EAT directly.

### 5.1. Traditional Image Processing Methods

Due to EAT's particular distribution, using traditional image processing methods to directly segment EAT automatically is quite challenging. Therefore, authors rely on

semi-automatic approaches such as Militello et al. [9]. The authors propose a time-saving GUI tool to enable experts to quickly segment EAT while reducing inter-observer variability. The experts manually determine the first and last slice of the heart region and then proceed to manually segment a region of interest on every 10th or 20th slice. The VoI of the remaining slices is then interpolated by the algorithm, with the ability of manual refinement if required. The algorithm further refines the segmentation with additional image processing steps like hole filling and morphological closing. The authors evaluate the method on 145 patients and a comparatively large number of slices (around 25,000 total slices), by slice count the largest datasets out of all the papers covered by this review. The authors report processing times of around 3.5 min for 60-slice scans and around 7.4 min for 235-slice scans, in both cases 4–8 times faster than manual segmentation. The method achieves a Dice coefficient of around 0.93. Considering the quantity of the used dataset, it is safe to say their method produces results equal to or better than inter-observer variability for this task, but still requires significantly more time than fully automatic methods. One potential improvement of the method would be to use a fully automatic method to perform the initial segmentation but allow experts to manually refine that automatic segmentation, thus saving even more time.

### 5.2. Machine Learning-Based Methods

There is a notable trend in using machine learning to fully automatically segment EAT directly. The most common approach is to heavily preprocess the images and then extract various features to feed into a random forest classifier or a neural network. An example of such an approach is present in Rodrigues et al. [45]. The input images are thresholded to the adipose tissue range and then translated to a standardized position with atlas-based segmentation of the retrosternal area. A selection of features is then extracted from the image. These features include the center of gravity of the image, x and y positions, and the gray level of the pixel, as well as additional features from the surrounding neighborhood of each pixel. Using these features as inputs, the authors train a machine learning model based on the random forest classifier to segment each pixel into one of three classes: EAT, paracardial fat, or pericardium. For EAT segmentation, the authors report a Dice score of 0.979 using the dataset from Rodrigues et al. [20]. Similar approaches can be seen in Rodrigues et al. [20] and Kazemi et al. [46].

Kazemi et al. [46] use contrast limited adaptive histogram equalization in addition to thresholding to preprocess the images. Their features also include a Gabor filter bank, a collection of Gabor filters at different scales and rotations. Instead of relying on only a random forest classifier, the features are used to train an ensemble of different classifiers. They achieve a Dice coefficient of 0.94, which indicates that selected features are less salient than those selected in Rodrigues et al. [20,45].

Other works use different types of optimization algorithms to segment the image. For instance, in Priya and Sudha [47], EAT and paracardial adipose tissue are segmented using region growing, and the regions are then merged with the help of fruitfly-based optimization. The authors report a Dice coefficient of 0.987 for EAT segmentation, slightly larger than Rodrigues et al. [20] on the same dataset. Another example of such an approach is presented in Zlokolica et al. [48]. The fuzzy c-means clustering to segment EAT from CT images. The method first segments a VoI of the heart on each CT slice using morphological operators and thresholding. Each slice then goes through a clustering step with a predefined number of clusters, where the features used to determine the clusters are the mean luminance and luminance variance of a patch of pixels. The clustering begins from a user-specified seed point placed on a single slice of the patient's CT scan. The resulting cluster of EAT is then post-processed by fitting an ellipse on the cluster and discarding the pixels outside the ellipse. The authors report Dice coefficients from 0.6 to 0.82 on 8 patients. 2 out of the 10 patients show unacceptable error rates of over 15% and were not included in the Dice coefficient measurements.

From the reported results it would seem that machine learning-based direct EAT segmentation performs very well, however it is not clear how representative these results are of real-world performance. The testing data in [45] consists of a random sampling of 34% of the slices of 16 patients (4 patients were discarded to reduce training times). This means that their model was trained and evaluated on different slices of the same patient, allowing potential overfitting of a patient's anatomy. It is not clear if [20,46] use the same validation approach. Ref. [47] also does not clarify if the authors used a separate training and testing dataset to evaluate the classification performance, or how many slices and patients the whole process was evaluated on.

### 5.3. Deep Learning-Based Methods

Similar to machine learning, deep learning is fitting for direct EAT segmentation since it can find very complex shapes in the image. However, deep learning does not require feature selection and reduces the need for heavy image preprocessing. Deep learning models for direct EAT segmentation are most commonly based on U-Net. For instance, Li et al. [38] used a U-Net-based network to segment EAT and paracardial fat. The U-Net is modified to include an additional pyramid pooling layer. The authors use the same dataset used in Rodrigues et al. [20] but apply data augmentation to artificially increase the amount of data. EAT and background classes are imbalanced, i.e., the number of EAT pixels is small compared to the number of background pixels. This imbalance can cause problems when training CNNs. To get around this problem, the authors use focal loss, which increases the weight of getting the less-populated class wrong. The authors report an area under the ROC curve (AUC) of 0.87 for EAT. The authors do not report a Dice coefficient so it is hard to determine the accuracy of their method, especially since AUC can misrepresent the effectiveness. As opposed to AUC, the Dice coefficient does not consider true negatives (the background class), which, in this case, represents a more accurate measurement than AUC.

A similar approach is presented in Benčević et al. [49], where the authors preprocess the input CT scans using the polar transform. The authors first train a stacked hourglass neural network to predict the optimal origin for the polar transform and then use that network to transform each input image. They then train a second U-Net-based segmentation network on the polar-transformed images. They achieve a relatively low Dice score of 0.78. This result is better than a U-Net baseline on the same dataset but worse than other neural network-based methods [36,37]. However, the authors comment that the results could be further improved by not segmenting EAT directly but instead first segmenting the pericardium. The results indicate that researching the polar transform and other forms of pre-processing could lead to a notable improvement in EAT segmentation results.

Most direct EAT segmentation methods work on 2D images, with the notable exception of He et al. [37]. They use the U-Net architecture on 3D patches of CT images. The model follows a modified U-Net architecture called deep attention U-Net, which includes attention gates. Since the U-Net has both coarse and fine-grained information, it can be modified to include attention gates that use coarse-grained features to suppress non-salient connections in the network, thus removing noisy data from the training process. The authors report a Dice coefficient of 0.85, cross-validated on 40 patients. When compared to results of direct 2D segmentation with a simple U-Net from Zhang et al. [36], the results from He et al. [37] achieve a better Dice coefficient, which indicates that training on 3D data could be beneficial. After all, EAT is a 3D structure so there could be a lot of mutual information on EAT between two consecutive slices.

## 6. PAT Segmentation Methods

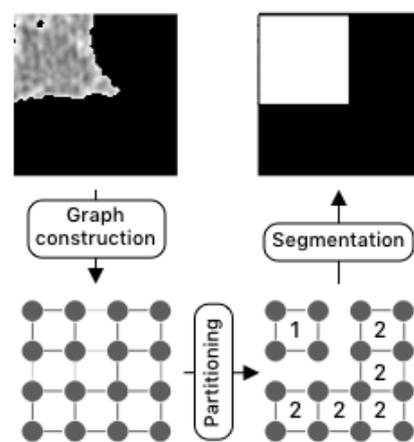
PAT segmentation methods usually rely on traditional image processing methods like thresholding and region growing, using various pre-programmed heuristics to identify common structures and segment PAT. More recent methods incorporate atlas-based segmentation. While PAT is a superset of EAT, its relatively smaller clinical importance

has resulted in a smaller number of published papers, compared to the number of papers describing EAT segmentation.

### 6.1. Traditional Image Processing Methods

Semi-automatic methods for PAT segmentation rely on a series of heuristics and thresholding based on manually selected limits. For instance, Dey et al. [50] proposes a method where an expert observer first selects the slices which limit the VoI to the heart. The algorithm then uses adaptive thresholding, region growing, and heuristics to segment out the lungs and other irrelevant structures and determine the heart region. From the heart region, fat voxels are calculated by thresholding. Once the limiting slices are manually selected, their algorithm can perform measurements within 20 s, compared to 9 to 15 min for manual segmentation. The authors report an accuracy of 94% validated on 105 patients, in line with inter-observer variability for PAT segmentation. However, in 8% of cases, the algorithm failed to segment PAT.

These methods can be made fully automatic by using atlas-based segmentation in place of manual initialization, as presented in Ding et al. [51]. Their method works by first segmenting the heart using atlas-based segmentation. Once the heart is segmented, an input image thresholded to the adipose tissue range is run through a graph segmentation algorithm. An overview of graph-based segmentation is presented in Figure 12. The image is represented as a 3D graph, where the weight on each edge of the graph adaptively measures the similarity between two neighboring voxels, using a similarity function that adapts to local features in the image. By finding low-weighted graph edges, the image can be split into segments. Using these segments and the segmented heart region, the method determines PAT on the image. The authors report a Dice coefficient of 0.82 evaluated on 6 patients. While the result is not indicative of real-world performance due to only evaluating 6 patients, it shows that it is possible to automatically segment PAT on MRI images.



**Figure 12.** An overview of graph-based segmentation. First, a graph is segmented from the image such that the vertices are a sampling of the points on the image, and the edges are weighted based on the similarity of the two points they connect. In the figure, darker edges are weighted more heavily. The graph is then partitioned into connected components using some criterion that discards low-weighted edges. Finally, the partitioned graph is used to segment the original image.

### 6.2. Machine Learning-Based Methods

The first method for segmenting PAT in the literature is based on fuzzy connectedness, presented by Bandekar et al. [52]. Their method uses a series of heuristics, thresholding, and sampling to segment the heart, lungs, and the rest of the torso on a CT scan. PAT is then segmented using fuzzy connectedness (FC) segmentation. FC finds several objects in an image starting from a specified seed point and, using a fuzzy affinity function, assigns a numerical value  $[0, 1]$  to each pixel. The authors compute the distributions of fat, non-fat,

and background classes dynamically using a training data set. This information is then used in FC segmentation to segment new images using the Mahalanobis similarity metric [53] as an affinity function. The authors evaluate their method on 23 patients and report a mean accuracy of 99.13%. However, the mean true positive rate is 85.63%, indicating that PAT segmentation is worse than overall accuracy. Their method requires less than 1 min to segment a single CT slice.

### 6.3. Deep Learning-Based Methods

Deep learning methods for segmenting PAT are similar to ones used for direct EAT segmentation. For instance, He et al. [54] use a 3D-based deep attention U-Net model as in He et al. [37], trained on CT images with labeled PAT and myocardium regions. The authors train and evaluate their model on 422 cardiac CT scans from six different centers, to our knowledge the largest cohort used to evaluate EAT and PAT segmentation methods. The authors experiment with training on patients with known CVDs and evaluating patients without CVDs, and vice-versa. Evaluation is done with leave-one-center-out cross-validation. The authors report a Dice coefficient higher than 0.88 for PAT segmentation, and a correlation of 0.99. The model is compared to a 3D U-Net model. For PAT segmentation, the model outperforms 3D U-Net in terms of the Dice coefficient by 0.02. The experiments show the potential for using the same deep-learning-based methods for both EAT and PAT segmentation, as they achieve similar results using the same model for both tasks.

## 7. Discussion

A tabular overview of PAT segmentation and quantification is shown in Table 1, while an overview of EAT segmentation and quantification methods is shown in Table 2. We have presented a wide range of methods for EAT and PAT segmentation. More traditional image processing methods like active contours and region growing can offer good, reproducible segmentation results but are often sensitive to noise or require manual intervention and verification. Machine learning and deep learning methods can provide fully automatic segmentation with robust segmentation performance but are often hard to interpret and fail in unforeseen ways. There is a comparatively small number of papers published on the subject of EAT and PAT segmentation when compared to other cardiac segmentation tasks [55].

In literature, EAT segmentation inter-observer variability ranged in percentage from 7% to 14% [6,23,24] and in Dice scores from 0.89 to 0.91 [7,25,26]. To be considered valuable for use in both a clinical setting and in studies, an automatic method for EAT quantification should achieve results better or at least the same as the inter-observer variability. The lack of annotated data, especially large datasets, means that it is hard to objectively compare different methods, and one can not simply look at Dice scores to determine how accurate a method is. In this analysis, we took into account not only reported results but also the number of patients the method was validated on, as well as other potential ways the results might not be representative. The vast majority of papers attempt to segment EAT and PAT on CT images, with the notable exception of Ding et al. [51], who use MRI images.

**Table 1.** Pericardial adipose tissue segmentation methods. DSC is the reported Dice coefficient. Segmentation time is the reported time to segment a single CT scan, unless specified otherwise.

Reference	Automatic	Method	Dataset	Segm. Time	Result
He et al. [54]	fully	deep learning	CT; 422 patients	-	DSC = $0.94 \pm 0.03$
Ding et al. [51]	fully	3D, atlas-based, graph-based	MRI; 6 patients	-	DSC = $0.82 \pm 0.06$
Dey et al. [50]	semi	region growing	CT; 105 patients	>20 s	acc. $94\% \pm 6\%$
Bandekar et al. [52]	fully	fuzzy connectedness	CT; 23 patients	1 min/slice	acc. $99\% \pm 0.38\%$

**Table 2.** Epicardial adipose tissue segmentation methods. DSC is the reported Dice coefficient,  $r$  is the Pearson correlation coefficient, AUC is the area under curve measurement and  $\rho$  is the Spearman correlation coefficient. Segmentation time is the reported time to segment a single CT scan.

	Reference	Automatic	Dataset	Segm. Time	Result
deep learning	Benčević et al. [49]	fully	CT; 20 patients 878 slices [20]	-	DSC = 0.78
	Benčević et al. [44]	fully	CT; 20 patients 878 slices [20]	-	DSC = 0.86
	He et al. [37]	fully	40 patients	-	DSC = $0.85 \pm 0.05$
	Zhang et al. [36]	fully	CT; 20 patients 878 slices [20]	-	DSC = $0.91 \pm 0.01$ $r = 0.93$
	Li et al. [38]	fully	CT; 20 patients 878 slices [20]	-	AUC = 0.87
	Commandeur et al. [7]	fully	CT; 250 patients ~13,750 slices	3–20 s	DSC = 0.82 IQR 0.78–0.86 $r = 0.97$
machine learning	Kazemi et al. [46]	semi	CT; 20 patients 878 slices [20]	-	DSC = 0.94
	Priya and Sudha [47]	fully	CTA; 30 patients	52 s	DSC = 0.90 $r = 0.98$
	Zlokolica et al. [48]	semi	CT; 10 patients	-	DSC = $0.69 \pm 0.08$
	Norlén et al. [26]	fully	CT; 30 patients	51.9 s	DSC = $0.91 \pm 0.04$ $r = 0.99$
	Spearman et al. [56]	semi	CT; 70 patients ~31,500 slices	133 s	$\rho = 0.97$
	Rodrigues et al. [20]	fully	CT; 20 patients 878 slices [20]	1.8 h	DSC = 0.97
	Rodrigues et al. [45]	fully	CT; 20 patients 878 slices [20]	24 h	DSC = 0.98
traditional methods	Militello et al. [9]	semi	CT; 145 patients ~24,150 slices	3.5–14.5 min	DSC = 0.93 $r = 0.95$
	Ding et al. [24]	fully	CT; 50 patients	60 s	DSC = 0.92 (0.88–0.95) $r = 0.97$
	Ding et al. [42]	fully	CT; 15 patients ~825 slices	60 s	DSC = 0.93 (0.88–0.95) $r = 0.96$
	Shahzad et al. [25]	fully	CT; 98 patients ~6272 slices	-	DSC = $0.89 \pm 0.03$ $r = 0.91$
	Barbosa et al. [23]	semi	CT; 10 patients	-	-
	Coppini [41]	semi	CT; 10 patients	-	-

### 7.1. Metrics Used to Evaluate Segmentation and Quantification

The papers mentioned in this review use several evaluation methods and metrics to present their segmentation results including (among others): pixel-level accuracy, sensitivity, and specificity; the Sørensen–Dice coefficient (Dice coefficient, DSC); the Jaccard index (Intersection over Union, IoU); and the area under the receiver operating characteristic curve (AUC).

However, some of these metrics are better than others at showing segmentation performance. As an example, consider the pixel-level accuracy. The pixel-level accuracy measurement is usually defined as in Equation (1).

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

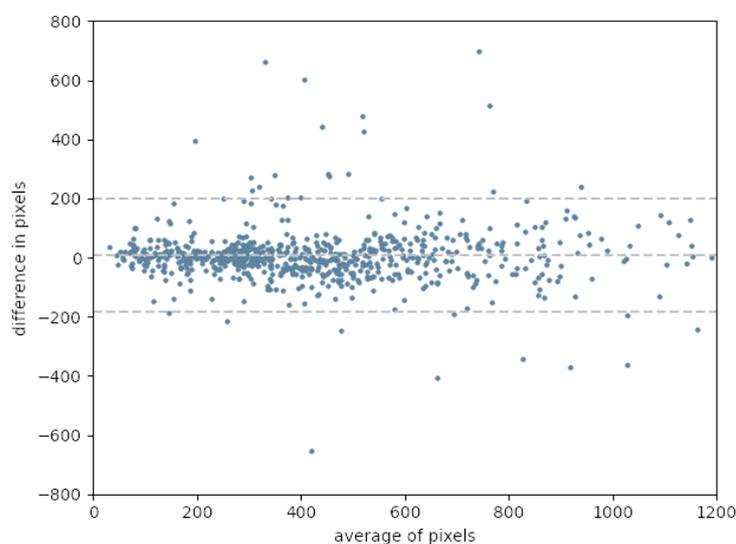
where  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  are the number of true positive, true negative, false positive, and false negative pixels, respectively. In the case of EAT or PAT segmentation, the number of positive pixels is usually much smaller than the number of negative (background) pixels, as the tissues themselves are thin and dispersed around the image. This means that a segmented mask that is fully negative (all pixels are 0) would still have high accuracy, as the number of true negatives would be much larger than the number of false positives. It is clear that in cases where the number of positive and negative pixels is highly imbalanced, pixel-level accuracy can be misleading.

Instead of using pixel-level accuracy, pixel-level sensitivity, the Dice coefficient, and the Jaccard index are metrics that are much better at representing true segmentation performance in these unbalanced examples. We strongly suggest to future authors of papers in this field to consider these metrics over the pixel-level accuracy.

Alongside segmentation accuracy and efficiency, in medical image segmentation, an important metric is segmentation precision or reliability [57]. Even if a segmentation algorithm is highly accurate for most examples, if it fails dramatically on other examples it will be hard to use the method in clinical practice. The simplest metric to evaluate reliability is by calculating and reporting the standard deviation of the segmentation metrics over the evaluation dataset. Other ways used to show reliability by the papers mentioned in this review include the inter-quartile range [7], or the minimum and maximum value of the metric [24,42]. Where reported by the authors, we have included these metrics in the Result columns in Tables 1 and 2.

Despite the importance of segmentation reliability, 10 papers mentioned in this review do not report any reliability metrics for segmentation performance. Future authors in this field should present the standard deviation of their metrics at a minimum, but also present more detailed reliability metrics or plots such as box plots across their evaluation dataset. This would add a dimension along which to more objectively compare methods.

For evaluating quantification performance, papers mentioned in this review mostly use correlation measures such as the Pearson correlation coefficient ( $r$ ) or the Spearman correlation coefficient ( $\rho$ ). While these metrics have some descriptive power of the quantification performance, they can potentially mislead if the sample size is large enough or if the measurements themselves have a low variability [58]. For evaluating quantification in medical imaging, a better evaluation tool is the Bland-Altman plot, which shows agreement and not correlation. An example of a Bland-Altman plot is shown in Figure 13. Measuring the agreement of EAT or PAT volume between expert measurements and a computerized segmentation method or between two segmentation methods is a more objective indicator of quantification performance than a single correlation measure.



**Figure 13.** An example of a Bland-Altman plot for EAT volume from [44].

Several papers mentioned in this review do not perform a Bland-Altman analysis of their results. We recommend that future authors use Bland-Altman analysis and other measurement agreement metrics when evaluating quantification performance.

### 7.2. Test Dataset Considerations

In this section, we describe how the methods mentioned in this review are evaluated using a test dataset. In this section, the test dataset, also known as the holdout dataset, refers to the dataset used to calculate evaluation metrics at the end of the model development process. In a machine-learning-based method, this dataset should be completely independent of the training and validation datasets, used to train models and to set the model's hyperparameters, respectively [59].

The best test dataset would be a sufficiently large set of data collected independently and after the model has been developed in real-world applications. In practice, this is often infeasible so researchers rely on splitting their originally available dataset into training, validation, and testing datasets, or using cross-validation to train and evaluate the model on several independent subsets of the original dataset.

In addition, when using learned models such as machine- or deep learning-based methods, it is vital that the test dataset is independent of the training dataset, i.e., that the researchers avoid data leakage. Data leakage happens when the training dataset contains examples from the test dataset [60]. If a learned model had access to the testing data during training, its performance on the test dataset will no longer be representative of any real-world performance.

When training machine learning-based models for EAT and PAT segmentation on 3D data like CT scans, another more subtle form of data leakage is possible. Some researchers perform a train, validation, and test split slice-wise, by randomly sampling a percentage of all slices from all patients into the train, validation, and test datasets. By doing so, different slices of the same patient end up in different datasets. Two neighboring slices of the same patient will be very similar, so a model trained on one or more slices of a patient will perform misleadingly well on their neighboring slices in the test dataset. This means that a slice-based test dataset is not independent of the training set and is not representative of any real segmentation or quantification performance. In real-world uses, the model would get slices from completely new patients which the model has not seen before.

Several papers mentioned in this work use slice-based dataset splitting including [45] and, while it is not explicitly stated, potentially also [20] (describes a very similar method with similar results to [45]), and [46] (uses 10-fold cross-validation which would not be feasible with only 20 patients). Ref. [47] does not provide sufficient information to determine how the validation dataset was obtained but has results similar to the above-mentioned methods. Note that these papers are also the ones that report much higher performance in terms of Dice coefficients and accuracy than any other paper in this review, suggesting that data leakage has occurred and that the results are not representative of real-world performance. The results of methods that split their datasets based on slices should not be directly compared to the results of other methods which perform their dataset split per patient.

We recommend that dataset splits should always be made per patient and not per slice since that better represents how a method would be used in a real setting. In addition, authors should consider using cross-validation, especially when the number of patients is relatively low like in the publicly available dataset from [20]. Finally, authors should clearly describe how they perform the validation and training dataset split so that their results are reproducible for future researchers.

### *7.3. Data Quality, Quantity and Diversity*

Tang et al. [61] show that there is a difference in brain morphology among different races and genetic backgrounds. In particular, the authors show that there are significant differences in cortical volume in various lobes of the brain between caucasian and Chinese cohorts. There is a possibility there are differences in the distribution of EAT in different ethnic, racial, and cultural groups. Additionally, the attenuation value of EAT in CTs can vary from patient to patient, and the distribution and volume vary depending on the patient's weight and overall health. To better understand the true performance of these methods, there is a need for larger datasets that include patients of different genders, races, geographical locations, health and fitness levels as well as pre-existing conditions.

While some papers mentioned in this review use comparatively large datasets of 100 patients or more [7,9,25,54], most validate their methods on smaller datasets, even as low as 10 patients. It is our opinion that using a dataset smaller than 100 patients is not sufficient to validate that a method generalizes well enough for medical use, neither on the same distribution of patients as in the dataset nor for patients of other backgrounds.

Despite the availability of unlabeled cardiac CT scans (e.g., [62]), due to the time-consuming nature of manually labeling EAT, there is not much publicly available data with labeled EAT and PAT volumes.

The one notable exception is [63], collected Rodrigues et al. [20]. This is a publicly available dataset of CT scans of 20 patients. This dataset is later used in many papers from this review [27,36–38,43–47,49]. While using a common dataset means a more objective measure of the algorithm's performance, this dataset still only contains 20 patients from the same geographic area.

With all this taken into account, it is clear there is a need for larger publicly available datasets of labeled EAT, especially CT scans. This can be addressed by a large-scale EAT quantification challenge, similar to Zhuang et al. [62] but focusing on EAT segmentation and quantification. Aside from collecting more data from expert observers, new image augmentation methods and methods for the automatic generation of images could help. Using adversarial image generation methods could also provide artificially created data useful for training neural networks [64,65].

#### 7.4. Direct EAT Segmentation vs. Segmentation via Pericardium Delineation

The two main ways to segment EAT are either to segment it directly, or to segment the pericardium and then segment EAT from the region inside the pericardium. Using the method of first delineating the pericardium, Zhang et al. [36] achieves a mean Dice coefficient of 0.912. This is a better result than Li et al. [38] achieved on the same dataset (AUC = 0.87), suggesting that deep learning methods benefit from delineating the pericardium before proceeding with EAT segmentation. This conclusion is also corroborated by the experiments in Zhang et al. [36]. When training other common segmentation networks to directly segment EAT, the best Dice coefficients the authors achieve are 76.59% for U-Net and 76.68% for Seg-Net.

#### 7.5. Fully- vs. Semi-Automatic Results

Fully automatic methods report a wide range of Dice coefficients for EAT segmentation ranging from 0.94 to 0.69. In terms of Dice coefficients, the reported results that are near or better than inter-observer results are [20,25,36,42,46,47]. However, in Priya and Sudha [47] and Ding et al. [42], the authors validate their method on only 30 and 15 patients, respectively, and [20,36,46] all use the same dataset of only 20 patients from the same geographic area. Given the small study size these works were validated on, it is hard to tell whether their results would generalize well to a wider population. Shahzad et al. [25] report a Dice coefficient of 0.89, which is around the lower end of reported inter-observer variability. Given that three out of 98 scans failed to be segmented and were removed from the analysis, we may conclude that their Dice coefficient would be slightly lower in real-world scenarios. After taking all of this into account, we conclude that the current state-of-the-art fully automatic EAT segmentation methods are not yet robust enough for use in studies or clinical settings, and are not better than manual segmentation. For fully automatic PAT segmentation, He et al. [54] evaluate their PAT segmentation method on 422 patients and report Dice coefficients 0.88 or greater, which is slightly lower than reported inter-observer variability.

Semi-automatic methods, on the other hand, seem to achieve better results. Militello et al. [9] report a Dice coefficient of 0.93, slightly better than reported inter-observer Dice coefficients, and validates the result on 145 patients with the largest number of slices of any work to our knowledge. The patients were also located in Sicily, a notably diverse region, which could also mean their results are more general than in other works. Still, the automatic part of their method not only saves time for the physicians but also offers fewer avenues for inter-observer variability errors and simple human errors. Kazemi et al. [46] report a Dice coefficient of 0.94 with another semi-automatic method, albeit validated on a much smaller cohort of 20 patients. We conclude that, currently, semi-automatic methods perform better than fully automatic methods for EAT segmentation. We

also conclude that semi-automatic methods may be better than fully manual segmentation, both in terms of saving time and improving the accuracy of the segmentation.

### 7.6. Trends and Future Challenges

In recent years, there have been more and more papers using deep learning to achieve EAT segmentation. This is a notable trend in the field and one that will probably continue as deep-learning-based methods develop and mature. Deep learning methods achieve state-of-the-art results in other cardiac imaging segmentation tasks [55], using very similar data to data used for EAT segmentation, which might indicate that it is possible to further improve the results of deep-learning-based methods.

Another important dimension on which to validate EAT segmentation methods is model interpretability. Deep learning methods are notoriously hard to interpret, and their results can be surprising and inconsistent with expectations for a given CT scan. Slight variations in noise or structure that are easy to spot by humans could produce widely different results. Knowing these limitations ahead of time is hard because deep learning models are usually “black boxes”. Once the model is trained, it is hard to determine exactly what the model is doing to achieve its results. Non-learning methods like active contours and region growing provide results that are more in line with expectations, and their limitations can often be tested and explained more easily. However, these approaches have lower limits on performance in other classification tasks compared to deep learning, and the same is probably true for EAT segmentation.

## 8. Conclusions

Developing fast and accurate ways to quantify EAT is an important task that could improve CVD risk assessment, CVD prevention, additional insight into the exact relationship between various CVDs (such as coronary atherosclerosis, metabolic syndrome, and atrial fibrillation), and EAT volume. Methods that are widely used in clinical settings are either inaccurate when using surrogate measures or slow when manually quantifying EAT volume from CT or MRI imaging.

After analyzing a wide range of methods of automatic EAT and PAT segmentation, ranging from traditional image processing methods like active contours and region growing to more novel approaches like deep learning, we conclude that there is still a need for more accurate EAT and PAT segmentation methods validated on larger and more diverse patient cohorts. Additionally, researching methods for better pericardium delineation could enable better EAT and PAT segmentation, as well as other tasks such as whole heart segmentation.

While the existing fully automatic methods show promising results and very fast segmentation times, they are still equal to or below inter-observer variability for EAT segmentation when validating on large patient cohorts. On the other hand, some semi-automatic methods perform better than the fully-automatic ones and show potential for clinical use, but might be too impractical to use in large studies. The biggest challenge to overcome when segmenting EAT is the lack of labeled datasets. More publicly available datasets, as well as new techniques in image synthesis, could help overcome this challenge.

**Author Contributions:** Conceptualization, M.B. and M.H.; writing—original draft preparation, M.B. and M.H.; writing—review and editing, M.B., M.H., I.G. and A.P.; visualization, M.B. and M.H.; supervision, I.G. and A.P.; project administration, I.G. and A.P.; funding acquisition, I.G. and A.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been supported in part by Croatian Science Foundation under the Project UIP-2017-05-4968 as well as the Flemish Government (AI Research Program).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

EAT	Epicardial adipose tissue
PAT	Pericardial adipose tissue
CVD	Cardiovascular disease
ANN	Artificial neural network
CNN	Convolutional neural network
AUC	Area under the ROC curve
IoU	Intersection over union
DSC	Sørensen–Dice coefficient

## References

- World Health Organization. *World Health Statistics 2016: Monitoring Health for the SDGs, Sustainable Development Goals*; World Health Organization: Geneva, Switzerland, 2016; p. 121.
- Iacobellis, G.; Lonn, E.; Lamy, A.; Singh, N.; Sharma, A.M. Epicardial fat thickness and coronary artery disease correlate independently of obesity. *Int. J. Cardiol.* **2011**, *146*, 452–454. [[CrossRef](#)] [[PubMed](#)]
- Raggi, P.; Alakija, P. Epicardial adipose tissue: A long-overlooked marker of risk of cardiovascular disease. *Atherosclerosis* **2013**, *229*, 32–33. [[CrossRef](#)]
- Mahabadi, A.A.; Berg, M.H.; Lehmann, N.; Kälsch, H.; Bauer, M.; Kara, K.; Dragano, N.; Moebus, S.; Jöckel, K.H.; Erbel, R.; et al. Association of Epicardial Fat With Cardiovascular Risk Factors and Incident Myocardial Infarction in the General Population. *J. Am. Coll. Cardiol.* **2013**, *61*, 1388–1395. [[CrossRef](#)] [[PubMed](#)]
- Iacobellis, G. Epicardial adipose tissue in contemporary cardiology. *Nat. Rev. Cardiol.* **2022**. [[CrossRef](#)] [[PubMed](#)]
- Marwan, M.; Achenbach, S. Quantification of epicardial fat by computed tomography: Why, when and how? *J. Cardiovasc. Comput. Tomogr.* **2013**, *7*, 3–10. [[CrossRef](#)] [[PubMed](#)]
- Commandeur, F.; Goeller, M.; Betancur, J.; Cadet, S.; Doris, M.; Chen, X.; Berman, D.S.; Slomka, P.J.; Tamarappoo, B.K.; Dey, D. Deep Learning for Quantification of Epicardial and Thoracic Adipose Tissue From Non-Contrast CT. *IEEE Trans. Med. Imaging* **2018**, *37*, 1835–1846. [[CrossRef](#)]
- Rozanski, A.; Gransar, H.; Shaw, L.J.; Kim, J.; Miranda-Peats, L.; Wong, N.D.; Rana, J.S.; Orakzai, R.; Hayes, S.W.; Friedman, J.D.; et al. Impact of Coronary Artery Calcium Scanning on Coronary Risk Factors and Downstream Testing. *J. Am. Coll. Cardiol.* **2011**, *57*, 1622–1632. [[CrossRef](#)]
- Militello, C.; Rundo, L.; Toia, P.; Conti, V.; Russo, G.; Filorizzo, C.; Maffei, E.; Cademartiri, F.; Grutta, L.L.; Midiri, M.; et al. A semi-automatic approach for epicardial adipose tissue segmentation and quantification on cardiac CT scans. *Comput. Biol. Med.* **2019**, *114*, 103424. [[CrossRef](#)]
- Sacks, H.S.; Fain, J.N. Human epicardial adipose tissue: A review. *Am. Heart J.* **2007**, *153*, 907–917. [[CrossRef](#)]
- Chenn, O.; Ahmad, I.; Hua, B.; Sockolow, J.A.; Klem, I.; Sacchi, T.; Heitner, J.F. Correlation of pericardial and mediastinal fat with coronary artery disease, metabolic syndrome, and cardiac risk factors. *J. Cardiovasc. Magn. Reson.* **2009**, *11*, 1–316. [[CrossRef](#)]
- Mahabadi, A.A.; Lehmann, N.; Kälsch, H.; Robens, T.; Bauer, M.; Dykun, I.; Budde, T.; Moebus, S.; Jöckel, K.H.; Erbel, R.; et al. Association of Epicardial Adipose Tissue With Progression of Coronary Artery Calcification Is More Pronounced in the Early Phase of Atherosclerosis. *JACC Cardiovasc. Imaging* **2014**, *7*, 909–916. [[CrossRef](#)] [[PubMed](#)]
- Gorter, P.M.; de Vos, A.M.; van der Graaf, Y.; Stella, P.R.; Doevendans, P.A.; Meijs, M.F.; Prokop, M.; Visseren, F.L. Relation of Epicardial and Pericoronary Fat to Coronary Atherosclerosis and Coronary Artery Calcium in Patients Undergoing Coronary Angiography. *Am. J. Cardiol.* **2008**, *102*, 380–385. [[CrossRef](#)] [[PubMed](#)]
- Goeller, M.; Achenbach, S.; Marwan, M.; Doris, M.K.; Cadet, S.; Commandeur, F.; Chen, X.; Slomka, P.J.; Gransar, H.; Cao, J.J.; et al. Epicardial adipose tissue density and volume are related to subclinical atherosclerosis, inflammation and major adverse cardiac events in asymptomatic subjects. *J. Cardiovasc. Comput. Tomogr.* **2018**, *12*, 67–73. [[CrossRef](#)] [[PubMed](#)]
- Chekakie, M.O.A.; Welles, C.C.; Metoyer, R.; Ibrahim, A.; Shapira, A.R.; Cytron, J.; Santucci, P.; Wilber, D.J.; Akar, J.G. Pericardial Fat Is Independently Associated With Human Atrial Fibrillation. *J. Am. Coll. Cardiol.* **2010**, *56*, 784–788. [[CrossRef](#)] [[PubMed](#)]
- Iacobellis, G. Epicardial and Pericardial Fat: Close, but Very Different. *Obesity* **2009**, *17*, 625–625. [[CrossRef](#)]
- Cheng, V.Y.; Dey, D.; Tamarappoo, B.; Nakazato, R.; Gransar, H.; Miranda-Peats, R.; Ramesh, A.; Wong, N.D.; Shaw, L.J.; Slomka, P.J.; et al. Pericardial Fat Burden on ECG-Gated Noncontrast CT in Asymptomatic Patients Who Subsequently Experience Adverse Cardiovascular Events. *JACC Cardiovasc. Imaging* **2010**, *3*, 352–360. [[CrossRef](#)]
- Wong, C.X.; Ganesan, A.N.; Selvanayagam, J.B. Epicardial fat and atrial fibrillation: Current evidence, potential mechanisms, clinical implications, and future directions. *Eur. Heart J.* **2016**, *38*, 1294–1302. [[CrossRef](#)]
- Nappi, C.; Ponsiglione, A.; Acampa, W.; Gaudieri, V.; Zampella, E.; Assante, R.; Cuocolo, R.; Mannarino, T.; Dell’Aversana, S.; Petretta, M.; et al. Relationship between epicardial adipose tissue and coronary vascular function in patients with suspected coronary artery disease and normal myocardial perfusion imaging. *Eur. Heart J. Cardiovasc. Imaging* **2019**, *20*, 1379–1387. [[CrossRef](#)]

20. Rodrigues, É.; Morais, F.; Morais, N.; Conci, L.; Neto, L.; Conci, A. A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography. *Comput. Methods Programs Biomed.* **2016**, *123*, 109–128. [[CrossRef](#)]
21. Carr, J.J.; Ding, J. Response to “Epicardial and Pericardial Fat: Close, but Very Different”. *Obesity* **2009**, *17*, 626–627. [[CrossRef](#)]
22. Iacobellis, G.; Willens, H.J. Echocardiographic Epicardial Fat: A Review of Research and Clinical Applications. *J. Am. Soc. Echocardiogr.* **2009**, *22*, 1311–1319. [[CrossRef](#)] [[PubMed](#)]
23. Barbosa, J.G.; Figueiredo, B.; Bettencourt, N.; Tavares, J.M.R. Towards automatic quantification of the epicardial fat in non-contrasted CT images. *Comput. Methods Biomech. Biomed. Eng.* **2011**, *14*, 905–914. [[CrossRef](#)] [[PubMed](#)]
24. Ding, X.; Terzopoulos, D.; Diaz-Zamudio, M.; Berman, D.S.; Slomka, P.J.; Dey, D. Automated pericardium delineation and epicardial fat volume quantification from noncontrast CT. *Med. Phys.* **2015**, *42*, 5015–5026. [[CrossRef](#)] [[PubMed](#)]
25. Shahzad, R.; Bos, D.; Metz, C.; Rossi, A.; Kirişli, H.; van der Lugt, A.; Klein, S.; Witteman, J.; de Feyter, P.; Niessen, W.; et al. Automatic quantification of epicardial fat volume on non-enhanced cardiac CT scans using a multi-atlas segmentation approach. *Med. Phys.* **2013**, *40*, 091910. [[CrossRef](#)]
26. Norlén, A.; Alvé, J.; Molnar, D.; Enqvist, O.; Norrlund, R.R.; Brandberg, J.; Bergström, G.; Kahl, F. Automatic pericardium segmentation and quantification of epicardial fat from computed tomography angiography. *J. Med. Imaging* **2016**, *3*, 034003. [[CrossRef](#)]
27. Rodrigues, É.; Pinheiro, V.; Liatsis, P.; Conci, A. Machine learning in the prediction of cardiac epicardial and mediastinal fat volumes. *Comput. Biol. Med.* **2017**, *89*, 520–529. [[CrossRef](#)]
28. Feeman, T.G. X-rays. In *The Mathematics of Medical Imaging: A Beginner’s Guide*; Springer International Publishing: Cham, Switzerland, 2015; pp. 1–11. [[CrossRef](#)]
29. Adams, R.; Bischof, L. Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16*, 641–647. [[CrossRef](#)]
30. Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. *Int. J. Comput. Vis.* **1988**, *1*, 321–331. [[CrossRef](#)]
31. Rohlfing, T.; Brandt, R.; Menzel, R.; Russakoff, D.B.; Maurer, C.R. Quo Vadis, Atlas-Based Segmentation? In *Handbook of Biomedical Image Analysis*; Springer: Boston, MA, USA, 2005; pp. 435–486. [[CrossRef](#)]
32. Bach Cuadra, M.; Duay, V.; Thiran, J.P., Atlas-based Segmentation. In *Handbook of Biomedical Imaging: Methodologies and Clinical Research*; Paragios, N., Duncan, J., Ayache, N., Eds.; Springer: Boston, MA, USA, 2015; pp. 221–244. [[CrossRef](#)]
33. Alpaydin, E. *Introduction to Machine Learning*, 2nd ed.; The MIT Press: Cambridge, MA, USA, 2010.
34. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)]
35. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015.
36. Zhang, Q.; Zhou, J.; Zhang, B.; Jia, W.; Wu, E. Automatic Epicardial Fat Segmentation and Quantification of CT Scans Using Dual U-Nets With a Morphological Processing Layer. *IEEE Access* **2020**, *8*, 128032–128041. [[CrossRef](#)]
37. He, X.; Guo, B.J.; Lei, Y.; Wang, T.; Liu, T.; Curran, W.J.; Zhang, L.J.; Yang, X. Automatic epicardial fat segmentation in cardiac CT imaging using 3D deep attention U-Net. In Proceedings of the Medical Imaging 2020: Image Processing, Houston, TX, USA, 17–20 February 2020; Išgum, I., Landman, B.A., Eds.; International Society for Optics and Photonics, SPIE: Bellingham, WA, USA, 2020; Volume 11313, pp. 589–595. [[CrossRef](#)]
38. Li, Z.; Zou, L.; Yang, R. A Neural Network-Based Method for Automatic Pericardium Segmentation. In Proceedings of the Proceedings of the 2nd International Conference on Computer Science and Software Engineering, CSSE 2019, Xi’an, China, 24–26 May 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 45–49. [[CrossRef](#)]
39. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.J.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.G.; Hammerla, N.; Kainz, B.; et al. Attention U-Net: Learning Where to Look for the Pancreas. *arXiv* **2018**, arXiv:1804.03999.
40. Hecht, H.S.; Cronin, P.; Blaha, M.J.; Budoff, M.J.; Kazerooni, E.A.; Narula, J.; Yankelevitz, D.; Abbara, S. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. *J. Cardiovasc. Comput. Tomogr.* **2017**, *11*, 74–84. [[CrossRef](#)] [[PubMed](#)]
41. Coppini, G. Quantification of Epicardial Fat by Cardiac CT Imaging. *Open Med. Inform. J.* **2010**, *4*, 126–135. [[CrossRef](#)] [[PubMed](#)]
42. Ding, X.; Terzopoulos, D.; Diaz-Zamudio, M.; Berman, D.S.; Slomka, P.J.; Dey, D. Automated epicardial fat volume quantification from non-contrast CT. In Proceedings of the Medical Imaging 2014: Image Processing, San Diego, CA, USA, 15–20 February 2014; Ourselin, S., Styner, M.A., Eds.; SPIE: Bellingham, WA, USA, 2014. [[CrossRef](#)]
43. Rodrigues, É.; Rodrigues, L.; Oliveira, L.; Conci, A.; Liatsis, P. Automated recognition of the pericardium contour on processed CT images using genetic algorithms. *Comput. Biol. Med.* **2017**, *87*, 38–45. [[CrossRef](#)] [[PubMed](#)]
44. Benčević, M.; Habijan, M.; Galić, I. Epicardial Adipose Tissue Segmentation from CT Images with A Semi-3D Neural Network. In Proceedings of the 2021 International Symposium ELMAR, Zadar, Croatia, 13–15 September 2021; pp. 87–90. [[CrossRef](#)]
45. Rodrigues, E.O.; Cordeiro de Morais, F.F.; Conci, A. On the Automated Segmentation of Epicardial and Mediastinal Cardiac Adipose Tissues Using Classification Algorithms. In *Studies in Health Technology and Informatics, Proceedings of the MEDINFO 2015: EHEALTH-ENABLED HEALTH—15th World Congress on Health and Biomedical Informatics (MEDINFO), Sao Paulo, Brazil, 19–23 August 2015*; Sarkar, I.N., Georgiou, A., Marques, P.M.D., Eds.; IOS Press: Amsterdam, The Netherlands, 2015; Volume 216, pp. 726–730. [[CrossRef](#)]

46. Kazemi, A.; Keshtkar, A.; Rashidi, S.; Aslanabadi, N.; Khodadad, B.; Esmaeili, M. Segmentation of cardiac fats based on Gabor filters and relationship of adipose volume with coronary artery disease using FP-Growth algorithm in CT scans. *Biomed. Phys. Eng. Express* **2020**, *6*, 055009. [[CrossRef](#)]
47. Priya, C.; Sudha, S. Adaptive Fruitfly Based Modified Region Growing Algorithm for Cardiac Fat Segmentation Using Optimal Neural Network. *J. Med. Syst.* **2019**, *43*, 104. [[CrossRef](#)]
48. Zlokolica, V.; Krstanović, L.; Velicki, L.; Popović, B.; Janev, M.; Obradović, R.; Ralević, N.M.; Jovanov, L.; Babin, D. Semiautomatic Epicardial Fat Segmentation Based on Fuzzy c-Means Clustering and Geometric Ellipse Fitting. *J. Healthc. Eng.* **2017**, *2017*, 1–12. [[CrossRef](#)]
49. Benčević, M.; Galić, I.; Habijan, M.; Babin, D. Training on Polar Image Transformations Improves Biomedical Image Segmentation. *IEEE Access* **2021**, *9*, 133365–133375. [[CrossRef](#)]
50. Dey, D.; Suzuki, Y.; Suzuki, S.; Ohba, M.; Slomka, P.J.; Polk, D.; Shaw, L.J.; Berman, D.S. Automated Quantitation of Pericardial Fat From Noncontrast CT. *Investig. Radiol.* **2008**, *43*, 145–153. [[CrossRef](#)]
51. Ding, X.; Pang, J.; Ren, Z.; Diaz-Zamudio, M.; Jiang, C.; Fan, Z.; Berman, D.S.; Li, D.; Terzopoulos, D.; Slomka, P.J.; et al. Automated pericardial fat quantification from coronary magnetic resonance angiography: Feasibility study. *J. Med. Imaging* **2016**, *3*, 014002. [[CrossRef](#)]
52. Bandekar, A.N.; Naghavi, M.; Kakadiaris, I.A. Automated Pericardial Fat Quantification in CT Data. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006. [[CrossRef](#)]
53. Mahalanobis, P.C. *On the Generalized Distance in Statistics*; National Institute of Science: Calcutta, India, 1936.
54. He, X.; Guo, B.J.; Lei, Y.; Wang, T.; Curran, W.J.; Liu, T.; Zhang, L.J.; Yang, X. Automatic quantification of myocardium and pericardial fat from coronary computed tomography angiography: A multicenter study. *Eur. Radiol.* **2020**, *31*, 3826–3836. [[CrossRef](#)] [[PubMed](#)]
55. Habijan, M.; Babin, D.; Galić, I.; Leventić, H.; Romić, K.; Velicki, L.; Pižurica, A. Overview of the Whole Heart and Heart Chamber Segmentation Methods. *Cardiovasc. Eng. Technol.* **2020**, *11*, 725–747. [[CrossRef](#)] [[PubMed](#)]
56. Spearman, J.V.; Meinel, F.G.; Schoepf, U.J.; Apfaltrer, P.; Silverman, J.R.; Krazinski, A.W.; Canstein, C.; Cecco, C.N.D.; Costello, P.; Geyer, L.L. Automated Quantification of Epicardial Adipose Tissue Using CT Angiography: Evaluation of a Prototype Software. *Eur. Radiol.* **2013**, *24*, 519–526. [[CrossRef](#)] [[PubMed](#)]
57. Udupa, J.K.; LeBlanc, V.R.; Zhuge, Y.; Imielinska, C.; Schmidt, H.; Currie, L.M.; Hirsch, B.E.; Woodburn, J. A framework for evaluating image segmentation algorithms. *Comput. Med. Imaging Graph.* **2006**, *30*, 75–87. [[CrossRef](#)]
58. Bland, J.M.; Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [[CrossRef](#)]
59. Xu, Y.; Goodacre, R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. *J. Anal. Test.* **2018**, *2*, 249–262. [[CrossRef](#)]
60. Kaufman, S.; Rosset, S.; Perlich, C. *Leakage in Data Mining*; ACM Press: New York, NY, USA, 2011. [[CrossRef](#)]
61. Tang, Y.; Hojatkashani, C.; Dinov, I.D.; Sun, B.; Fan, L.; Lin, X.; Qi, H.; Hua, X.; Liu, S.; Toga, A.W. The construction of a Chinese MRI brain atlas: A morphometric comparison study between Chinese and Caucasian cohorts. *NeuroImage* **2010**, *51*, 33–41. [[CrossRef](#)]
62. Zhuang, X.; Li, L.; Payer, C.; Štern, D.; Urschler, M.; Heinrich, M.P.; Oster, J.; Wang, C.; Smedby, Ö.; Bian, C.; et al. Evaluation of algorithms for Multi-Modality Whole Heart Segmentation: An open-access grand challenge. *Med. Image Anal.* **2019**, *58*, 101537. [[CrossRef](#)]
63. Visual Lab, Computing Institute IC/UFF. Cardiac Fat Database—Computed Tomography. 2015. Available online: <https://visual.ic.uff.br/en/cardio/ctfat/> (accessed on 16 October 2020).
64. Welander, P.; Karlsson, S.; Eklund, A. Generative Adversarial Networks for Image-to-Image Translation on Multi-Contrast MR Images—A Comparison of CycleGAN and UNIT. *arXiv* **2018**, arXiv:1806.07777.
65. Costa, P.; Galdran, A.; Meyer, M.I.; Abràmoff, M.D.; Niemeijer, M.; Mendonça, A.M.; Campilho, A. Towards Adversarial Retinal Image Synthesis. *arXiv* **2017**, arXiv:1701.08974.