IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 18, 2020, accepted July 2, 2020, date of publication July 9, 2020, date of current version July 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008190

Automatic Epicardial Fat Segmentation and Quantification of CT Scans Using Dual U-Nets With a Morphological Processing Layer

QI ZHANG[1], JIANHANG ZHOU[1], BOB ZHANG[1], (Senior Member, IEEE), WEIJIA JIA[2], (Fellow, IEEE), AND ENHUA WU[3], (Member, IEEE)

[1]PAMI Research Group, Department of Computer and Information Science, Faculty of Science and Technology, University of Macau, Taipa, Macau
[2]BNU-UIC Joint AI Research Institute, Beijing Normal University, Zhuhai 519087, China
[3]Faculty of Science and Technology, University of Macau, Taipa, Macau

Corresponding author: Bob Zhang (bobzhang@um.edu.mo)

This work was supported by the University of Macau (File no. MYRG2019-00006-FST).

ABSTRACT The epicardial fat plays a key role in the development of many cardiovascular diseases. It is necessary and useful to precisely segment this fat from CT scans in clinical studies. However, it is not feasible to manually segment this fat in clinical practice, as the workload and cost for technicians or physicians is quite high. In this work, we propose a novel method for automatic segmentation and quantification of epicardial fat from CT scans accurately. In detail, dual U-Nets with the morphological processing layer is used for this goal. The first network is based on the U-Net framework to detect the pericardium, before segmenting its inside region. A morphological layer is concatenated as the following layer of the first network, to refine and obtain the ideal inside region of the pericardium. While the second network is also applied using U-Net as its backbone to find and segment the epicardial fat of the processed inside region from the pericardium using the first network. Our proposed method obtains the highest mean Dice similarity (91.19%), correlation coefficient (0.9304) compared to other state-of-art methods on a cardiac CT dataset with 20 patients. The results indicate our proposed method is effective for quantifying epicardial fat automatically.

INDEX TERMS Cardiac fat, CT, deep learning, image segmentation, medical imaging analysis.

I. INTRODUCTION

The epicardial fat is a local visceral fat deposit, often existing between the pericardium and the myocardium, which can surround the coronary arteries as well as the entire heart directly [1], [2]. Many research works and studies indicate that this fat has correlation with the development of many heart diseases. For instance, authors believe that epicardial fat can cause the formation of coronary atherosclerosis directly through the local generation of the inflammatory factors [3], [4]. Researchers also have found that the epicardial fat can play a dominant role in the coronary artery disease [5]. In [6], the epicardial fat is regarded as a predictor for heart disease events, compared to other common risk factors by applying 998 candidates in a MESA study. Thus, performing segmentation and predicting the volumes of epicardial fat is a useful and interesting task for evaluating the risk of heart diseases, and predicting major adverse cardiovascular events. However, the workload and time-cost of manually segmenting and quantifying epicardial fat is very high for technicians or physicians, as the size, shape and position of the epicardial fat varies between the heart and the pericardium [7] (shown in Figs. 1-2, the pixels with red color and green color indicate the epicardial fat and mediastinal fat, respectively). Therefore, it is necessary and desired to develop a novel method to automatically segment and quantify this kind of fat.

There are many works on the development of segmenting and quantifying the epicardial fat at present. Some approaches are based on expert knowledge to detect the pericardium and quantifying the epicardial fat manually. For instance, researchers attempted to apply landmark points for initializing the entire heart segmentation [8]. In [9], authors proposed a method to trace the pericardium and extract the heart from the surrounding organs. Ding *et al.* also proposed a similar method to use landmark points to detect the

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

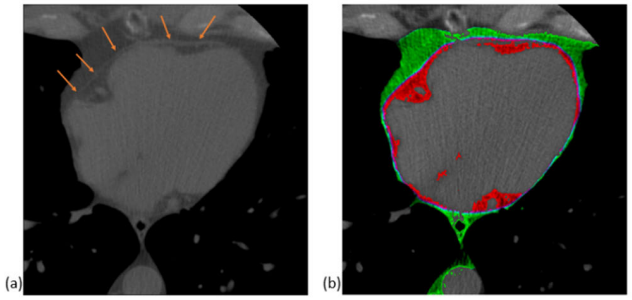

FIGURE 1. The pericardium is presented in (a) indicated by orange arrows, the pixels with red color in (b) are the epicardial fat in the cardiac CT scans.

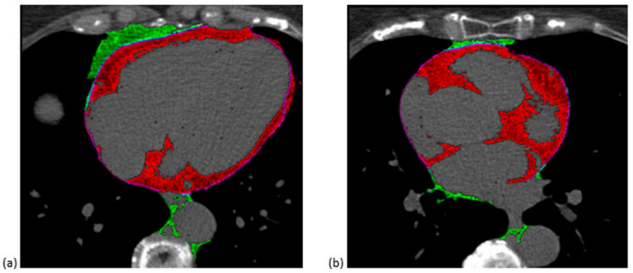

FIGURE 2. The various sizes, shapes and positions of the epicardial fat of the cardiac CT scans shown in (a) and (b) with red color.

whole heart, before further evaluating the epicardial fat [10]. Although these methods often obtain good performance for segmenting and quantifying the epicardial fat, the procedures are very complex, requiring expert participation.

Recently, some semi-automatic methods for epicardial fat quantification have been proposed. For instance, authors attempted to select the top and bottom slices of the heart and then selected 5-7 control points to anchor the position of the pericardium. Thus, the epicardial fat within the pericardium could be segmented and further evaluated [11].

Besides this, several methods are proposed to implement the segmentation and quantify the cardiac fat automatically [12]–[20]. In [21], the authors applied an atlas-based method for initializing the contour of the pericardium, and then analyzed the epicardial fat of the inside region of the pericardium (IRP). However, the border of the pericardium is sometimes not accurate, leading to some errors in epicardial fat segmentation. Rodrigues *et al.* investigated the feature extraction method comprising of classification and registration algorithms for automatic segmentation and quantification of epicardial fat. This method can obtain an ideal performance in the Dice similarity index [22], where the procedures are some what complicated. For the results of cross-validation on another patient, it can be observed that the epicardial fat is over-segmented on the cardiac CT image, which affects the effectiveness of segmentation. With the fast development of deep learning, researchers also investigated deep learning techniques for epicardial fat segmentation and quantification. Commandeur *et al.* first applied a convolutional neural network to obtain epicardial-paracardial masks of the heart. Then, they detected the contour of the pericardium and acquired the final masks of the epicardial and thoracic fat [23], [24]. Their method attempted to segment two kinds of cardiac fats simultaneously, which can cause some errors in the segmentation. Thus, there is still room for further improvement in segmenting and quantifying the epicardial fat.

To segment and quantify the epicardial fat precisely, here, we propose a novel method based on the U-Net framework, applying dual U-Nets with a morphological layer on cardiac CT scans. The U-Net is a popular framework for deep learning models, it often obtains ideal performance in image segmentation, especially in the area of medical image processing [25]–[28]. The whole pipeline of our proposed method is shown in Fig. 3. It can be noted that the first sub-network is based on the U-Net framework to detect the pericardium, before segmenting *IRP*. The segmented inside region often has some holes, irregular contour and noisy small parts on the outside (the coarse segmented *IRP* with blue color in Fig. 3). Thus, a novel layer with morphological processing is concatenated as the final following layer of the first sub-network to refine and obtain the ideal inside region of the pericardium. While the second sub-network is used to analyze the segmented part, i.e., the *IRP* from the first sub-network, and then segment and quantify the epicardial fat. Thus, the main contributions of our proposed work are follows:

- A precise automatic epicardial fat segmentation and quantification method is proposed.
- To the best of our knowledge in epicardial fat segmentation, the proposed method is implemented for the first time based on a double U-Net framework along with a morphological processing layer.
- The proposed approach is more accurate than other state-of-art segmentation approaches.

II. MATERIALS AND METHODS

As mentioned from Fig. 3, it can be observed that the cardiac CT images are input to the first sub-network to generate the raw inside regions of the pericardium, before applying a morphological processing layer to produce the images of the corresponding refined inside regions. Then, the second sub-network segments and quantifies the epicardial fat.

A. THE FIRST SUB-NETWORK ARCHITECTURE

The first sub-network architecture is based on the U-Net framework, which is the baseline and popular for medical image semantic segmentation tasks. The structure of this framework has two major characteristics: U-shape structure and skip connection [26], shown in Fig. 4. The U-shape structure can provide the useful semantic information of the segmented object in the whole image, which can reflect the relationship between the segmented object and its related environment. Thus, this structure is helpful for detecting the targeted objects. At the same time, the skip connection can

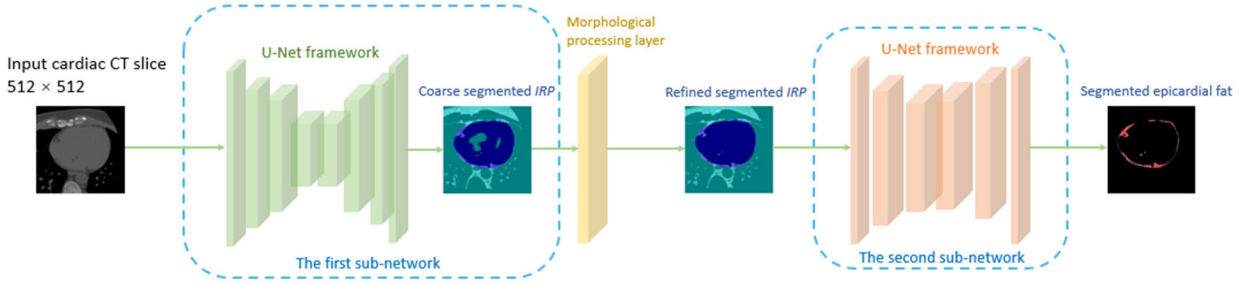

FIGURE 3. The whole pipeline of our proposed method. It contains two sub-networks with a morphological layer to segment and quantify the epicardial fat.

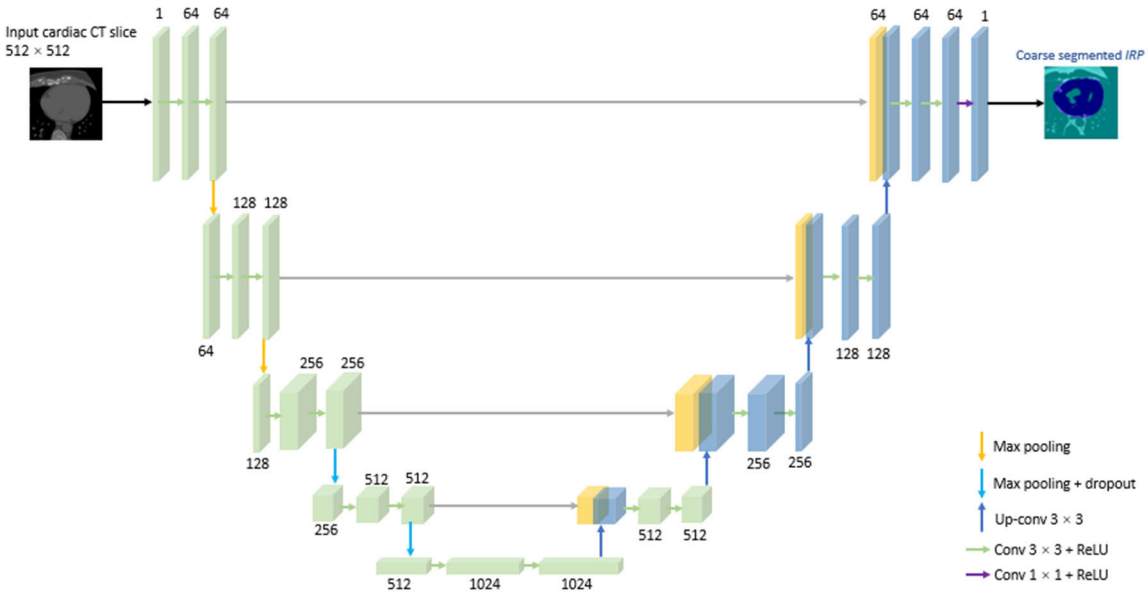

FIGURE 4. The architecture of the first sub-network. The yellow boxes indicate the feature maps used to combine with the corresponding feature maps (green boxes) after up-sampling.

transfer the high-resolution information from the encoder to the decoder of the same level. It can provide more fine features for segmentation. Thus, there are three main advantages of applying U-Net for medical imaging segmentation. The U-Net is robust and sturdy in small-size datasets; it is an effective and efficient method in medical image segmentation, can obtain the ideal performance; its structure is concise and easy to reform for further improvement [29].

In Fig. 4, the size of input raw cardiac CT image is 512×512, while the output image is a binary mask with the same size to cover the inside region of the pericardium (pixels with blue color). The outside region of *IRP* (green color) is set to zero value in order to facilitate the subsequence processing by the morphological processing. The green lines in the Fig. 4 link the encoder and decoder of the same level, which can fusion the shallow features and deep features together. To obtain the ideal inside region of the pericardium, a morphological processing layer is used to process the coarse *IRP* from the first sub-network. More details about this morphological layer can be found in the Section 2.B.

The training dataset contains around half of the whole dataset, i.e., the images of ten patients, with the remaining parts regarded as the testing dataset. For the learning rate, we attempted to evaluate various settings ($\alpha = 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}$), to find an ideal model with better performance. Fig. 5 represents the training/testing accuracy with different learning rates.

It can be seen (from Fig. 5) that the proposed method obtains the best performance with the learning rate of $\alpha = 10^{-4}$, compared to the other settings. Also, the convergence is faster than the other learning rates (i.e. the accuracy of training and testing is higher), when applying $\alpha = 10^{-4}$. Thus, the learning rate ($\alpha = 10^{-4}$) is selected for training the first sub-network in this study.

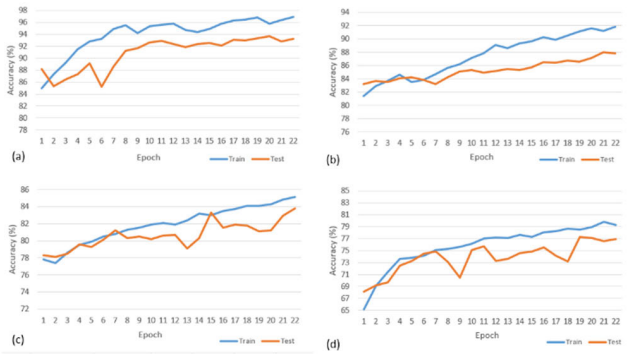

FIGURE 5. The performance of the training/testing datasets with different learning rates: (a) $\alpha = 10^{-4}$; (b) $\alpha = 10^{-5}$; (c) $\alpha = 10^{-6}$; (d) $\alpha = 10^{-7}$ for our proposed method.

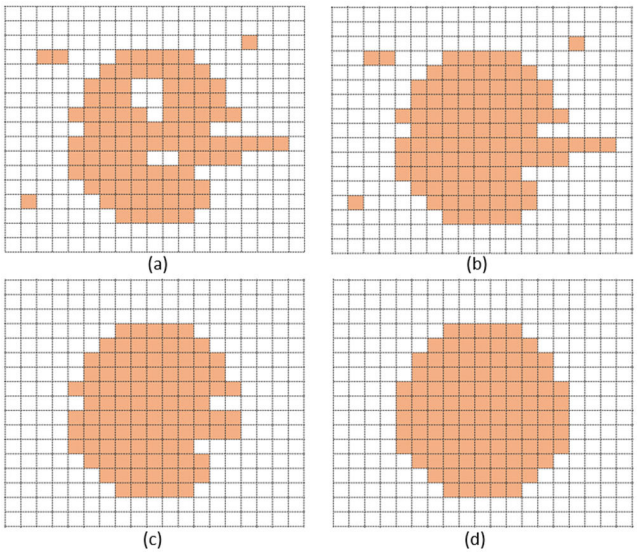

FIGURE 6. The performance of refinement by various procedures. (a) represents the coarse segmented *IRP* (pixels with orange color) without morphological processing; (b) is processed by the morphological reconstruction (Algorithm 1); (c) is obtained through the open operation; (d) can be generated by the close operation.

B. THE MORPHOLOGICAL PROCESSING LAYER

Although the simple U-Net framework in the first sub-network can detect and segment the inside region of the pericardium, the initial segmented region is still coarse with some noise. For instance, it can be observed that there are some holes, irregular contours and noisy small parts on the outside of the targeted *IRP* shown in Fig. 3. These problems on the coarse segmented *IRP* would affect the performance of quantifying epicardial fat in the second sub-network. To solve them, here, we propose a morphological processing layer [30], [31], concatenated as the following layer of the first sub-network. This layer can refine the raw segmented mask in order to generate an ideal inside region of the pericardium for the second sub-network.

The performance diagram of using this layer is represented in Fig. 6. Fig. 6(a) indicates the coarse binary mask M obtained from the first sub-network. Figs. 6(b)-(d) show some refinement algorithms to generate an ideal mask, which will

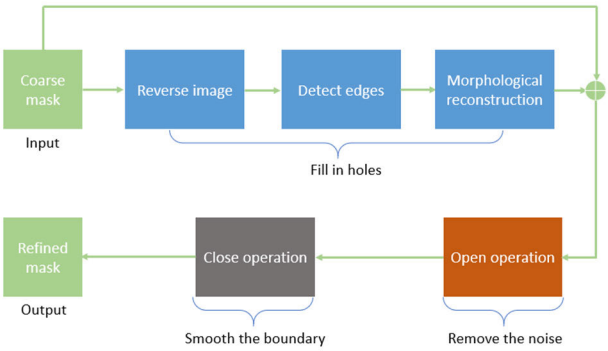

FIGURE 7. The pipeline of refining the coarse images in the morphological processing layer. The output is a refined mask covered on the corresponding CT image.

be introduced later. Here, we define two definitions for the *IRP*, as follows:

Definition 1: The connected component, i.e., *C,* contains a series of pixels, which have foreground pixels, i.e., *p,* possessing a connectivity of eight-connected neighborhoods within a random pixel q in the C.

Definition 2: The segmented *IRP* is the largest connected component within any binary mask.

Thus, it can be easy to find that the targeted segmented *IRP* can be regarded as the largest connected component and located in the center based on Definition 1 and 2, in the Fig. 6(a). Besides this, there are some noisy small parts outside the *IRP*. The *IRP* also contains some inside holes and irregular contours, which we need to deal with.

To obtain the refined *IRP* region, the specific procedures are applied in our proposed morphological processing layer, expressed in Fig. 7. The coarse binary mask produced from the basic U-Net framework is delegated as the input image. The first step is filling the inside holes of the targeted *IRP*. It contains 3 sub-procedures, i.e., reverse image, detect edges, and morphological reconstruction, respectively. The second step is based on the open operation [32] to eliminate some noisy small parts outside the *IRP*. The last step is based on the close operation [32] to smooth the boundary of the segmented *IRP*. Thus, the refined segmented *IRP* is regarded as the output from the whole pipeline in the morphological processing layer.

To fill in the inside holes of the segmented region, a morphological imaging processing method, i.e., morphological reconstruction [33] is applied. There are two kinds of images: maker image L and a template image T in the morphological reconstruction. The morphological reconstruction is an approach that can apply morphological transformations for keeping the targeted region defined within the template image from the marker image (shown in Algorithm 1). It can be considered conceptually as the repeated dilations of an image, i.e., the marker image, until the contour of the marker image fits with the template image. The reverse image R can be obtained from the given image I, which converts the hole region to the foreground by transferring the non-hole region

to the background:

$$R = 1 - I \tag{1}$$

Then, we apply an edge detection for R, where every pixel in the edge contains a series of points in a set E. Thus, any pixels that are not part of the edges are fixed to zero, which can generate the new image L:

$$L(x, y) = \begin{cases} 0, & otherwise, \\ 1, & L(x, y) \notin E. \end{cases} \tag{2}$$

The image L can be regarded as a marker image, T is described as the template image (i.e., the reverse image R generated from the given image I) in the morphological reconstruction of Algorithm 1. The new image H without holes is then generated based on images L and R:

$$H = MR(L,R,c,t) \tag{3}$$

where t indicates the tolerance to regulate the reconstruction iterative operation, and c signifies the size of the structural element. Thus, images I and H are overlapped to generate a new mask F with the holes filled (shown in Fig. 6(b)), i.e.,

$$F = I + H \tag{4}$$

For removing the noisy small parts outside *IRP*, we propose the open operation:

$$M \circ f = \bigcup \{f_d | f_d \subseteq M\} \tag{5}$$

where f represents the structural element, M describes the mask image, and f_d indicates the origin point with filter f. From this, we can see that the noisy small elements outside *IRP* are removed in Fig. 6(c).

To smooth the boundary and internal noise, here, we apply the close operation as follows:

$$M \circ f = \bigcup \{d | f_d \subseteq W\} \tag{6}$$

where W can be expressed as $W = \{d | f_d \cap M \neq \phi$, which indicates the results obtained from M dilated by the filter f. The corresponding effect of the close operation is expressed in Fig. 6(d), which is much better than the raw input mask in Fig. 6(a). The whole algorithm of the morphological processing layer can be presented in Algorithm 2.

C. THE SECOND SUB-NETWORK ARCHITECTURE

To further analyze and quantify the epicardial fat of the segmented *IRP* from the morphological layer in the first sub-network, the second sub-network is designed in this study. This network is also based on a U-Net architecture, which is shown in Fig. 8. The input image is a segmented *IRP* (512 × 512), which is obtained by the refined binary mask covering the corresponding CT image. Please note that the input image is not a binary image, but a segmented *IRP* part of the corresponding CT image, which would be processed by the second sub-network. At the same time, the output from the second sub-network is a segmented epicardial fat of the same size. As the input image is refined and processed by the first sub-network, here, we apply three encoders as well as three decoders of the same level, and set the same learning

Algorithm 1 The Morphological Reconstruction

Input: The maker L, Template T, Tolerance t, Size of the structural element c
Output: The binary image I
1: Initiating the structural element

$$K = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{r \times r} \quad \text{and the image}$$

$$I = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{m \times n}$$

2: **do**
3: $I_{p+1} = (I_p \oplus K) \cap T$
4: **While** $\sum_{i+1}^{n} \sum_{j+1}^{m} |I(i,j) - T(i,j)| \geq t$

Algorithm 2 The Layer With Morphological Processing

Input: The Filter size θ for the open operation, Filter size σ for the close operation, Threshold γ, and Mask image M,
Output: The refined image *FE*
1: prepare the structural elements,

$$s_1 = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{\theta \times \theta}, \quad s_2 = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{\sigma \times \sigma}$$

2: Produce the image R based on (1)
3: Apply edges detection step
4: Obtain the Reverse pixels S in image R based on (2)
5: Produce the holes-filled image H through (3)
6: Overlap the holes-filled image H and mask image I based on (4)
7: Apply the open operation, i.e., $O = open(F, s_1)$ based on (5)
8: Apply the close operation, i.e., $FE = close(O, s_2)$ based on (6)
9: **Return** Refined image *FE*

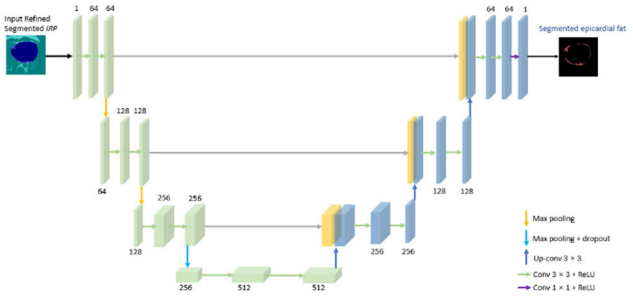

FIGURE 8. The architecture of the second sub-network. The yellow boxes indicate the feature maps used to combine with the corresponding feature maps (green boxes) after up-sampling.

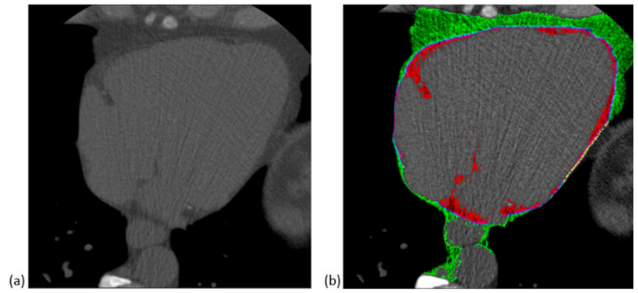

FIGURE 9. The sample of (a) raw cardiac CT scan; (b) depicts the pixels with red color, i.e., epicardial fat, is labeled as the ground truth by the experts manually.

TABLE 1. The patients demographics.

Patients	Women	Men
Ages	Mean: 53	Mean: 55.4
Slices	Total amount: 878	
	Per patient median:43.9	Per patient median:42
Manufactures	Philips: 9	Siemens: 11

rate as before. Besides this, the training dataset contains half the segmented images from the whole dataset, while the rest is regarded as the testing dataset. Therefore, we can ultimately obtain the segmented epicardial fat from this network.

III. RESULTS

In this section, we will apply our proposed method for segmenting and quantifying epicardial fat and validate its effectiveness. A cardiac CT scans dataset is first described before being applied for evaluating the performance of the proposed method. Then, the related experimental settings for each sub-network and morphological processing layer will be presented. Next, some experimental metrics for quantitatively evaluating the results are proposed. Finally, several state-of-art methods, i.e., FCN, U-Net, Seg-Net [34]–[36] are used to compare with our proposed approach.

A. DATA DESCRIPTION

In this work, we used the cardiac CT dataset from [22]. To properly evaluate the cardiac fat of the CT scans, all of the raw images in this dataset were transferred to $(-200, -30)$ the HU range to keep the cardiac fat more obvious. This dataset contains 20 patients, with 878 images, and were all manually segmented by two experts, shown in Fig. 9. All cardiac CT images are gray scale images in this study (and thus use a single channel). The pixels with red color indicate the epicardial fat, while the pixels with green color represent the mediastinal fat in the images. The demographic information of this dataset is presented in Table 1. The patients with different sizes, shapes and positions of the epicardial fat is shown in Fig. 10. All of the collected data was obtained within the ethical standards explained in the Declaration of Helsinki and more details can be accessed in [37].

B. EXPERIMENTAL SETTINGS

For the first sub-network, half of the whole dataset, i.e., ten patients with CT images are used for training, while the rest of the dataset (another ten patients with CT images) is applied for testing. Then, all of the generated images of *IRP* are refined through the morphological processing layer. The second sub-network subsequently uses the segmented images of the same part from the previous processing of the training data, before generating the corresponding epicardial fat mask compared to the ground truth marked by the experts. All related experiments were performed on a PC with an Intel(R) i7-4770 CPU, 16 GB RAM, accelerated by a NVIDIA Geforce GTX 1070 graphics card. Besides this, all experiments in this work were randomly repeated ten times on the afore-mentioned dataset.

In the training phase of the first sub-network, the learning rate, and number of epochs were set to $\alpha = 10^{-4}$, $N = 200$, respectively. While the learning rate, and number of epochs were set to $\alpha = 10^{-4}$, $N = 100$, for the second sub-network. Here, we have applied the Adam optimizer [38] for optimizing the two sub-networks, which is a robust and popular approach for training our model. The proposed model is fine-tuned carefully to obtain an ideal performance in this study. For the morphological processing layer, the filter size and tolerance of the morphological reconstruction were set to $r = 3$, $t = 0.001$, while the open operation filter and close operation filter were applied with $\rho = 4$ and $\sigma = 4$, respectively. Besides this, the single U-Net, FCN and Seg-Net are also trained and tested with the same training and testing dataset. The learning rate is set to $\alpha = 10^{-4}$ with 200 epochs for these single models. Besides this, the number of convolutional layers used in the single U-Net and Seg-Net remain the same for comparison. All mentioned methods are implemented and run on MATLAB 2019b.

C. EXPERIMENTAL METRICS

To evaluate the performance of segmenting the epicardial fat, the Intersection over Union (IOU) is applied in this study [39]. For the ground-truth image G and the predicted image P from our proposed method, the pixel $p_{ij} = 1$ indicates the positive class, and the pixel $p_{ij} = 0$ represents the negative class. Then, we can define it as:

$$IOU = \frac{|GP \cap PP|}{|GP| + |PP| - |GP \cap PP|'} \tag{7}$$

where GP and PP are the series of positive pixels, i.e., the epicardial fat region, in images G and P, respectively. Here, $|GP|$ and $|PP|$ signify the total number of pixels in the image GP and PP. The IOU can reflect the similarity between the predicted image and the ground truth image through the intersection region divided by the union region of the two images.

In addition to IOU, the Dice similarity score [40] is also applied for measuring the segmented epicardial fat, which can

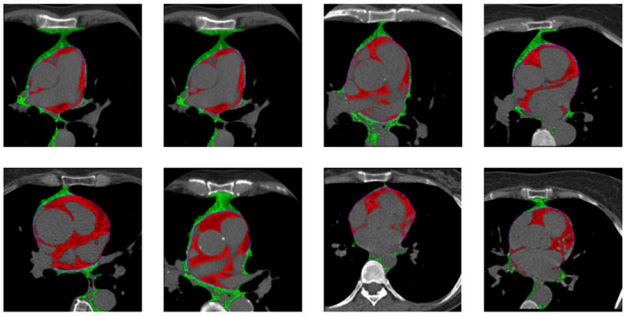

FIGURE 10. The sample CT images from 8 patients with different sizes, shapes and positions of the epicardial fat with red color.

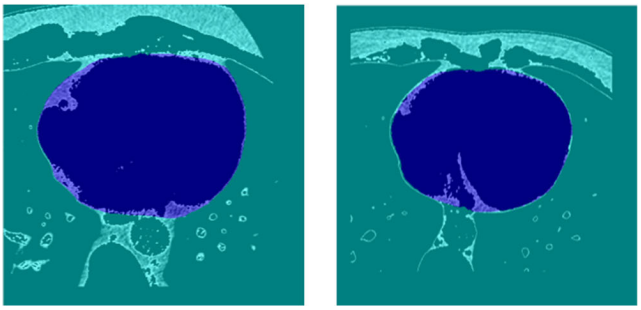

FIGURE 11. Two samples of segmented IRP (pixels with blue color) using our proposed first sub-network.

be defined as:

$$Dice\,(GP, PP\,) = \frac{2|GP \cap PP|}{|GP| + |PP|} \tag{8}$$

The definition of GP and PP are the same to the IOU in (7). The Dice similarity score is a common metric to evaluate the results of segmentation in medical imaging analysis.

To quantify the volumes of the epicardial fat, the Pearson correlation coefficient ρ is applied in this work [41], which can be expressed as:

$$\rho = \frac{\sum_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2 \sum_{i=1}^{n}(b_i - \bar{\bar{b}})^2}} \tag{9}$$

where a_i and b_i denote the predicted/original values of instance i in the dataset, while $\bar{a}$ and $\bar{b}$ indicate the mean of the predicted/original values, and n is the total number of instances.

D. EXPERIMENTAL RESULTS

To evaluate the segmented *IRP* through the first sub-network with the morphological processing layer, the mean IOU and mean Dice similarity score were applied to the testing dataset. The ground truth mask to do the *IRP* segmentation is labeled manually based on the pericardium from the CT images, as the epicardial fat (red color) is located inside of the pericardium as shown in Fig. 1. The related results show that our proposed approach can obtain ideal performances in segmenting *IRP*, reaching 95.05% and 91.14% in the mean Dice similarity score and mean IOU, respectively. The samples of the segmented results of *IRP* are represented in Fig. 11, which also validates the effectiveness of the proposed first-subnetwork with a morphological processing layer.

The results of segmenting and quantifying epicardial fat through the second sub-network are shown in Table 2. Besides this, other state-of-the-art segmentation methods, i.e., FCN, U-Net and Seg-Net [34]–[36] are also assessed here. From Table 2, it can be seen that our proposed method reaches the highest performance in segmenting and quantifying epicardial fat, obtaining 84.25%, 91.19%, and 0.9304 in mean IOU, mean Dice score and Person correlation coefficient, respectively. The segmented samples of epicardial fat

TABLE 2. The results of our proposed approach and other state-of-the-art segmentation methods.

Methods	Mean IOU (%)	Mean Dice score (%)	Pearson correlation coefficient
Proposed method	84.25 ± 1.52	91.19 ± 1.41	0.9304
Commandeur et al. [23]	/	82.30	0.9240
FCN [34]	56.39 ± 3.48	71.21 ± 3.03	0.7394
U-Net [35]	62.54 ± 3.41	76.59 ± 2.92	0.7927
Seg-Net [36]	62.67 ± 3.45	76.68 ± 2.97	0.7931

by applying our proposed method and other state-of-the-art models are also shown in Fig. 12. The conventional U-Net and FCN methods often regarded the surrounding parts of the epicardial fat as the predicted label, which caused higher errors and a lower index in the mean IOU and mean Dice score in Table 2. Furthermore, it was easy for Seg-Net to obtain under-segmentation results shown in Fig. 12, which also affects the final results, displayed in Table 2. The training time of our proposed method, Commandeur *et al.* [23], FCN, single U-Net and Seg-Net are around 13.5 hours, 11 hours, 7.5 hours, 8 hours, and 8.5 hours, respectively.

IV. DISCUSSION

Our proposed method contains two parts, the first part contains a network based on the U-Net framework with a morphological processing layer, while the second part aims to extract the epicardial fat by using the processed images from the first part. Extensive results of our approach compared to other state-of-the-art methods are shown in Table 2 and Fig. 12, verifying its effectiveness and quality. As the epicardial fat is located in the inside region of the pericardium (shown in Fig. 1), the key is to detect the pericardium for segmentation. By using the first sub-network we proposed, it can avoid the interference of the other surrounding parts, which can improve the performance of the whole method. Other state-of-the-art methods segment the epicardial fat often mistakenly identifying and segmenting its surrounding tissue, decreasing their performances. The segmented samples in Figs. 10 (c)-(e) can also confirm this argument.

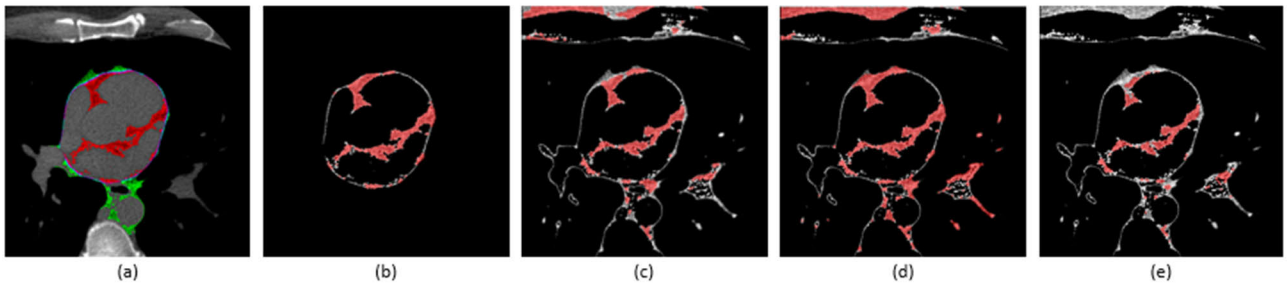

FIGURE 12. The samples of (a) ground truth labelled cardiac CT image, (b) the segmented result through our proposed method, (c) the segmented result by U-Net, (d) the segmented result by FCN, and (e) the segmented result from Seg-Net. The (c)-(e) show wrong segmentation of many parts outside of the *IRP*.

The morphological processing layer used as the final layer of the first sub-network can refine the coarse segmentation of the pericardium, obtaining an ideal *IRP* for the second sub-network. There are three parts in the morphological processing layer, i.e., morphological reconstruction, open operation and close operation with different parameter settings. Here, we aim to keep a balance between its effectiveness and efficiency by fine-tuning different filter sizes and tolerances to generate the desired results.

The novelty of this work is that the outside organs of an *IRP* can affect the performance of segmenting the epicardial fat without using the morphological layer. It can be noticed that our proposed dual U-Nets with the morphological layer obtains the highest results in mean Dice score (91.19%) and correlation coefficient (0.9304), compared to other state-of-the-art methods such as the single U-Net, FCN and Seg-Net. Even comparing with the works of others in segmenting epicardial fat by CNN, our approach is still better, the mean Dice score is 91.19% using our method, while the mean Dice score of [23] is 82.30% in Table 2. Beside this, the segmented samples in Fig. 10 can also confirm this argument. These results can support the novelty and effectiveness of our proposed method.

From Table 2, it can be noticed that the performance of a single U-Net is still slightly lower than the Seg-Net (Mean Dice score: 76.59% vs 76.68%). However, the training time of Seg-Net is also slightly longer than a single U-Net. The training time of a single U-Net is around 8 hours, while the Seg-Net needs about 8.5 hours. Considering the training time and the similar performances of a single U-Net and Seg-Net, we attempt to apply the dual U-Net for segmenting the epicardial fat. It is an interesting topic to investigate dual Seg-Net or dual FCN with the morphological processing layer, and compare it with our proposed method. Further analysis of these three methods with a morphological layer are warranted in the future.

The limitation of our proposed method is that the two networks with a morphological layer are connected together, which requires more computation time compared to the single deep learning models such as FCN (training times: 13.5 hours – proposed method vs 7.5 hours – FCN). Researchers [14] have attempted to apply the thresholding methods to processing the CT images, which aims to reduce the interference of other organs outside of the heart. The time cost of this method is much smaller and is easier to implement. Inspired from their approach, it is reasonable to attempt and apply the thresholding techniques to replace the first sub-network before the morphological layer. Thus, the time cost is much lower than before, with the possibility of a similar performance. Further evaluation of this is necessary in the future. Besides this, the procedure shown in Fig. 3 is also more slightly complicated than other end-to-end methods, where further analysis to simplify our proposed method is warranted.

It can be noted that there are some green parts existing in Fig. 12 (a), which indicates the mediastinal fat outside of the pericardium. As this fat has various shapes and sizes, it is not easy to segment via some state-of-the-art models. Thus, it would be an interesting topic to implement a novel method based on our proposed approach to segment this fat as part of our future work.

V. CONCLUSION

In this paper, we presented a novel method to segment and quantify the epicardial fat in cardiac CT scans. This method can detect the inside region of the pericardium, refine the targeted mask and segment the epicardial fat automatically. The results of this method are in strong agreement with the manually marked areas by experts and performs much better compared to other state-of-the-art methods. In conclusion, the proposed approach can be used as a tool to reduce the processing time of segmenting the epicardial fat, with the potential to be applied for clinical use in the future.

REFERENCES

[1] Y. Wu, A. Zhang, D. J. Hamilton, and T. Deng, "Epicardial fat in the maintenance of cardiovascular health," *Methodist DeBakey Cardiovascular J.*, vol. 13, no. 1, p. 20, 2017.

[2] J. Salazar, E. Luzardo, J. C. Mejías, J. Rojas, A. Ferreira, J. R. Rivas-Ríos, and V. Bermúdez, "Epicardial fat: Physiological, pathological, and therapeutic implications," *Cardiol. Res. Pract.*, vol. 2016, pp. 1–15, 2016.

[3] A. A. Mahabadi, N. Reinsch, N. Lehmann, J. Altenbernd, H. Kälsch, R. M. Seibel, R. Erbel, and S. Möhlenkamp, "Association of pericoronary fat volume with atherosclerotic plaque burden in the underlying coronary artery: A segment analysis," *Atherosclerosis*, vol. 211, no. 1, pp. 195–199, 2010.

[4] N. Bettencourt, A. M. Toschke, D. Leite, J. Rocha, M. Carvalho, F. Sampaio, S. Xará, A. Leite-Moreira, E. Nagel, and V. Gama, "Epicardial adipose tissue is an independent predictor of coronary atherosclerotic burden," *Int. J. Cardiol.*, vol. 158, no. 1, pp. 26–32, Jun. 2012.
[5] N. Alexopoulos, D. S. McLean, M. Janik, C. D. Arepalli, A. E. Stillman, and P. Raggi, "Epicardial adipose tissue and coronary artery plaque characteristics," *Atherosclerosis*, vol. 210, no. 1, pp. 150–154, May 2010.
[6] J. Ding, F.-C. Hsu, T. B. Harris, Y. Liu, S. B. Kritchevsky, M. Szklo, P. Ouyang, M. A. Espeland, K. K. Lohman, M. H. Criqui, M. Allison, D. A. Bluemke, and J. J. Carr, "The association of pericardial fat with incident coronary heart disease: The multi-ethnic study of atherosclerosis (MESA)," *Amer. J. Clin. Nutrition*, vol. 90, no. 3, pp. 499–504, Sep. 2009.
[7] S. W. Rabkin, "Epicardial fat: Properties, function and relationship to obesity," *Obesity Rev.*, vol. 8, no. 3, pp. 253–261, May 2007.
[8] G. L. Wheeler, R. Shi, S. R. Beck, and C. D. Langefeld, "Pericardial and visceral adipose tissues measured volumetrically with computed tomography are highly associated in type 2 diabetic families," *Investigative Radiol.*, vol. 40, no. 2, pp. 97–101, 2005.
[9] G. A. Rosito, H. S. Markus, M. L. Bots, M. Rosvall, and M. Sitzer, "Clinical perspective," *Circulation*, vol. 117, no. 5, pp. 605–613, 2008.
[10] J. Ding, S. B. Kritchevsky, T. B. Harris, G. L. Burke, R. C. Detrano, M. Szklo, J. J. Carr, and M.-E. Study of Atherosclerosis, "The association of pericardial fat with calcified coronary plaque," *Obesity*, vol. 16, no. 8, pp. 1914–1919, Aug. 2008.
[11] D. Dey, Y. Suzuki, S. Suzuki, M. Ohba, P. J. Slomka, D. Polk, L. J. Shaw, and D. S. Berman, "Automated quantitation of pericardiac fat from noncontrast CT," *Investigative Radiol.*, vol. 43, no. 2, pp. 145–153, Feb. 2008.
[12] H. H. Hu, J. Chen, and W. Shen, "Segmentation and quantification of adipose tissue by magnetic resonance imaging," *Magn. Reson. Mater. Phys., Biol. Med.*, vol. 29, no. 2, pp. 259–276, Apr. 2016.
[13] C. Priya and S. Sudha, "Adaptive fruitfly based modified region growing algorithm for cardiac fat segmentation using optimal neural network," *J. Med. Syst.*, vol. 43, no. 5, p. 104, May 2019.
[14] V. H. C. de Albuquerque, D. de A. Rodrigues, R. F. Ivo, S. A. Peixoto, T. Han, W. Wu, and P. P. Rebouças Filho, "Fast fully automatic heart fat segmentation in computed tomography datasets," *Computerized Med. Imag. Graph.*, vol. 80, Mar. 2020, Art. no. 101674.
[15] X. He, B. J. Guo, Y. Lei, T. Wang, and T. Liu, "Automatic epicardial fat segmentation in cardiac CT imaging using 3D deep attention U-Net," *Proc. SPIE*, vol. 11313, 2020, Art. no. 113132D.
[16] F. Zhao, H. Hu, Y. Chen, J. Liang, X. He, and Y. Hou, "Accurate segmentation of heart volume in cta with landmark-based registration and fully convolutional network," *IEEE Access*, vol. 7, pp. 57881–57893, 2019.
[17] É. Rodrigues, V. H. A. Pinheiro, P. Liatsis, and A. Conci, "Machine learning in the prediction of cardiac epicardial and mediastinal fat volumes," *Comput. Biol. Med.*, vol. 89, pp. 520–529, Oct. 2017.
[18] V. Zlokolica, L. Krstanović, L. Velicki, B. Popović, M. Janev, R. Obradović, N. M. Ralević, L. Jovanov, and D. Babin, "Semiautomatic epicardial fat segmentation based on fuzzy c-means clustering and geometric ellipse fitting," *J. Healthcare Eng.*, vol. 2017, Sep. 2017, Art. no. 5817970.
[19] E. O. Rodrigues, A. Conci, F. F. C. Morais, and M. G. Perez, "Towards the automated segmentation of epicardial and mediastinal fats: A multi-manufacturer approach using intersubject registration and random forest," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Mar. 2015, pp. 1779–1785.
[20] D. Aarthy, C. Priya, and S. Sudha, "Deep learning for quantification of epicardial fat from non-contrast CT," in *Proc. IEEE Int. Conf. Intell. Techn. Control, Optim. Signal Process. (INCOS)*, Apr. 2019, pp. 1–6.
[21] R. Shahzad, D. Bos, C. Metz, A. Rossi, and H. Kirişli, "Automatic quantification of epicardial fat volume on non-enhanced cardiac CT scans using a multi-atlas segmentation approach," *Med. Phys.*, vol. 40, no. 9, 2013, Art. no. 091910.
[22] O. Rodrigues, F. Morais, N. Morais, L. Conci, L. Neto, and A. Conci, "A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography," *Comput. Methods Programs Biomed.*, vol. 123, pp. 109–128, Jan. 2016.
[23] F. Commandeur, M. Goeller, J. Betancur, S. Cadet, M. Doris, X. Chen, D. S. Berman, P. J. Slomka, B. K. Tamarappoo, and D. Dey, "Deep learning for quantification of epicardial and thoracic adipose tissue from non-contrast CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 8, pp. 1835–1846, Aug. 2018.
[24] F. Commandeur, M. Goeller, A. Razipour, S. Cadet, M. M. Hell, J. Kwiecinski, X. Chen, H.-J. Chang, M. Marwan, S. Achenbach, D. S. Berman, P. J. Slomka, B. K. Tamarappoo, and D. Dey, "Fully automated CT quantification of epicardial adipose tissue by deep learning: A multicenter study," *Radiol., Artif. Intell.*, vol. 1, no. 6, Nov. 2019, Art. no. e190045.
[25] S. Li, M. Dong, G. Du, and X. Mu, "Attention Dense-U-Net for automatic breast mass segmentation in digital mammogram," *IEEE Access*, vol. 7, pp. 59037–59047, 2019.
[26] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Munich, Germany: Springer, 2015, pp. 234–241.
[27] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "RIC-unet: An improved neural network based on unet for nuclei segmentation in histology images," *IEEE Access*, vol. 7, pp. 21420–21428, 2019.
[28] L. Ding, K. Zhao, X. Zhang, X. Wang, and J. Zhang, "A lightweight U-Net architecture multi-scale convolutional network for pediatric hand bone segmentation in X-ray image," *IEEE Access*, vol. 7, pp. 68436–68445, 2019.
[29] J. Zhou, Q. Zhang, B. Zhang, and X. Chen, "TongueNet: A precise and fast tongue segmentation system using U-Net with a morphological processing layer," *Appl. Sci.*, vol. 9, no. 15, p. 3128, Aug. 2019.
[30] E. R. Dougherty, and R. A. Lotufo, *Hands-on Morphological Image Processing*, vol. 59. Bellingham, WA, USA: SPIE press, 2003.
[31] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin, Germany: Springer, 2013.
[32] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*. Boston, MA, USA: Cengage, 2014.
[33] R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2006.
[34] Y. Shen, Z. Fang, Y. Gao, N. Xiong, C. Zhong, and X. Tang, "Coronary arteries segmentation based on 3D FCN with attention gate and level set function," *IEEE Access*, vol. 7, pp. 42826–42835, 2019.
[35] Y. Zhou, W. Huang, P. Dong, Y. Xia, and S. Wang, "D-UNet: A dimension-fusion u shape network for chronic stroke lesion segmentation," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, Sep. 6, 2019.
[36] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[37] *Cardiac Fat Database–Computed Tomography*. Accessed: Apr. 2, 2020. [Online]. Available: http://visual.ic.uff.br/en/cardio/ctfat/
[38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[39] S. Kosub, "A note on the triangle inequality for the jaccard distance," *Pattern Recognit. Lett.*, vol. 120, pp. 36–38, Apr. 2019.
[40] J. Bertels, M. Berman, D. Vandermeulen, F. Maes, R. Bisschops, and M. B. Blaschko, "Optimizing the Dice score and Jaccard index for medical image segmentation: Theory and practice," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Shenzhen, China: Springer, vol. 2019, pp. 92–100.
[41] L. Sheugh and S. H. Alizadeh, "A note on pearson correlation coefficient as a metric of similarity in recommender system," in *Proc. AI Robot. (IRANOPEN)*, Apr. 2015, pp. 1–6.

QI ZHANG received the M.Sc. degree in electrical and computer engineering from the University of Macau, Macau, in 2018, where he is currently pursuing the master's degree with the Department of Computer and Information Science. His research interests include medical image analysis, medical biometrics, deep learning, and image processing.

JIANHANG ZHOU received the B.S. degree from Nanjing Forestry University (NJFU), Nanjing, China, in 2018. He is currently pursuing the M.S. degree in computer science with the Department of Computer and Information Science, Faculty of Science and Technology, University of Macau.

BOB ZHANG (Senior Member, IEEE) received the B.A. degree in computer science from York University, Toronto, ON, Canada, in 2006, the M.A.Sc. degree in information systems security from Concordia University, Montreal, QC, Canada, in 2007, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2011. He was with the Center for Pattern Recognition and Machine Intelligence and a Postdoctoral Researcher with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA. He is currently an Associate Professor with the Department of Computer and Information Science, University of Macau, Macau. His research interests include biometrics, pattern recognition, and image processing. He is a Technical Committee Member of the IEEE Systems, Man, and Cybernetics Society. He is also an Associate Editor of *IET Computer Vision*.

WEIJIA JIA (Fellow, IEEE) received the B.Sc. and M.Sc. degrees from Center South University, China, in 1982 and 1984, respectively, and the Master of Applied Science and Ph.D. degrees from the Polytechnic Faculty of Mons, Belgium, in 1992 and 1993, respectively, all in computer science. From 1993 to 1995, he joined the German National Research Center for Information Science (GMD) in Bonn (St. Augustine) as a Research Fellow. From 1995 to 2013, he worked with the City University of Hong Kong as a Professor. His contributions have been recognized as optimal network routing and deployment; vertex cover; anycast and QoS routing, and sensors networking; knowledge relation extractions; and NLP and edge computing. He is currently the Chair Professor and the Director of the BNU-UIC Joint AI Research Institute (Zhuhai). Guangdong, China. He has been the Zhiyuan Chair Professor with Shanghai Jiaotong University, China. Prior to joining BNU-UIC, he has served as the Deputy Director of the State Key Laboratory of Internet of Things for Smart City, University of Macau. He has over 500 publications in the prestige international journals/conferences and research books and book chapters. He is the Distinguished Member of CCF. He has received the best product awards from the International Science and Tech. Expo (Shenzhen), in 2011/2012, and the 1st Prize of Scientific Research Awards from the Ministry of Education of China, in 2017 (list 2). He has served as an Area Editor for various prestige international journals, the Chair, and a PC Member/Keynote Speaker for many top international conferences.

ENHUA WU (Member, IEEE) received the B.Sc. degree from Tsinghua University, Beijing, in 1970, and the Ph.D. degree from the University of Manchester, U.K., in 1984. He has been working with the State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, since 1985, and a Full Professor with the University of Macau, since 1997. He stayed for teaching with Tsinghua University, until 1980. His research interests include realistic image synthesis, physically-based simulation, and virtual reality. He is a member of ACM and a Fellow of the China Computer Federation (CCF). He has been invited to be the Chair/Honorary Chair/PC Chair recently for various conferences, such as SIGGRAPH Asia16, ACM VRCAI16/18/19, CASA18, and CGI20. He has served as an Associate Editor for TVC, CAVW, *Visual Informatics*, IJIG, and IJSI. He has been an Associate Editor-in-Chief of the *Journal of Computer Science and Technology* (JCST), since 1995.