An Effective CNN Method for Fully Automated Segmenting Subcutaneous and Visceral Adipose Tissue on CT Scans

ZHENG WANG,[1,2] YU MENG,[3] FUTIAN WENG,[1] YINGHAO CHEN,[1] FANGGEN LU,[4] XIAOWEI LIU,[5] MUZHOU HOU (iD),[1] and JIE ZHANG[4]

[1]School of Mathematics and Statistics, Central South University, Changsha 410083, China; [2]School of Information Science and Engineering, Hunan First Normal University, Changsha 410205, China; [3]The Gastroenterology Department of General Hospital, Shenzhen University, Shenzhen 518055, China; [4]The Gastroenterology Department of Second Xiangya Hospital, Central South University, Changsha 410011, China; and [5]The Gastroenterology Department of First Xiangya Hospital, Central South University, Changsha 410008, China

Abstract—One major role of an accurate distribution of abdominal adipose tissue is to predict disease risk. This paper proposes a novel effective three-level convolutional neural network (CNN) approach to automate the selection of abdominal computed tomography (CT) images on large-scale CT scans and automatically quantify the visceral and subcutaneous adipose tissue. First, the proposed framework employs support vector machine (SVM) classifier with a configured parameter to cluster abdominal CT images from screening patients. Second, a pyramid dilation network (DilaLab) is designed based on CNN, to address the complex distribution and non-abdominal internal adipose tissue problems of biomedical image segmentation in visceral adipose tissue. Finally, since the trained DilaLab implicitly encodes the fat-related learning, the transferred DilaLab learning and a simple decoder constitute a new network (DilaLabPlus) for quantifying subcutaneous adipose tissue. The networks are trained not only all available CT images but also with a limited number of CT scans, such as 70 samples including a 10% validation subset. All networks are yielding more precise results. The mean accuracy of the configured SVM classifier yields promising performance of 99.83%, while DilaLabPlus achieves a remarkable performance improvement an with average of 98.08 ± 0.84% standard deviation and 0.7 ± 0.8% standard deviation false-positive rate. The performance of DilaLab yields average 97.82 ± 1.34% standard deviation and 1.23 ± 1.33% standard deviation false-positive rate. This study demonstrates considerable improvement in feasibility and reliability for the fully automated recognition of abdominal CT slices and segmentation of selected abdominal CT in subcutaneous and visceral adipose tissue, and it has a high agreement with a manually annotated biomarker.

INTRODUCTION

The prevalence of obesity around the world is high. Abdominal adiposity is associated with health-related problems and has reached epidemic proportions. Obesity is well-known as a risk factor for significant comorbidity with numerous different diseases, including coronary artery disease, type II diabetes, hypertension, gallstones, pulmonary hypertension, and some cancers.[33,46] Abdominal adipose tissue is further classified as subcutaneous adipose tissue and visceral adipose tissue, and this classification is important in clinical practice. Various studies have demonstrated that the separate measurement of subcutaneous and visceral adipose tissue is crucial for abdominal adipose tissue quantification.[15,42,50] Measurements and assessments are often based on computed tomography (CT), which provides reproducibility for obtaining accurate adiposity-related features of the abdominal section.[37,68] Visceral adipose tissue is directly linked to various diseases.[8] In recent years, some studies have found that the ratio or regional area of visceral and subcutaneous abdominal adipose tissue is a strong independent predictor of the development of numerous

Address correspondence to Muzhou Hou, School of Mathematics and Statistics, Central South University, Changsha 410083, China and Jie Zhang, The Gastroenterology Department of Second Xiangya Hospital, Central South University, Changsha 410011, China. Electronic mails: houmuzhou@sina.com, jiezhang@csu.edu.cn

Zheng Wang and Yu Meng contributed equally to this work.

diseases, which could allow the generation of precise clinical biomarkers to better assist in diagnosis of patients or providing a more accurate prognosis.[6,32,64]

Segmentation of visceral and subcutaneous fat on volumetric CT images is critical to identifying biomarkers for clinical diagnosis prediction and assessment of prognosis. Many therapies based on biomarkers identification have recently been developed for several cancers, including breast,[2] lung,[17] epithelial ovarian,[56] and colorectal.[43] Pioneering works of visceral and subcutaneous fat segmentation performed manual,[73,78] semiautomatic[1,35,44,58,65] and automated segmentation.[13,18,23,27,38,62,70,72] Hence, these methods have a number of limitations:

(1) The segmentation is investigated at a specific region (e.g., at the abdominal region or pelvic regions), or the domain is not mentioned;
(2) Some of models require human intervention (e.g., drawing the abdominal wall, placing a seed, or using semiautomated validation);
(3) Previous studies may not be consistent for various patients (e.g., optimal values of morphological or deformable parameters) or require large amounts of training data;
(4) Although the resolution of the raw CT image is 512×512, the preceding works employed lower resolution CT patches (e.g., 64×64 resolution), which can lead to image distortion and produce inaccurate segmentation in clinic practice;
(5) Some previously published methods are not capable of distinguishing narrow visceral fat associated with SAT or isolated parts of VAT (i.e., positive adipose tissue in this study);
(6) Anatomical-level features are not sufficiently considered (e.g., inconspicuous intestinal capillaries, intestinal cavity, intramuscular fat in this work).

It is challenging to obtain high accuracy using limited raw resolution CT volume-based fully automated segmentation of abdominal fat because of complicated distribution of adipose tissue in viscera and the potential for misclassification due to the intestinal cavity and non-abdominal internal adipose tissue in the abdomen. Precise quantification is critical in biomedical image segmentation to assist the prediction of disease risk. Even marginal segmentation errors may have an adverse impact in clinical settings. In addition, an erroneously segmented mask may cause pathological treatment in subsequent computer-aided diagnosis. Incorrect measurement of visceral fat, e.g., incorrect accumulation, can led to incorrect diagnosis of glucose and lipid metabolic disorders or cardiovascular disease in obese populations.[19,39,47,54] Therefore, achieving the objectives of the study would increase the credibility of medical diagnoses. To accomplish these aims, we sought to develop a effective quantitative framework that provides highly accurate, consistent and automated segmentation for visceral and subcutaneous fat on automatically acquired abdominal CT images from volumetric CT.

Recently, with outstanding improvements in convolutional neural network (CNN) models and high-speed parallel computing unit hardware, state-of-the-art semantic segmentation structures have yielded impressive results, e.g., FCN,[59,61] SegNet,[4] PSPNet,[81] and Deeplab.[10–12] Simultaneously, deep learning techniques have achieved promising performance in a number of computer vision domains. The dense convolution block[26] exploits the potential of convolution through feature reuse, which is easy to train, and has better parameter efficiency. Memorized concatenation[16] incorporates local contextual information between the encoder module and decoder module with additional shortcut skip operation to recover spatial information lost the through downsampling propagation. The pyramid dilated convolution module[22,29,80] can aggregate more effective contextual-level information and enriches it gradually until feature hierarchies are fused. A couple of previous works successfully adopted CNN models or techniques to address biomedical image analysis.[21,67,74,79] For example, Brebisson *et al.*[7] dealt with feature capture at different scales in brain anatomical segmentation of 3D and orthogonal 2D intensity patches; Zhen *et al.*[40] developed an end-to-end deep model using integrated local and global context related-features to predict protein secondary structures; and Jia *et al*[41] proposed Gaussian-pyramid multiscale input and brainstem-region feature enhancement to resolve challenging tasks in brainstem glioma segmentation.

In this paper, the proposed framework consists of three-level CNNs. The first level performs automatic abdominal CT image acquisition on volumetric CT slices using an SVM classifier.[48,60] In the second level, a effective network is developed similar to the encoder-decoder architecture[59,81] (termed *DilaLab*) to segment VAT on the collected CT images from the first step. SAT is the adiposity outside the visceral wall. It would therefore be beneficial if we could leverage the transfer technique (known as transfer learning[51,63]), which transfers the learning from the pretrained *DilaLab* network to the subcutaneous adiposity segmentation task. At the third level, the localization learning of *DilaLab* and the simple decoder constitute a quantification model for SAT on CT images collected in the first stage (termed *DilaLabPlus*). This framework achieves pixel-level accuracy without losing image resolution and increasing network complexity. The

proposed study achieves fully automated segmentation and consistently better performance compared with previous methods as well as more reliable predictions. This computer-aided diagnosis scheme and the evaluation of its performance are presented in the following sections.

MATERIALS AND METHODS

The work proposes a fully automated segmentation scheme for VAT and SAT of automatically selected CT images from volumetric CT slices. The proposed framework contains three components: an SVM classifier, *DilaLab* and *DilaLabPlus*. The process overview of this study is illustrated in Fig. 1. The contributions of the study are mainly in the following areas. First, this paper develops the *DilaLab* network to distinguish narrow parts of visceral fat, inconspicuous intestinal capillaries and intestinal cavity features. *DilaLab*'s architecture is built upon a simple encoder-decoder network and expanded to address prediction limitations in the visceral area. Second, *DilaLab*'s learning is transferred into a newly invented network *DilaLabPlus*, which demonstrates adaptability to quantify the SAT distribution on abdominal CT images. Both investigate the network segments of inner and outer abdominal adipose tissue simultaneously and yield considerably better segmentation accuracy. Lastly, to further fully automate measurement of abdominal adipose tissue for each patient without manually identifying the abdominal CT images, an Support Vector Machine[20,60] classifier is configured. It automatically discriminates abdomen CT images, which are collected for the quantification procedure.

Image Characteristics

Under an image data collection and retrospective study protocol approved by Second XiangYa Hospital of Central South University, we randomly assembled a CT image dataset acquired from patients who received treatment at the same hospital. All used CT images had similar tech specs: scanned by a Siemens Healthcare CT machine (produced 2015–2017); X-ray power output set to 100 kVp; data collected with a diameter of 500 mm; image slice thickness or spacing 1.5 mm; and reconstruction by the Siemens Healthcare Standard image reconstruction kernel in the slice-wise mode.

The CT volume was stored under the DICOM 3.0 protocol. The attenuation values of adipose tissue on CT were in the Hounsfield unit range [− 190, − 30.[77] DICOM files were filtered by the above range to achieve a meaningful result, and irrelevant organs and objects were bypassed. The 512×512 pixel resolution CT volume was converted into a 3-dimensional matrix for further processing. The converted CT images were assigned to the proposed framework for the training and testing dataset, and the split proportion of the dataset is set to 0.1. In the testing phase, CT images from patients were randomly selected and the corresponding predictions were compared with manual annotations performed by skilled individuals.

A adipose tissue from the inner visceral area is surrounded by highly varied human organs, such as the liver, pancreas, spleen and kidney. To inspect adipose tissue in the inner viscera, three frequently encountered issues for abdominal tissue are shown in Fig. 2. The notion of global anatomical-level context is defined, and its importance is argued below:
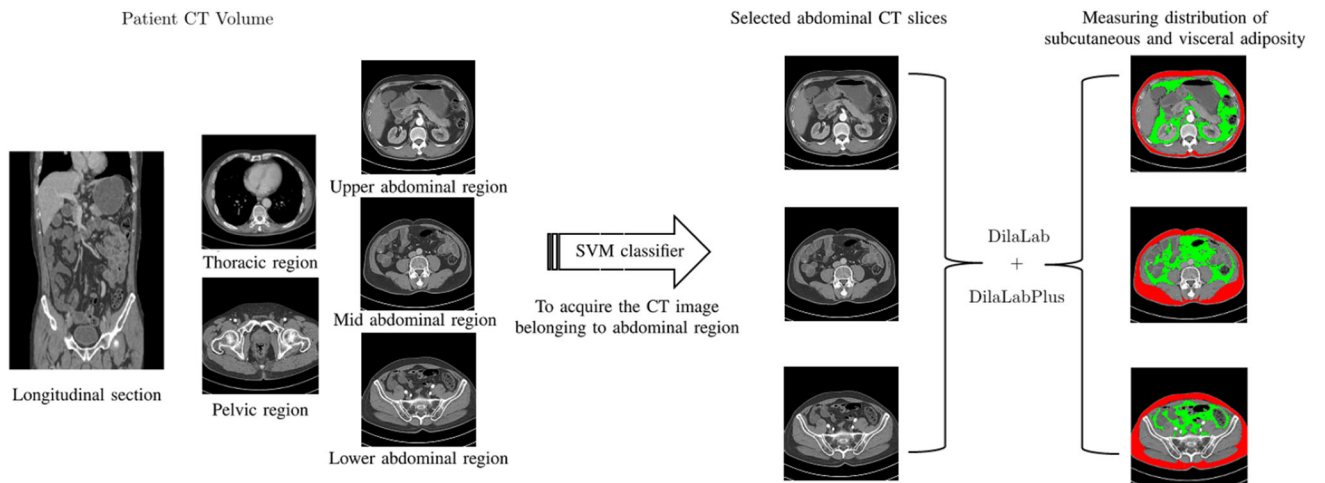

FIGURE 1. Overview of the proposed framework: automated quantification of adiposity on automatically acquired abdominal CT images of volumetric CT slices.
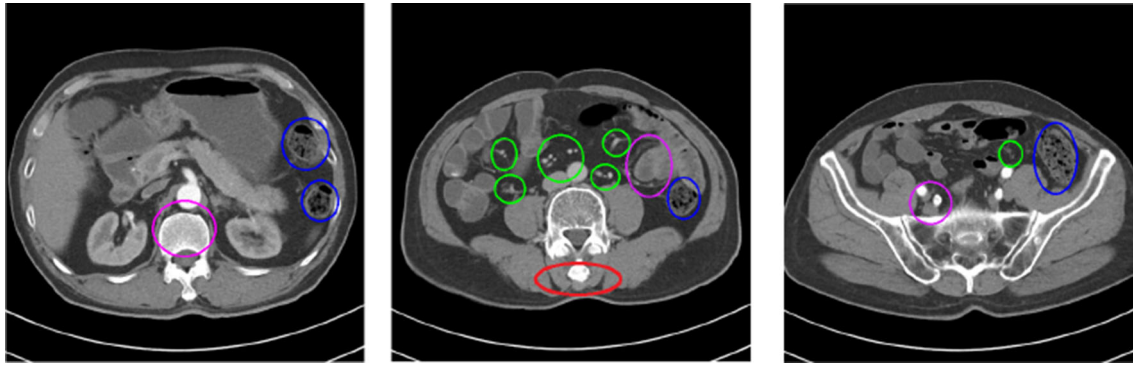
FIGURE 2. Representative issues on CT images.

– *Misclassification* As an example, the blue-outlined tissue is the intestinal cavity, and the inner tissues are easily misclassified as VAT since their pixel-level features are similar to those of VAT. This similarity increases the opportunity for mistaken identification. This study seeks to exploit the capability of anatomical information, which is crucial, especially for understanding the visceral area of abdominal tissue.

– *Complex Distribution* Several narrow parts of visceral fat, e.g., the pink-outlined regions (Gap), are denoted as positive visceral fat and are difficult to find because they are inconspicuous, narrow and isolated. The green-outlined regions (intestinal capillaries tissue) denote negative visceral fat and can be easily misclassified, as the neighborhood effect may surpass the receptive field. This issue arises because intestinal capillaries coexist with inner abdomen fat or VAT is partially isolated at the narrow scale. The complex distribution (i.e., boundaries) of abdominal adipose tissue possesses parts of adipose tissue at the narrow-scale and are an arduous challenge.

– *Non-Abdominal Internal Adipose Tissue* It is important to independently segment abdominal adipose tissue and thus the proposed framework should be suitable to extract non-abdominal internal fat (e.g., intramuscular fat, the red-outlined tissue). *DilaLabPlus* employs the learning of the trained *DilaLab* model and effectively allows global anatomical reconstruction using a simple *decoder*.

– *Abdomen Area Identification* The volumetric CT slices of each patient can be further classified into several subdomains i.e., the longitudinal section, pelvic region and thoracic region and the abdominal region (see columns 1 to 3 in Fig. 1). The proposed approach should 'learn' how to discriminate the volumetric CT slices instead of manually identifying abdominal CT images and collect the extracted abdominal CT images for subsequent automatic segmentation of adipose tissue.

Important Observations

Many semantic segmentation architectures[4,12,59,61,81] have led to a series of breakthroughs for semantic segmentation. This section describes observations for representative segmentation cases. Our objective is to compare the performance of *DilaLab* and state-of-the-art models that are suitable for the challenging task of VAT segmentation of abdominal CT images. For this comparison, we also reproduce several state-of-the-art networks, such as FCNs,[61] SegNet,[4] PSPNets,[81] and DeepLabV3 + .[12]

Base on the observation of visceral fat segmentation (shown in Fig. 3), PSPNets works more efficient and produces competitive segmentation, although with a positive boundary loss. The results of FCNs are clearly underdeveloped, even after losing the smooth portion and positive boundaries, giving rise to a negative boundary and undesired intestinal cavity. SegNet and DeepLabV3 + are the deepest models with the highest time consumption, but the inclusion of the intestinal cavity is undesired in the results. To summarize these results, typical failure cases are partially or completely related to the anatomic relationship and local information of multiscale receptive fields. The experiments show that these pioneering models are inefficient for the task. *DilaLab* shows a significant performance improvement and superior ability to differentiate complex boundaries and the undesired intestinal cavity, compared with other large models. The proposed *DilaLab* with a suitable local-anatomic-level prior can achieve a remarkable performance improvement in anatomical localization parsing.

According to these observations the loss and misclassification of inner abdomen fat provide no benefit
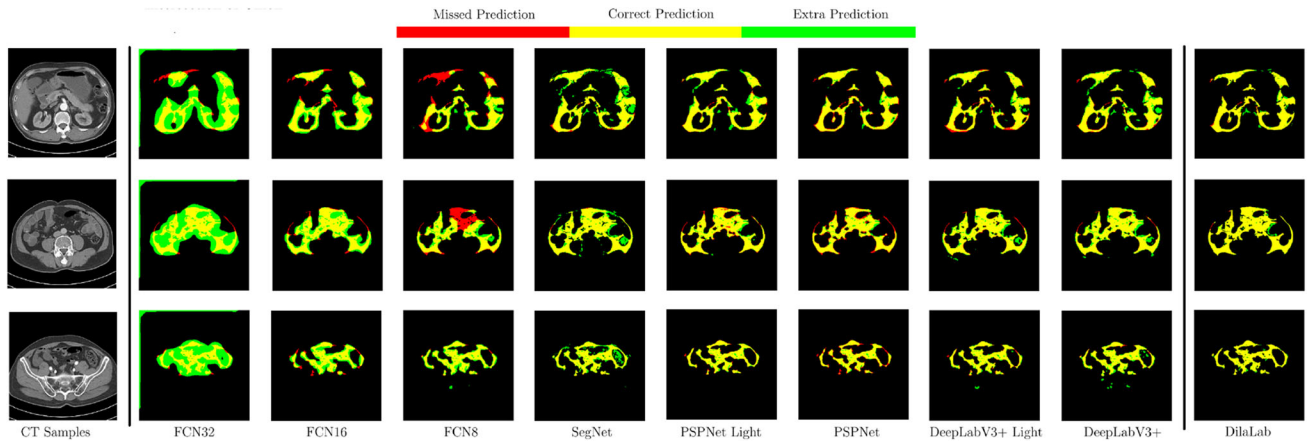
FIGURE 3. Segmentation results on abdominal CT samples for visceral fat. Yellow depicts the correctly predicted visceral fat segmentation, red demonstrates the missed visceral fat prediction, and green indicates extra predictions of visceral fat.

for prediction. Therefore, the complex distribution and non-abdominal internal adipose tissue must be distinguished. For this reason, *DilaLab* is introduced in Sect. 2.4 to perform an effective local-anatomic-level semantic segmentation.

SVM Classifier

The proposed framework should automatically select abdominal CT images from volumetric CT slices for each patient rather than requiring manual identification by an observer. Specifically, the abdomen area was placed below the lung area (i.e., upper bound of interest) and above the femur (i.e., lower bound of interest). CT images between the two bounds were predicted as positive; images outside of these bounds were predicted as negative. Conventional machine learning approaches are employed as classifiers, due to the consideration of efficiency and applicability. Many effective classifiers have been used for the field of medical image analysis. Decision Tree[57] method has been proposed to classify pre-diagnosed database of brain images and achieved 95% accuracy. *Kim et al.*[31] developed a novel X-ray image classification approach using Random Forest, it has obtained the best precision percentage of 93.1% on IRMA 2017 dataset. *Robert et al.*[49] used AdaBoost algorithm to overcome challenging tasks in automated classification of lung bronchovascular anatomy from chest CT images. These types of classifier have been demonstrated to perform well in various medical image classification tasks.

In addition, the support vector machine[66] (i.e., SVM) method has achieved impressive performance in images classification,[48,60,76] e.g., in breast cancer histopathology image classification with 96.19% accuracy.[71] The features Histogram of Oriented Gra-

dients[14] (i.e., HOG) is a common feature to describe local image texture in the field of computer vision and pattern recognition, and has been successfully applied in computer vision community[5,9,36,69], such as classification of breast mass,[55] which presented 97.7% precision. Hence, we employed an efficient computer version algorithm that combines the HOG descriptor and the SVM classifier (termed SVM-HOG in this study). HOG SVM-HOG is popular to exploit due to the small number of parameters to tune, and the capable to simultaneously classify multiple categories. Under an orientation with 9, 8 pixels per cell and 2 cells per block, the descriptor convert the CT slices into HOG feature data. The support vector machine adopts a one-against one strategy with 10 as the optimum value of the error penalty C[52] through a radial basis function.

For identification, an SVM is trained with a set of HOG features data, which convert CT slices belonging to separate categories (i.e., longitudinal section, pelvic and thoracic and abdomen), builds a model and subsequently processes the automatic identification of abdominal CT images using that model. The implemented SVM-HOG successfully distinguishes whether the input CT image belongs to the abdomen or not. The accuracy obtained from the validation data agrees with the parameter selection. For an objective comparison, the other approaches compared with SVM-HOG are also implemented through scikit-learn in python.[53] The max depth tree and number of base learners were set as 5 and 10 in Random Forest method respectively. Decision Tree also obtained 5 as the optimum value of max depth tree. And the AdaBoost method adopted decision tree as base classifier with the optimum value of base learners 10. The other parameters are gained by default values. The comparison between different classifier (In Table 1) shows

TABLE 1. Comparison with different classifiers.

Classifier	Training time (s)	Train score	Test accuracy	Classification report			
				Precision	Recall	f1-score	Support
Decision tree + HOG	25.96	0.9808	0.9808	0.98	0.98	0.98	1100
Random forest + HOG	**17.22**	0.8269	0.8269	0.88	0.83	0.82	1100
AdaBoost + HOG	193.07	0.6346	0.6346	0.49	0.63	0.52	1100
SVM-HOG	55.40	**0.9983**	**0.9983**	**1**	**0.99**	**0.99**	**1100**

that, the proposed SVM-HOG yields a promising performance with a mean accuracy of 99.83% and the bold values depict the best performance of the corresponding metrics. It also outperforms than other classifiers in terms of train score, and precision, recall f1-score in 1100 testing samples. The acquired abdominal CT images can be used for fully automated segmentation of VAT and SAT.

The developed framework and the proposed tasks are optimized together in an end-to-end approach. With successive forward propagation, it can simultaneously acquire the abdomen's volume, and generates the subcutaneous and visceral abdominal fat distribution.

DilaLab Network

In this section, *DilaLab* is proposed for completely automated inner abdomen fat segmentation. This network is modified and extended on a more elegant encoder-decoder architecture inspired by Fisher *et al.*[80] and Ronneberger *et al.*[59] DilaLab feasible for the anatomical context and local information of multiscale receptive fields. *DilaLab* consists of 13 convolutional blocks that are based on an encoder-center-decoder structure. A 512×512 resolution CT image is used as the input. The encoder module in *DilaLab* performs the contraction, the center module assembles the enlarged feature maps, which convolve with the dilation factor at a diverse rate, and the decoder module expands its input. Figure 4 illustrates the *DilaLab* architecture schematically. *DilaLab* employs a pyramid dilation module (middle) with a contracting encoder module (left side) and an expanding decoder module (right side).

The contracting encoder module consists of 6 *encoder* blocks, which are built based on a dense architecture. Each *encoder* contains repeated application of batch normalized convolution and nonoverlapping Max-Pooling. The convolution repetition aims to exploit the potential of feature channels to capture important local contextual information. The convolution is implemented with the same padding operation, an element-wise rectified-linear nonlinearity unit (ReLU) activation,[75] 3×3 convolution kernel and

feature channels (also known as filters). The initial value of the filter is set to 16. the batch normalization operation[28] reduces the internal covariance and accelerates the process dramatically. Max-Pooling uses 2×2 size with a stride of 2 to downsample the higher resolution. The number of filters is doubled at each contracting encoder step. The contracting encoder is fed from the input of the network and propagates its outcomes to the middle module.

To aggregate anatomically contextual information from a multiscale receptive field, this study designs a pyramid dilation module by leveraging dilated convolution. Dilated convolution is suitable for assembling multiscale anatomical contexts without losing resolution and for increasing the exponential growth of the feature map without increasing the number of parameters. A dilated convolution applies 512 filters, a padding operation and a ReLU activation. The dilation rate must be equal to a power of 2 (i.e., 1, 2, 4, 8.). This pyramid dilation module sequentially connects four dilated convolutions with diverse dilation rates. The densely anatomical-level aggregation promote the segmentation accuracy and performance. The assembled feature maps under four different dilation scales are summed up to feed the expanding decoder module.

Every appropriate *decoder* in the expanding decoder module upsamples the fed feature map with size 2×2 (black box in Fig. 5). Furthermore, the convoluted output of the *encoder* (green box in Fig. 5) and the upsampled feature maps of the corresponding *decoder* are concatenated (dark gray objects in Fig. 5). The concatenated result convolves two times with same padding operation, a ReLU activation, 3×3 convolution kernel and relevant feature channels from the filter bank. The expanding decoder module gradually reconstructs the lower resolution to higher resolution. The concatenating layer at each expanding decoder module is indispensable because it processes the contextual and localization information together.

DilaLabPlus Model

Precisely determining subcutaneous adipose tissue is crucial for quantifying the outer adipose tissue in vis-
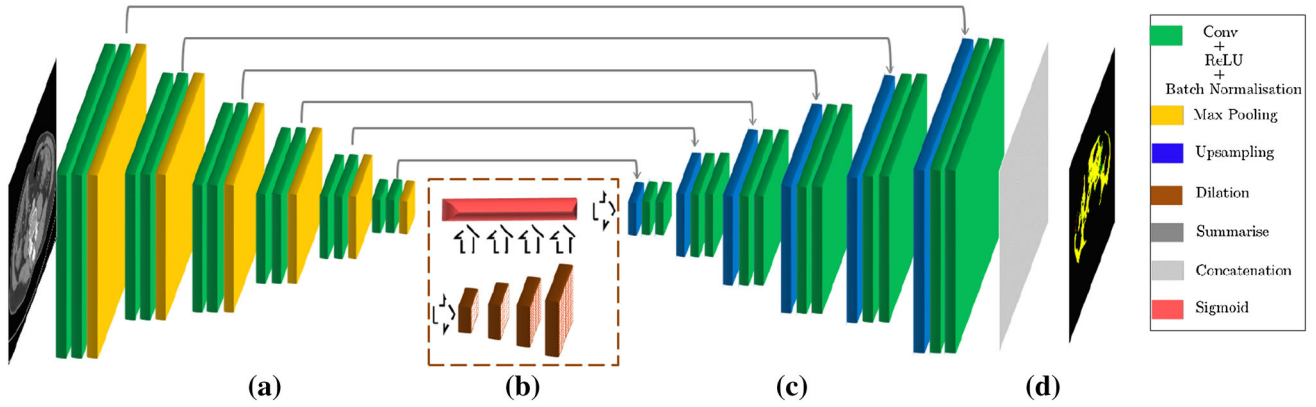
FIGURE 4. An illustration of the proposed *DilaLab* network. The *DilaLab* network is based on an encoder-decoder architecture and consists of a densely convolutional block, pyramid dilation module and memorized concatenation. It performs convolution with a filter bank to enlarge the feature map. (a) By supplying successive blocks, the contracting subnetwork (left side) downsamples the resolution and propagates the feature maps. (b) The pyramid dilation module (middle) increases the anatomic context information. (c) Subsequently, the expanding subnetwork (right side) upsamples and propagates concatenated local context information. (d) Last but not least, at the final layer the output from the expanding subnetwork is fed into a sigmoid classifier to predict the final segmentation.
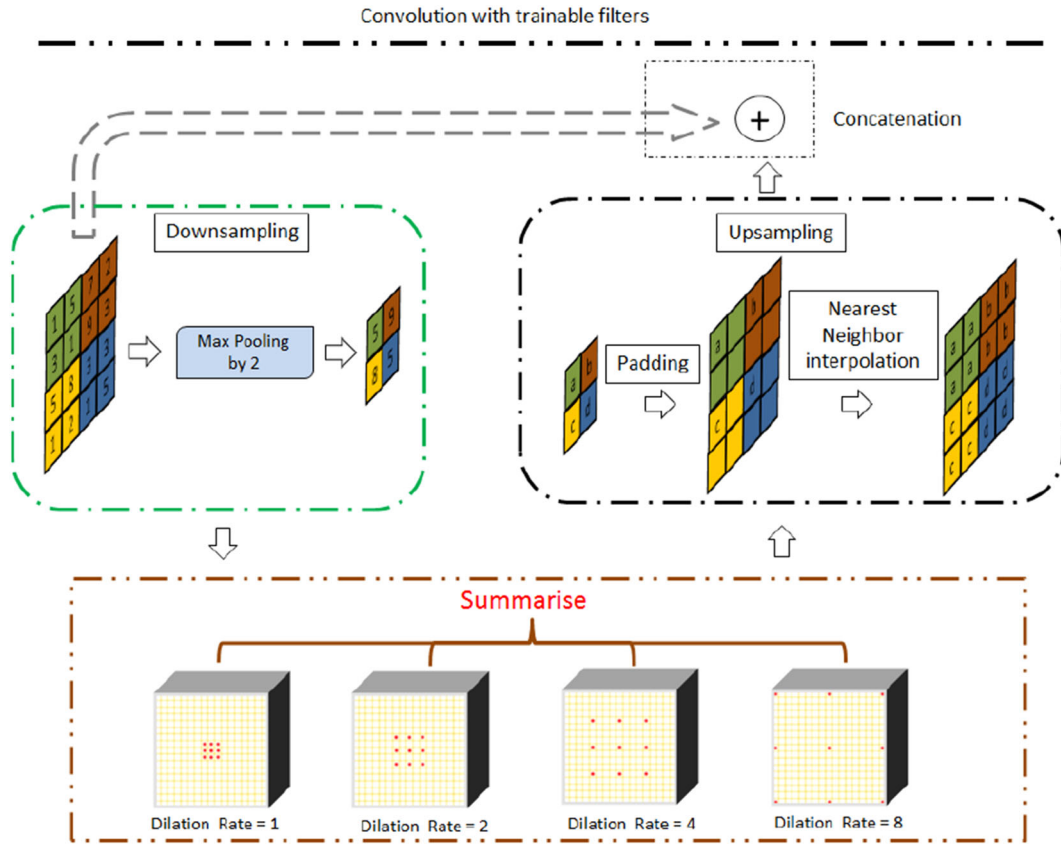

FIGURE 5. An example of summaries, upsampling and concatenation. The brown box illustrates the summary of the feature maps, which are convoluted outcomes with diverse dilated receptive fields. The black box illustrates the upsampling strategy, *a, b, c and d* corresponding to values in a feature map. The green and black boxes are input vectors for the concatenation operator. the black arrow propagates the feature maps forward.

cera. *DilaLabPlus* utilizes a suitable strategy to provide a clear representation of adipose tissue compared to *DilaLab* (see Fig. 6).

We found that the *DilaLab* model lacks a suitable strategy for SAT segmentation that utilizes global anatomical-level clues, leading to the inclusion of non-

abdominal internal adipose tissue (as described in Sect. 2.1). This problem should be addressed because the elimination of non-abdominal internal adipose tissue would improve credibility from a clinical perspective and is crucial for increasing the precision of outer abdomen fat segmentation. Furthermore, the *DilaLab* architecture contains more than 30 million parameters, which is overdeveloped for subcutaneous fat segmentation. In contrast to local anatomical-level segmentation, subcutaneous fat perception favors the incorporation of global and smooth feature information. One method for addressing this issue is to leverage learning from *DilaLab*, which is a technique called transfer learning. In accordance with our task, transfer learning is used to utilize the same type between the source and target.

In this section, the segmentation task is to incorporate suitable global anatomical-level features. We use the *DilaLabPlus* network to quantify subcutaneous adipose tissue with an effective transfer learning algorithm (as schematically depicted in Fig. 7), rather than training the full network from scratch. *DilaLabPlus* is implemented according to the following steps:

(i)   The assembled learning of the first dilated convolution from the pretrained *DilaLab* is transferred as the training input for *DilaLabPlus*.

(ii) The transferred learning propagates the aggregated localization information to a simple yet effective *decoder* module, which reconstructs the global and smooth context information for precisely quantifying subcutaneous abdominal tissue.

(iii) The learning from the final convolution of each *DilaLab encoder* is transferred to concatenate the feature map from each corresponding step of the expanding *decoder* so that general feature information can be efficiently captured.

Table 2 compares floating point operations (FLOPs), number of parameters (params), train rate (milliseconds per step), train time (seconds) and test rate (seconds per image) between DilaLab and DilaLabPlus. For FLOPs and parameters, DilaLabPlus decreases over 70% than DilaLab. Meanwhile, it processes images faster than DilaLab. It is a situation, *DilaLabPlus* works much more efficiently to recognize global anatomical-level information without loss of border pixels and shortens the training duration on abdominal CT images.
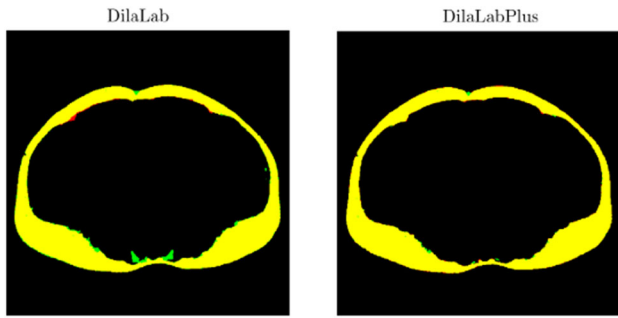

FIGURE 6. Representative prediction of subcutaneous adipose tissue. The yellow color depicts the section of overlap between the prediction and ground truth. Extra prediction: green. Missed ground truth: red.

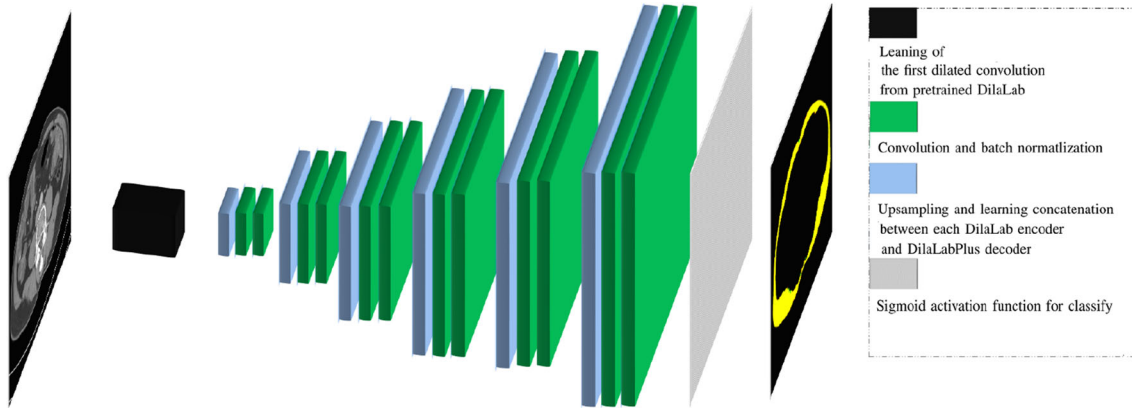FIGURE 7. Schematic depiction of the DilaLabPlus network. The *DilaLabPlus* model performs the learning of the first dilated convolution from the pretrained *DilaLab* and feeds into a *decoder* module. Each step of the expanding *decoder* concatenates the learning from the final convolution of the corresponding *DilaLab encoder*, which leads to significantly improved results for SAT segmentation.

TABLE 2. Runtime comparison.

Model	Model analysis report				
	FLOPs	Params	Train rate (ms/setp)	Train time (s)	Test rate (s/im)
DilaLab	13.8604×10^8	6.8253×10^8	73	683.0077	4.6741
DilaLabPlus	3.9424×10^8	1.9695×10^8	32	455.6625	5.8758

EXPERIMENTS

We initially obtained 5000 CT images(780 from the longitudinal section, 1206 from the abdominal region, 1489 from the pelvic region and 1525 from the thoracic region) from 20 random patients. This CT dataset is used to train the proposed networks as well as the comparative methods.

Implementation Details

In this study, to make the results to be more reliable, we split the CT dataset into two groups: training set (included validation set) and testing set. Training set contains total of 3900 images, which has 600 from the longitudinal section and 1100 from each category of the rest. Based on the train_test_split method[25] in scikit-learn module,[53] validation set is randomly selected from training set and is only done on the training phase. In details, we first shuffle the training set, then split the whole training examples into training set and validation set by the ratio size (e.g., test_size with 0.1 in this study) using train_test_split method. Therefore, the validation subset contains 390 images. And the rest of 1100 images from experimental data are supplied to testing set.

The proposed three level network-constituted framework is used to perform the training on a Dell xps8930 server (hexa-core 3.20 GHz processor, 16 GB RAM and one NVIDIA GeForce GTX 1070 video card). The proposed methods were implemented in Python using the Keras framework with a TensorFlow backend.

The proposed and mentioned networks are trained by the RMSprop optimizer.[24] RMSprop experiments with a momentum term γ of 0.9, while good suggested value for the learning rate is η is 0.001. The form of RMSprop (Eq. 1) is derived as follows:

$$\begin{aligned} E[g^2]_t &= \gamma E[g^2]_{t-1} + (1-\gamma)g_t^2 \\ \theta_{t-1} &= \theta_t - \frac{\eta}{\sqrt{E[g^2]_t + \epsilon}} g^2 \end{aligned} \qquad (1)$$

The feature outcome from the implemented methods is fed to a trainable nonlinearity sigmoid classifier. The prediction probability function is given by Eq. 2-1 with a parameter vector θ for every data vector x, and set $F_{sigmoid}$ (Eq. 2-2) is the logistic of the objective function w.r.t. the parameter z. Finally, the prediction probability $\hat{y}(x)$ simply replaces the z variable of the sigmoid to obtain the binary classification function (Eq. 2-3):

$$\hat{y}(x) = \theta_0 x_0 + \theta_1 x_1 + \cdots + \theta_n x_n \qquad (2-1)$$

$$F_{sigmoid}(z) = \frac{1}{1 + e^{-z}} \qquad (2-2)$$

$$F_{sigmoid}(x) = \frac{1}{1 + e^{-(\theta_0 x_0 + \theta_1 x_1 + \cdots + \theta_n x_n)}} \qquad (2-3)$$

With the sigmoid classifier, the prediction could be interpreted as a probability between 0 and 1 to obtain the actual segmentation. The implemented networks are monitored by minimizing the Dice Similarity Coefficient loss function (L_{dsc})[34] through 10 epochs of training split to 200 steps. The L_{dsc} for an output X with ground truth Y derived from the binary case is given by Eq. 3:

$$L_{dsc}(X, Y) = 1 - \frac{2 \sum_{j=1}^n x_i y_i}{\sum_{j=1}^n x_i + \sum_{j=1}^n y_i} \qquad (3)$$

where $x_i \in X$ denotes the *i*th flattened prediction probability through a sigmoid classifier, and y_i is indicative of the flattened corresponding ground truths of Y. As an objective function, L_{dsc} does not consider the errors of false prediction (as shown in Fig. 2). To improve generalizability, we introduce auxiliary metrics to benchmark these errors, such as false-positive rate (R_{fp}), false-negative rate (R_{fn}) and specificity, which supplement the L_{dsc} function. These metrics are more biased towards accuracy. These metrics are given by the following formula (Eq. 4):

$$\begin{aligned} R_{fp}(X, Y) &= \frac{x_i - x_i y_i}{\hat{y}_i} \\ R_{fn}(X, Y) &= \frac{\hat{x}_i - \hat{x}_i \hat{y}_i}{y_i} \end{aligned} \qquad (4)$$

where $\hat{x}_i \in X$ denotes the negative prediction probability and $\hat{y}_i \in Y$ is the true negative label. The detailed description of the metrics refers to.[3,45] In the training phase, image augmentation is executed on those CT

images through five types of transformation operators according to the following transformation matrix (Eq. 5):

$$T_{\text{flip}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$T_{\text{zoom}} = \begin{bmatrix} z_x & 0 & 0 \\ 0 & z_y & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$T_{\text{shear}} = \begin{bmatrix} 1 & R\sin(\theta) & 0 \\ 0 & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$T_{\text{shift}} = \begin{bmatrix} 1 & 0 & r \times s_x \\ 0 & 1 & c \times t_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$T_{\text{rotation}} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

z_x, z_y of T_{zoom} was uniformly sampled from (0.05, 1). T_{shear} involves the variables R and θ. R is sampled as -1 or 1 with probability $\frac{1}{2}$, and θ is uniformly sampled in range $(-0.05, 0.05)$. In the transformation matrix T_{shift}, where r is the row number of the image and c is the column number of the image, s_x and t_y are uniformly sampled in the range $(-0.05, 0.05)$. θ in T_{rotation} is uniformly sampled from the range $(-0.05, 0.05)$.

Results

To further analyze the proposed framework, we conduct experiments for recently widely adopted semantic segmentation models. The segmentation results for the selected methods are shown in Fig. 3.

As listed in Table 3, the bold line corresponds to the proposed *DilaLab*, which achieves more promising results than the other models. Its predicted masks outperform the conventional manual annotation. This proves the ability of the *DilaLab* network to parse local-anatomical features from the input image and predict it precisely and smoothly. *DilaLab* achieved a mean pixel-wise prediction accuracy of 97.82% with a standard deviation 1.34%, a false-positive rate for mean with 1.23%, a false-negative rate for mean with 0.75% and a specificity for mean with 98.36%.

In Table 4, the bold number denotes the performance of *DilaLabPlus*, which works better than the other approaches, such as[4,12,61,81] and *DilaLab*. The mean pixel-wise prediction accuracy is equal to 98.08% with a standard deviation 0.63%, a false-positive rate for mean with 0.7%, a false-negative rate for mean with 0.53% and a specificity mean with 99.43%.

The proposed networks successfully and effectively identified potential abdominal adipose tissue between the visceral wall. In Table 5, we compare the average automated segmentation results of the proposed networks and the previous methods,[27,72] which shows that the performances of *DilaLab* and *DilaLabPlus* are statistically better than the other methods tested. Several examples of segmentation results generated by the proposed framework are illustrated in Fig. 8.

With Limited Amount of Training Data

To evaluate the stability of the proposed scheme, the following experiments were performed with *DilaLab* and *DilaLabPlus*. In the experiments, an extremely small training subset containing 70 randomly selected images from the abdomen category was used, including a 10% validation subset images. After 10 epochs (iterations through the training subset), the training was stopped due to nonconvergence of further improvement in accuracy. The results from the differ-

TABLE 3. Quantitative performance report for VAT.

Networks	Evaluation of VAT			
	Pixel-wise accuracy	False-positive rate	False-negative rate	Specificity
FCN32	0.8412 ± 0.0320	0.3305 ± 0.0329	0.0986 ± 00421	0.8483 ± 0.0298
FCN16	0.9139 ± 0.0340	0.2980 ± 0.0249	0.1735 ± 0.0416	0.8706 ± 0.0211
FCN8	0.9556 ± 0.0349	0.2577 ± 0.0255	0.1551 ± 0.0306	0.9301 ± 0.0212
SegNet	0.9577 ± 0.4119	0.0558 ± 0.0513	0.3089 ± 0.0300	0.9509 ± 0.0307
PSPNet light	0.9692 ± 0.0270	0.0333 ± 0.0287	0.0306 ± 0.0231	0.9734 ± 0.0216
PSPNet	0.9137 ± 0.0026	0.0625 ± 0.0027	0.0934 ± 0.0024	0.9137 ± 0.0026
DeepLabV3+ light	0.9789 ± 0.0138	0.0293 ± 0.0374	0.0424 ± 0.0355	0.9823 ± 0.0151
DeepLabV3+	0.9788 ± 0.0135	0.0213 ± 0.0129	0.0300 ± 0.0203	0.9786 ± 0.0151
DilaLab	0.9782 ± 0.0134	0.0123 ± 0.0133	0.0075 ± 0.0023	0.9836 ± 0.0191

TABLE 4. Numerical performance report for SAT.

Networks	Evaluation of SAT			
	Pixel-wise accuracy	False-positive rate	False-negative rate	Specificity
FCN32	0.9280 ± 0.0240	0.0765 ± 0.0259	0.0632 ± 0.0046	0.9321 ± 0.0224
FCN16	0.8755 ± 0.0021	0.2398 ± 0.0064	0.0654 ± 0.0057	0.8721 ± 0.0061
FCN8	0.8756 ± 0.0018	0.2499 ± 0.0013	0.0848 ± 0.0262	0.8731 ± 0.0010
SegNet	0.9727 ± 0.0254	0.0272 ± 0.0239	0.0249 ± 0.0360	0.9728 ± 0.0265
PSPNet light	0.9355 ± 0.0732	0.0960 ± 0.0881	0.1383 ± 0.1078	0.9415 ± 0.0789
PSPNet	0.9401 ± 0.0730	0.0598 ± 0.0730	0.1042 ± 0.0953	0.9401 ± 0.0730
DeepLabV3+ light	0.9878 ± 0.0110	0.0186 ± 0.0285	0.0327 ± 0.0378	0.9901 ± 0.0056
DeepLabV3+	0.9875 ± 0.0102	0.0128 ± 0.0108	0.0214 ± 0.0188	0.9874 ± 0.0113
DilaLab	0.9784 ± 0.0161	0.0088 ± 0.0143	0.0061 ± 0.0032	0.9903 ± 0.0319
DilaLabPlus	0.9833 ± 0.0084	0.0070 ± 0.0084	0.0053 ± 0.0023	0.9943 ± 0.0100

TABLE 5. Automated segmentation results on abdominal adipose tissue.

Methods	Evaluation of VAT		
	Pixel-wise accuracy	Dice coefficient	Pearson correlation
Wang et al.[72]	0.9682 ± 0.0218	0.9150 ± 0.0624	×
Hui et al.[27]	×	×	0.7916 ± 0.0121
DilaLab	0.9782 ± 0.0134	0.9278 ± 0.0158	0.9083 ± 0.0546

Methods	Evaluation of SAT		
	Pixel-wise accuracy	Dice coefficient	Pearson correlation
Wang et al.[72]	0.9682 ± 0.0218	0.9797 ± 0.0145	×
Hui et al.[27]	×	×	0.9446 ± 0.0026
DilaLabPlus	0.9833 ± 0.0084	0.9808 ± 0.0063	0.9688 ± 0.0220

ent training data setups are presented in Fig. 9a for the *DilaLab* learning curve and Fig. 9b for the *DilaLabPlus* learning curve. The comparison illustrates that, with 70 training images, the performance is very similar to what was obtained with all available training samples (as shown in Tables 6 and 7).

The results of *DilaLabPlus* yielded a mean pixel-wise prediction accuracy of 98.23% with a standard deviation of 0.63%, a false-positive rate for mean with 0.55%, a false-negative rate for mean with 0.57% and a specificity for mean with 99.95%. *DilaLab*'s performance achieved a mean pixel-wise prediction accuracy of 97.56% with a standard deviation of 2.29%, a false-positive rate for mean with 1.49%, a false-negative rate for mean with 0.78% and a specificity for mean with 98.91%. The experimental results indicate a tendency towards stability even for an extremely small training subset. The proposed framework represents a generalized platform that can be applied for fully automated segmentation of visceral and subcutaneous adipose tissue on CT volumes in abdominal sites.

DISCUSSION

The completely automated segmentation of visceral adipose tissue and subcutaneous adipose tissue data from CT volumes is an exceedingly challenging task due to the large variability of adipose tissue in the visceral area as well as the potential misclassification of inner and outer abdominal fat. In addition, precise adipose tissue segmentation can perform as a qualified bio-marker in clinical practice. Factors such as loss of adipose tissue, abnormal adipose tissue distribution and mutated adipose tissue have diagnostic and therapeutic implications for patients. Previous research in this area required individual participation, such as manual selection, placement of segmentation seeds, boundary conditions, and designation of viscera boundaries or small patches. In this study, we proposed a completely automated segmentation framework and demonstrated its feasibility for overcoming these challenges without any artificial participation and subjectivity. The effective framework possesses unique characteristics as outlined below.
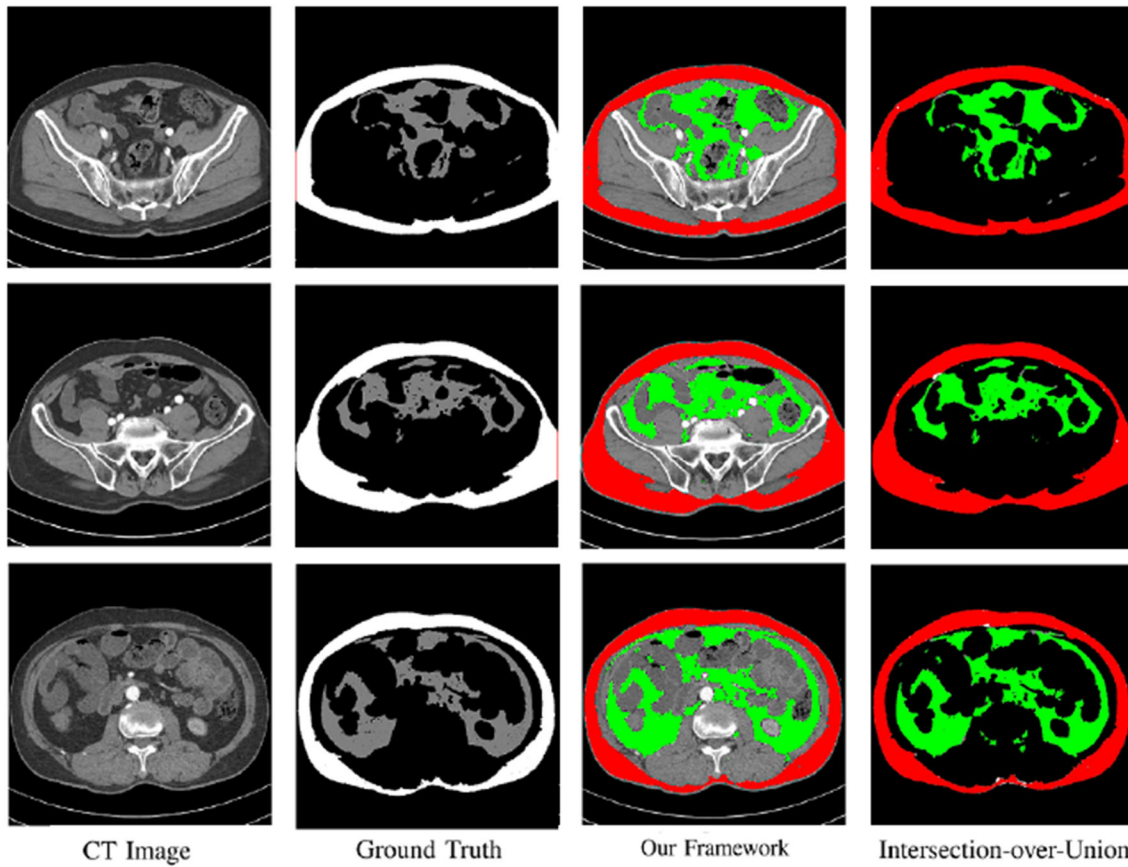
FIGURE 8. Example of overlap between ground truth and segmentation in VAT and SAT. VAT in segmentation: red; SAT in segmentation: green; VAT in ground truth: gray; SAT in ground truth: white.

First, our SVM-HOG was developed for automated abdominal CT image acquisition. This classifier has proven capable of reviewing immense amounts of CT images and generating superior classification accuracy compared to the results of intervention (i.e., achieving a classification accuracy of 99.83%.). Therefore, SVM-HOG provides reliability and is capable of processing large-scale datasets, especially for biomedical image classification and biomedical data analysis with high efficiency.

Second, the previously developed schemes might be not sensitive to local-anatomical-level information (e.g., non-abdominal internal adipose tissue, complex distribution and tissue misclassification). The visceral area is key aspect of these issues. To solve these problems, we developed an effective CNN, namely *DilaLab*, that is based on an encoder decoder network, and it adopts a pyramid dilation summary module, and feature maps concatenation operation, convolution repetition and special upsampling, etc. Consequently, *DilaLab* provides a highly qualitative approach by detecting the undesired intestinal cavity and recognizing complex boundaries (i.e., finding the positive and excluding the negative).

Third, this paper formulated the *DilaLabPlus* network for the subcutaneous adipose tissue segmentation task. *DilaLabPlus* adopts transfer learning by considering *DilaLab* learning, which increases the local coordinates and neighborhood information for the outcome. To precisely categorize subcutaneous adipose tissue (particularly easily misclassified tissues), the decoder subnetwork was redesigned with a simple decoder structure that is capable of assembling more general output information.

In summary, our effective framework involves three classifier modules, the SVM-HOG classifier, *DilaLab* model and *DilaLabPlus* network. The objective is to realize a completely automated segmentation of abdominal adipose tissue without artificial participation, subjectivity and local-anatomical-level loss. Moreover, our proposed framework is much more productive in terms of computation and memory. This framework involves a sequential two-step process that includes (1) the acquisition of abdominal CT slices from screening patients and (2) automatic segmentation of visceral adipose tissue and subcutaneous adipose tissue from each acquired CT image. The experimental results further demonstrate the feasibility

(a)

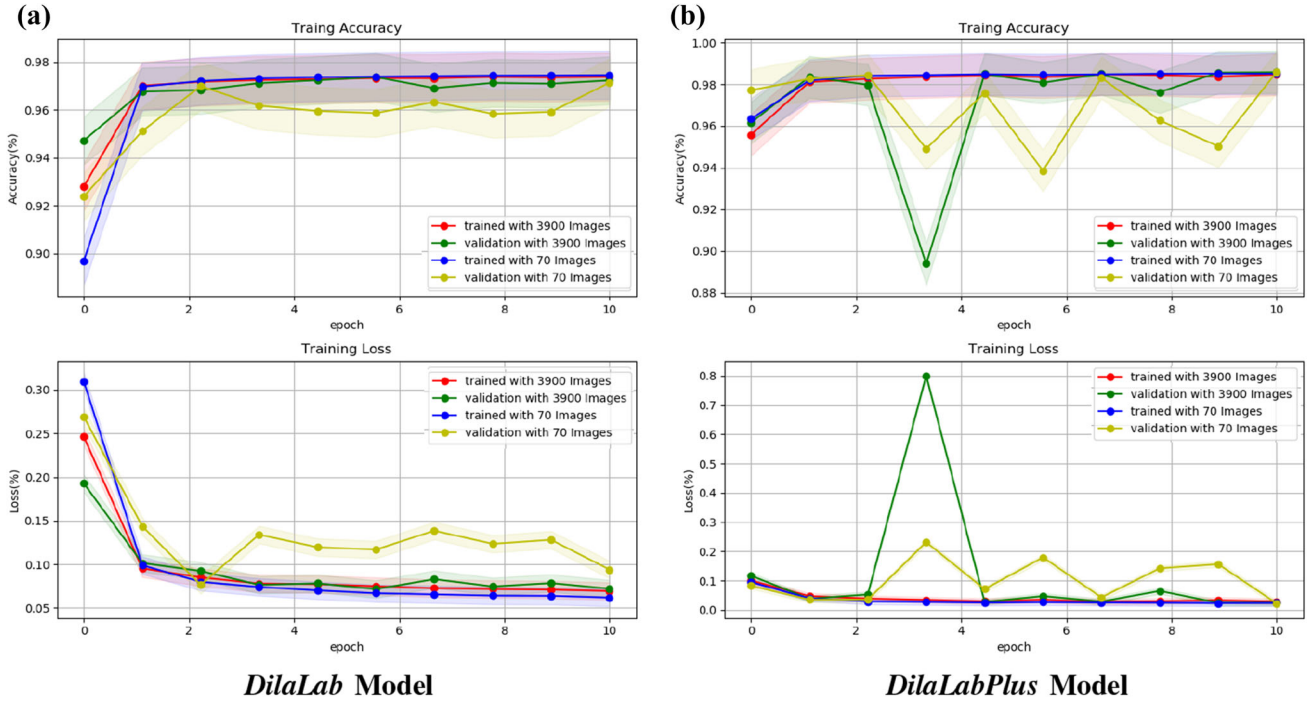**(b)**

DilaLab Model   *DilaLabPlus* Model

FIGURE 9. Plot showing performance in the training and validation subsets. Accuracy and Loss are plotted against the training step for visceral adipose tissue (a) and subcutaneous adipose tissue (b) respectively, during the training length of the binary-class classifier over the course of 2000 steps. The first row represents accuracy performance, and the last row depicts loss qualities. Training with 3900 datasets: red. Validation with 3900 datasets: green. Training with 70 images: blue. Validation with 70 images: yellow.

TABLE 6. Evaluation of DilaLab with different training samples.

Number of samples	Evaluation of visceral adipose tissue			
	Pixel-wise accuracy	False-positive rate	False-negative rate	Specificity
1100 Samples	0.9782 ± 0.0134	0.0123 ± 0.0133	0.0075 ± 0.0023	0.9836 ± 0.0191
70 Samples	0.9756 ± 0.0229	0.0149 ± 0.0240	0.0078 ± 0.0017	0.9891 ± 0.0174

TABLE 7. Evaluation of DilaLabPlus with different training samples.

Number of samples	Evaluation of subcutaneous adipose tissue			
	Pixel-wise accuracy	False-positive rate	False-negative rate	Specificity
1100 Samples	0.9833 ± 0.0084	0.0070 ± 0.0084	0.0053 ± 0.0023	0.9943 ± 0.1001
70 Samples	0.9823 ± 0.0063	0.0055 ± 0.0057	0.0057 ± 0.0011	0.9995 ± 0.0138

and superior accuracy of our framework for completely automated segmentation of visceral adipose tissue and subcutaneous adipose tissue from CT volumes. This framework can facilitate abdominal adipose tissue biomarker processing in future clinical practice and could provide a good method for assisting patient diagnosis and appropriate treatment.

APPENDIX A: EXPERIMENTS ON NIH DEEPLESION

We further report segmentation of visceral and subcutaneous adipose tissue results on the DeepLesion dataset.[30] This dataset was collected from the PACS of a major medical institute. It contains over 30K radio-
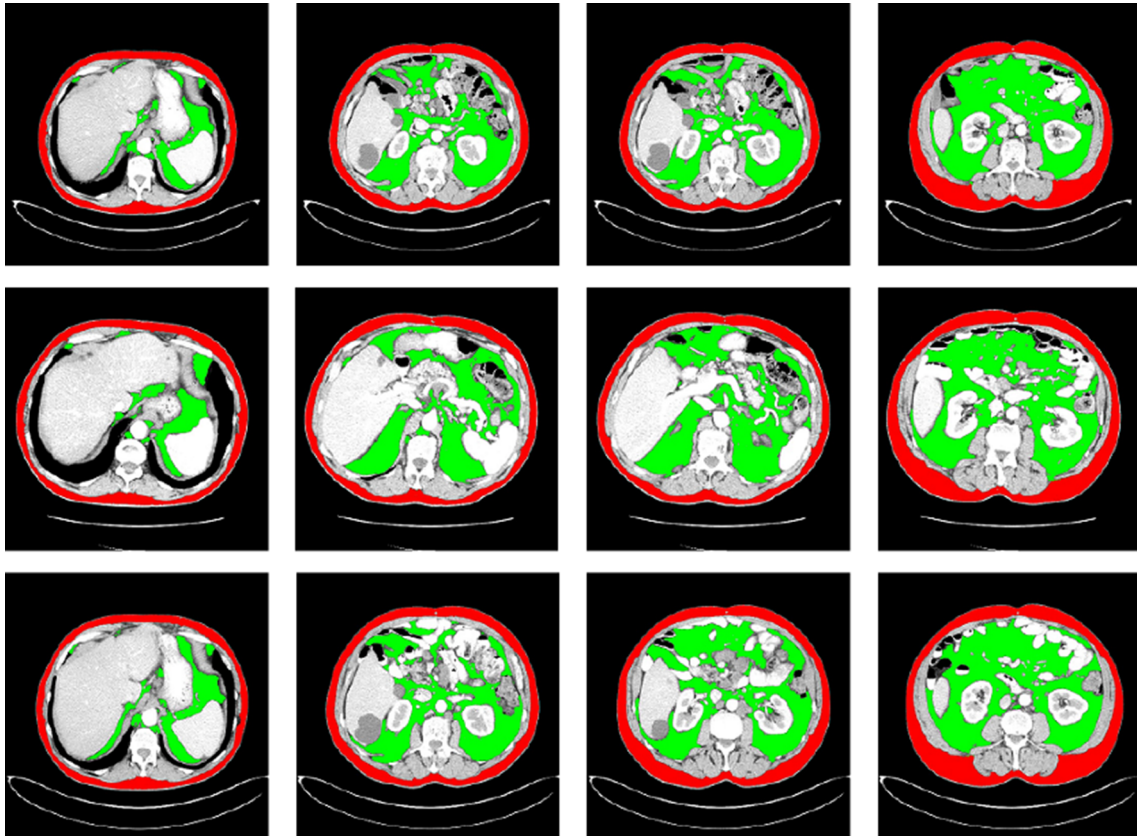
FIGURE 10. The proposed framework results on NIH DeepLesion datasets. Red color depicts VAT; Green color depicts SAT.

logical images (mostly 512 × 512). To validate our method, 290 abdominal CT images are randomly selected to split into three subsets: training subset (total of 203 images), including a validation subset, and testing subset (total of 87 images). All images are stored in unsigned 16 bit, and manually label by radiologists from the Second XiangYa Hospital of Central South University.

All of the abdominal CT images are standardized to 512 × 512. Then, apply our proposed framework without modifications, including the computer conditions and parameters of experiments. Total of 12 images from testing subset are randomly selected to display the performance of our proposed framework. Figure 10 shows the results with red color in VAT and green color in SAT.

ACKNOWLEDGMENTS

BMES BIOMEDICAL ENGINEERING SOCIETY

CONFLICT OF INTEREST

All authors declared that they have no conflict of interests.

REFERENCES

[1]Agarwal, C., A. H. Dallal, M. R. Arbabshirani, A. Patel, and G. Moore, Unsupervised quantification of abdominal fat from CT images using greedy snakes. In: Society of Photo-optical Instrumentation Engineers, 2017, p. 101332T.

[2]Aghaei, F., M. Tan, A. B. Hollingsworth, and B. Zheng. Applying a new quantitative global breast MRI feature analysis scheme to assess tumor response to chemotherapy. *J. Magn. Resonance Imaging* 44(5):1099–1106, 2016.

[3]Athanassiadi, K., A. Makrygianni, E. Balis, N. Alevizopoulos, M. Vaslamatzis, and C. Vourlakou. False-positive and false-negative rate after positron emission tomography/computer tomography scan for mediastinal staging in non-small-cell lung cancer. *Eur. Respir. J.* 42(1):93–100, 2014.

[4]Badrinarayanan, V., A. Kendall, and R. Cipolla. SegNet: a deep convolutional encoder-decoder architecture for scene segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39(99):2481–2495, 2015.

[5]Balasubramanian, T., S. Krishnan, M. Mohanakrishnan, K. R. Rao, C. V. Kumar, and K. Nirmala. Hog feature

based SVM classification of glaucomatous fundus image with extraction of blood vessels. In: India Conference, 2017, pp. 1–4.

[6]Boris, G., P. Jean Michel, B. Franck, L. Sylvain, G. Sverine, C. Jean-Pierre, K. Denis, H. Patrick, B. Christophe, and C. Bruno. Visceral fat area is an independent predictive biomarker of outcome after first-line bevacizumab-based treatment in metastatic colorectal cancer. *Gut* 59(3):341–347, 2010.

[7]Brebisson, A. D. and G. Montana. Deep neural networks for anatomical brain segmentation. In: Computer Vision & Pattern Recognition Workshops, vol. 2015-October, 2015, pp. 20–28.

[8]Caprio, S. Relationship between abdominal visceral fat and metabolic risk factors in obese adolescents. *Am. J. Hum. Biol.* 11(2):259–266, 1999.

[9]Chandra, M. A. and S. S. Bedi. Survey on svm and their application in image classification. *Int. J. Inform. Technol.* 2:1–11, 2018.

[10]Chen, L. C., G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS. *IEEE Trans. Pattern Anal. Mach. Intell.* 40(4):834–848, 2018.

[11]Chen, L. C., G. Papandreou, F. Schroff, and H. Adam. Rethinking atrous convolution for semantic image segmentation. arXiv:1706.05587, 2017.

[12]Chen, L. C., Y. Zhu, G. Papandreou, F. Schroff, and H. Adam. Encoder–decoder with atrous separable convolution for semantic image segmentation. In: European Conference on Computer Vision, pp. 833–851, 2018.

[13]Commandeur, F., M. Goeller, J. Betancur, S. Cadet, M. Doris, C. Xi, 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. Imaging* 37(8):1835–1846, 2018.

[14]Dalal, N. and B. Triggs. Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision & Pattern Recognition, vol. 2, 2005.

[15]Després, J. P., I. Lemieux, J. Bergeron, P. Pibarot, P. Mathieu, E. Larose, J. RodésCabau, O. F. Bertrand, and P. Poirier. Abdominal obesity and the metabolic syndrome: contribution to global cardiometabolic risk. *Arterioscler. Thromb. Vasc. Biol.* 28(6):1039, 2008.

[16]Drozdzal, M., E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal. The importance of skip connections in biomedical image segmentation. arXiv:1608.04117, 2016.

[17]Emaminejad, N., W. Qian, Y. Guan, M. Tan, Y. Qiu, H. Liu, and B. Zheng. Fusion of quantitative image and genomic biomarkers to improve prognosis assessment of early stage lung cancer patients. *IEEE Trans. Biomed. Eng.* 63(5):1034–1043, 2016.

[18]Estrada, S., R. Lu, S. Conjeti, X. Orozco-Ruiz, J. Panos-Willuhn, M. M. B. Breteler, and M. Reuter. FatSegNet: a fully automated deep learning pipeline for adipose tissue segmentation on abdominal dixon MRI. *CoRR.* arXiv:abs/1904.02082, 2019.

[19]Fujioka, S., Y. Matsuzawa, K. Tokunaga, and S. Tarui. Contribution of intra-abdominal fat accumulation to the impairment of glucose and lipid metabolism in human obesity. *Metab. Clin. Exp.* 36(1):54–59, 1987.

[20]Furey, T. S., Cristianini, N., Duffy, N., D. W. Bednarski, M. Schummer, and D. Haussler. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 16(10):906–914, 2000.

[21]S. Hai, F. Liu, Y. Xie, F. Xing, S. Meyyappan, and Y. Lin. Region segmentation in histopathological breast cancer images using deep convolutional neural network. In: IEEE International Symposium on Biomedical Imaging, 2015, pp. 55–58.

[22]He, Y., M. Keuper, B. Schiele, and M. Fritz. Learning dilation factors for semantic segmentation of street scenes. In: LNCS, vol. 10496, pp. 41–51, 2017.

[23]Hill, J. E., M. Fernandez-Del-Valle, R. Hayden, and S. Mitra. An automated segmentation for direct assessment of adipose tissue distribution from thoracic and abdominal dixon-technique mr images. In: Society of Photo-optical Instrumentation Engineers, vol. 10133, 2017, p. 1013315.

[24]Hinton, G. and T. Tieleman. Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. *COURSERA* 4:26–30, 2012.

[25]https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split/.

[26]Huang, G., Z. Liu, V. D. M. Laurens, and K. Q. Weinberger. Densely connected convolutional networks. In: European Conference on Computer Vision, vol. 2017-January, pp. 2261–2269, 2016.

[27]Hui, S. C. N., T. Zhang, L. Shi, D. Wang, and W. C. W. Chu. Automated segmentation of abdominal subcutaneous adipose tissue and visceral adipose tissue in obese adolescent in MRI. *Mag. Reson. Imaging* 45:97–104, 2017.

[28]Ioffe, S. and C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167, 2015.

[29]Jegou, S., M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio. The one hundred layers tiramisu: fully convolutional dense nets for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1175–1183, 2016.

[30]Ke, Y., X. Wang, L. Le, R. M. Summers, Y. Ke, X. Wang, L. Le, R. M. Summers, Y. Ke, and X. Wang. Deeplesion: automated deep mining, categorization and detection of significant radiology image findings using large-scale clinical lesion annotations. arXiv:1710.01766, 2017.

[31]Kim, S. H., J. H. Lee, B. Ko, and J. Y. Nam. X-ray image classification using random forests with local binary patterns. In: International Conference on Machine Learning & Cybernetics, vol. 6(July), 2010, pp. 3190–3194.

[32]Kissebah, A. H. and A. N. Peiris. Biology of regional body fat distribution: relationship to non-insulin-dependent diabetes mellitus. *Diabetes Metab. Rev.* 5(2):83–109, 2010.

[33]Kissebah, A. H., N. Vydelingum, R. Murray, D. J. Evans, A. J. Hartz, R. K. Kalkhoff, and P. W. Adams. Relation of body fat distribution to metabolic complications of obesity. *J. Clin. Endocrinol. Metab.* 54(2):254–254, 1982.

[34]Klein, S., U. van-der Heide, I. Lips, M. van Vulpen, M. Staring, and J. Pluim. Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information. *Med. Phys.* 35(4):1407–1417, 2008.

[35]Kullberg, J., A. Hedström, J. Brandberg, R. Strand, L. Johansson, G. Bergström, and H. Ahlström. Automated analysis of liver fat, muscle and adipose tissue distribution from CT suitable for large-scale studies. *Sci. Rep.* 7(1):10425, 2017.

[36]Kumar, K. V. V. and P. V. V. Kishore. Indian classical dance mudra classification using hog features and svm classifier. *Int. J. Electr. Comput. Eng.* 7(5):2537, 2018.

[37]Kvist, H., L. Sjoestroem, B. Chowdhury, M. Alpsten, B. Arvidsson, L. Larsson, and A. Cederblad. Determination of total adipose tissue and body fat in women by computed tomography, 40k, and tritium. *Am. J. Physiol.* 250(6 Pt 1):E736, 1986.

[38]Langer, T., A. Hedstrom, K. Morwald, D. Weghuber, A. Forslund, P. Bergsten, H. Ahlstrom, and J. Kullberg. Fully convolutional networks for automated segmentation of abdominal adipose tissue depots in multicenter water-fat MRI. *Magn. Reson. Med.* 81(4)2736–2745, 2019.

[39]Larsson, B., K. Svärdsudd, L. Welin, L. Wilhelmsen, P. Björntorp, and G. Tibblin. Abdominal adipose tissue distribution, obesity, and risk of cardiovascular disease and death: 13 year follow up of participants in the study of men born in 1913. *Br. Med. J.* 288(6428)1401–1404, 1984.

[40]Li, Z. and Y. Yu. Protein secondary structure prediction using cascaded convolutional and recurrent neural networks. In: Conference: International Joint Conference on Artificial Intelligence (IJCAI), New York, 2016.

[41]Liu, J., F. Chen, C. Pan, M. Zhu, X. Zhang, L. Zhang, and H. Liao. A cascaded deep convolutional neural network for joint segmentation and genotype prediction of brainstem gliomas. IEEE Trans. Bio-Med. Eng. 99:1, 2018.

[42]Makrogiannis, S., G. Caturegli, C. Davatzikos, and L. Ferrucci. Computer-aided assessment of regional abdominal fat with food residue removal in CT. *Acad. Radiol.* 20(11)1413–1421, 2013.

[43]Martinezuseros, J. and J. Garciafoncillas. Obesity and colorectal cancer: molecular features of adipose tissue. *J. Transl. Med.* 14(1):1–12, 2016.

[44]Mensink, S. D., J. W. Spliethoff, R. Belder, J. M. Klaase, R. Bezooijen, and C. H. Slump. Development of automated quantification of visceral and subcutaneous adipose tissue volumes from abdominal CT scans. In: Medical Imaging 2011: Computer-Aided Diagnosis, vol. 7963, no. 0, p. 79632Q, 2011.

[45]Moses, L. E., D. Shapiro, and B. Littenberg. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat. Med.* 12(14):1293–1316, 1993.

[46]Mun, E. C., G. L. Blackburn, and J. B. Matthews. Current status of medical and surgical therapy for obesity. *Gastroenterology* 120(3):669–681, 2001.

[47]Nakamura, T., K. Tokunaga, I. Shimomura, M. Nishida, S. Yoshida, K. Kotani, A. H. M. W. Islam, Y. Keno, T. Kobatake, and Y. Nagai. Contribution of visceral fat accumulation to the development of coronary artery disease in non-obese men. *Atherosclerosis* 107(2):239–246, 1994.

[48]Neumann, D., T. Langner, F. Ulbrich, D. Spitta, and D. Goehring. Online vehicle detection using haar-like, LBP and HOG feature based image classifiers with stereo vision preselection. In: Proceedings of the on Intelligent Vehicles Symposium, 2017.

[49]Ochs, R., J. Goldin, F. Abtin, H. Kim, K. Brown, P. Batra, D. Roback, M. Mcnitt-Gray, and M. Brown. Automated classification of lung bronchovascular anatomy in CT using adaboost. *Med. Image Anal.* 11(3):315–324, 2007.

[50]Ogden, C. L., M. D. Carroll, B. K. Kit, and K. M. Flegal. Prevalence of childhood and adult obesity in the united states, 2011–2012. *JAMA* 311(8):806, 2014.

[51]Oquab, M., L. Bottou, I. Laptev, and J. Sivic. Learning and transferring mid-level image representations using convolutional neural networks. In: IEEE Conference on Computer Vision & Pattern Recognition, 2014, pp. 1717–1724.

[52]Palacharla, P. K. Machine learning driven model inversion methodology to detect reniform nematodes in cotton. Dissertations & Theses - Gradworks, 2011.

[53]Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, and G. Louppe. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12(10):2825–2830, 2013.

[54]Peiris, A. N., M. S. Sothmann, R. G. Hoffmann, M. I. Hennes, C. R. Wilson, A. B. Gustafson, and A. H. Kissebah. Adiposity, fat distribution, and cardiovascular risk. *Ann. Int. Med.* 110(11):867–872, 1989.

[55]Pomponiu, V., H. Hariharan, B. Zheng, and D. Gur. Improving breast mass detection using histogram of oriented gradients. In: Medical Imaging: Computer-Aided Diagnosis, vol. 9035, 2014, p. 90351R.

[56]Qiu, Y., M. Tan, S. Mcmeekin, T. Thai, K. Ding, K. Moore, H. Liu, and B. Zheng. Early prediction of clinical benefit of treating ovarian cancer using quantitative ct image feature analysis. *Acta Radiol.* 57(9):1149, 2016.

[57]Rajendran, P. and M. Madheswaran. Hybrid medical image classification using association rule mining with decision tree algorithm. *Comput. Sci.* 3(10):1173–1178, 2010.

[58]Romero, D., J. C. Ramirez, and A. Marmol. Quantification of subcutaneous and visceral adipose tissue using CT. In: IEEE International Workshop on Medical Measurement & Applications, 2006, pp. 128–133.

[59]Ronneberger, O., P. Fischer, and T. Brox. U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing & Computer-assisted Intervention, vol. 9351, 2015, pp. 234–241.

[60]Saha, S., A. Mahmud, A. A. Ali, and M. A. Amin. Classifying digital X-ray images into different human body parts. In: International Conference on Informatics, 2016, pp. 67–71.

[61]Shelhamer, E., J. Long, and T. Darrell. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39(4):640–651, 2014.

[62]Shen, N., X. Li, S. Zheng, L. Zhang, Y. Fu, X. Liu, M. Li, J. Li, S. Guo, and H. Zhang. Automated and accurate quantification of subcutaneous and visceral adipose tissue from magnetic resonance imaging based on machine learning. *Magn. Reson. Imaging.* https://doi.org/10.1016/j.mri.2019.04.007, 2019.

[63]Sinno Jialin, P. and Y. Qiang. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22(10):1345–1359, 2010.

[64]Slaughter, K. N., T. Thai, S. Penaroza, D. M. Benbrook, E. Thavathiru, K. Ding, T. Nelson, D. S. Mcmeekin, and K. N. Moore. Measurements of adiposity as clinical biomarkers for first-line bevacizumab-based chemotherapy in epithelial ovarian cancer. *Gynecol. Oncol.* 133(1):11–15, 2014.

[65]Spasojević, A., O. Stojanov, T. L. Turukalo, and O. Sveljo, Estimation of subcutaneous and visceral fat tissue volume on abdominal MR images, 2015, pp. 217–220.

[66]Suykens, J. A. K. and J. Vandewalle. Least squares support vector machine classifiers. *Neural Process. Lett.* 9(3):293–300, 1999.

[67]Tajbakhsh, N., S. R. Gurudu, and J. Liang. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Trans. Med. Imaging* 35(2):630–644, 2016.

[68]Tokunaga, K., Y. Matsuzawa, K. Ishikawa, and S. Tarui. A novel technique for the determination of body fat by computed tomography. *Int. J. Obes.* 7(5):437–445, 1983.

[69]Van, D. W. S., J. L. Schönberger, J. Nuneziglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and T. S. Contributors. Scikit-image: image processing in python. *PeerJ* 2(2):e453, 2014.

[70]Walid, Z., T. Brown, A. Murtada, and S. Ali. The application of deep learning to quantify SAT/VAT in human abdominal area. In: Advances in Science and Engineering Technology International Conferences (ASET), 2019, pp. 1–5.

[71]Wang, P., X. Hu, Y. Li, Q. Liu, and X. Zhu. Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Process. 122:1–13, 2016.

[72]Wang, Y., Y. Qiu, T. Thai, K. Moore, L. Hong, and B. Zheng. A two-step convolutional neural network based computer-aided detection scheme for automatically segmenting adipose tissue volume depicting on CT images. Comput. Methods Program. Biomed. 144:97–104, 2017.

[73]Wang, Y., T. Thai, K. Moore, K. Ding, S. Mcmeekin, H. Liu, and B. Zheng. Quantitative measurement of adiposity using CT images to predict the benefit of bevacizumab-based chemotherapy in epithelial ovarian cancer patients. *Oncol. Lett.* 12(1):680–686, 2016.

[74]Wu, X. An iterative convolutional neural network algorithm improves electron microscopy image segmentation. *Comput. Sci.*, pp. 1–9, 2015.

[75]Xie, J., L. Yang, S. C. Zhu, and N. W. Ying. A theory of generative convnet. In: International Conference on International Conference on Machine Learning, 2016.

[76]Yi, L. and Y. F. Zheng. One-against-all multi-class SVM classification using reliability measures. In: IEEE International Joint Conference on Neural Networks, vol. 2, 2013, pp. 849-854.

[77]Yoon, D. Y., J. H. Moon, H. K. Kim, C. S. Choi, S. K. Chang, E. J. Yun, and Y. L. Seo. Comparison of low-dose ct and mr for measurement of intra-abdominal adipose tissue 1: a phantom and human study. *Acad. Radiol.* 15(1):62–70, 2008.

[78]Yoshizumi, T., T. Nakamura, M. Yamane, A. H. Islam, M. Menju, K. Yamasaki, T. Arai, K. Kotani, T. Funahashi, and S. Yamashita. Abdominal fat: standardized technique for measurement at CT. *Radiology* 211(1):283–286, 1999.

[79]Yu, L., H. Chen, Q. Dou, J. Qin, and P. A. Heng. Integrating online and offline 3D deep learning for automated polyp detection in colonoscopy videos. *IEEE J. Biomed. Health Inform.* 21(1):65–75, 2017.

[80]Yu, F. and V. Koltun. Multi-scale context aggregation by dilated convolutions. In: Conference Paper at ICLR, pp. 1–9, 2016.

[81]Zhao, H., J. Shi, X. Qi, X. Wang, and J. Jia. Pyramid scene parsing network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2017-January, 2017, pp. 6230–6239.