Magnetic Resonance Imaging 64 (2019) 28–36

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original contribution

Automated and accurate quantification of subcutaneous and visceral adipose tissue from magnetic resonance imaging based on machine learning

Ning Shen[a,1], Xueyan Li[a,*,1], Shuang Zheng[b], Lei Zhang[b], Yu Fu[b], Xiaoming Liu[a], Mingyang Li[a], Jiasheng Li[a], Shuxu Guo[a], Huimao Zhang[b,**]

[a] State Key Laboratory on Integrated Optoelectronics, College of Electronic Science and Engineering, Jilin University, 130012 Changchun, China
[b] Department of Radiology, the First Hospital of Jilin University, 130021 Changchun, China

ABSTRACT

Accurate measuring of subcutaneous adipose tissue (SAT) and visceral adipose tissue (VAT) is vital for the research of many diseases. The localization and quantification of SAT and VAT by computed tomography (CT) expose patients to harmful ionizing radiation. Magnetic resonance imaging (MRI) is a safe and painless test. The aim of this paper is to explore a practical method for the segmentation of SAT and VAT based on the iterative decomposition of water and fat with echo asymmetry and least square estimation-iron quantification (IDEAL-IQ) technology and machine learning. The approach involves two main steps. First, a deep network is designed to segment the inner and outer boundaries of SAT in fat images and the peritoneal cavity contour in water images. Second, after mapping the peritoneal cavity contour onto the fat images, the assumption-free K-means + + with a Markov chain Monte Carlo (AFK-MC^2) clustering method is used to obtain the VAT content. An MRI data set from 75 subjects is utilized to construct and evaluate the new strategy. The Dice coefficients for the SAT and VAT content obtained from the proposed method and the manual measurements performed by experts are 0.96 and 0.97, respectively. The experimental results indicate that the proposed method and the manual measurements exhibit high reliability.

1. Introduction

Obesity is the excessive accumulation of adipose tissue (AT) in the body and is known to be a key factor in multiple chronic conditions. Body-fat distribution is more dangerous to health than the overall amount body fat, and abdominal obesity has been reported to have the greatest health risks among the different types of obesity [1]. Based on statistics from the Chinese Center for Disease Control and Prevention, the mean body mass index (BMI) in Chinese adults aged 18–69 years increased from 22.7 kg/m^2 in 2004 to 23.7 kg/m^2 in 2010 [2]. An epidemiological study reported that 10.6% of 1.38 billion Chinese people have diabetes, and diabetic patients make up 15.4% and 21.1% of overweight and obese individuals, respectively [3]. Abdominal fat is associated with cardiovascular risk factors [4], and the triglyceride content in plasma and the liver is closely related to abdominal fat content [5–7]. There is a strong correlation between the VAT content and insulin resistance [8,9]. Fat not only manifests itself as the “fat” that can be seen, but it also accumulates around the internal organs and liver cells.

Abdominal fat analysis includes the assessment of two components: SAT and VAT. There are many ways to assess obesity in clinical trials; BMI is one of the simplest methods. However, it is impossible to measure the values of SAT and VAT. Bioelectrical impedance analysis (BIA) [10,11] is a common method to estimate body composition, particularly body fat. However, only a few studies report the actual amount of VAT. The development of imaging technologies, such as CT and MRI, has promoted the ability to measure fat.

Based on a concrete range of attenuation values associated with fat, muscle and bone tissue, CT can scan the entire body or a specific region of the body and generate a cross-section image, which can clearly distinguish the three tissues. However, patients are exposed to additional radiation when these measurements are taken. Tong et al. [12] explored a fuzzy model for the segmentation of the SAT and VAT components of chest fat. Kim YJ et al. [13] detected SAT and VAT levels using a separation mask based on muscles of the human body. Wang et al. [14] proposed a CNN framework consisting of two steps with a selection CNN and a segmentation-CNN to classify SAT and VAT. Weston et al. [15] used deep learning to perform abdominal VAT and SAT segmentation.

At present, body fat assessment can also be performed via the multiecho Dixon technique. The Dixon technique uses a spectral model to derive chemical shift-separated water and fat images from multiple

* Correspondence to X. Li: College of Electronic Science and Engineering, Jilin University, No. 2699 of Qianjin Street, Changchun, Jilin 130012, China.
** Correspondence to H. Zhang: Department of Radiology, the First Hospital of Jilin University, No. 71 of Xinmin Street, Changchun, Jilin 130021, China.
E-mail addresses: shenning17@mails.jlu.edu.cn (N. Shen), leexy@jlu.edu.cn (X. Li), liuxm16@mails.jlu.edu.cn (X. Liu), jali17@mails.jlu.edu.cn (J. Li), sxguo@jlu.edu.cn (S. Guo), huimaozhanglinda@163.com (H. Zhang).

[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.mri.2019.04.007
Received 18 January 2019; Received in revised form 2 April 2019; Accepted 17 April 2019
0730-725X/

source images acquired at different echo times [16]. The three-point Dixon (3PD) method is capable of proper fat and water signal assignment [17]. R2* mapping obtained according to T2* can analyze iron deposition in the liver and other parts [18]. For the diagnosis and quantification of fat, MRI is more sensitive than CT. The state-of-the-art technology of the 3PD method is IDEAL-IQ. The algorithm is designed to accommodate multiple spectral peaks of fat and leads to more accurate triglyceride fat fraction modeling and estimation [19,20]. By removing the effect from multiple chemical components, IDEAL-IQ improves the accuracy of AT quantification [21]. In a single scan, IDEAL-IQ produces six sequences: water images, fat images, fat fraction images, R2* relaxation images, in-phase images and out-phase images [22,23]. This technology has been indicated to be a very convenient and powerful method to measure fat fraction, and it is helpful for accurately assessing changes in fat content during disease progression. Wang et al. [24] separated the SAT and VAT regions in fat-water separation MRI based on a clustering method. Hui et al. [25] proposed an approach for the segmentation of SAT, VAT and bone marrow adipose tissue. Sun et al. [26] integrated some medical image processing methods to quantify AT. Sadananthan et al. [27] developed a segmentation method to classify deep and superficial subcutaneous tissue.

In our study, we present a new method for SAT and VAT segmentation in IDEAL-IQ sequences. A deep neural network model was trained to extract the SAT pattern. The AFK-MC^2 algorithm [28,29] was employed for accelerating clustering to obtain the VAT pattern. We address an independent, precise model for SAT and VAT segmentation with learning. By comparing a large amount of test data with manually measured results, the technical and clinical utility of the proposed method is verified and evaluated.

We describe the materials and methods in Section 2, the experiments and the results in Section 3, and we provide a discussion in Section 4. Finally, we conclude the paper in Section 5.

2. Material and methods

2.1. Subjects

In this study, all 75 subjects (male: 27, female: 48, age range: 29–94 years, mean age = 53.44 ± 11.79 years) were from The First Hospital of Jilin University. Abdominal MRIs from the subjects were utilized for modeling and testing.

Our research was retrospective, and the data are from March 2018 to September 2018. The hospital's ethics committee approved all studies included in this analysis. To limit possible postprandial effects, all the subjects fasted for 4 h before the MRI examinations. The duration for one test was approximately 5 min. The authors had no control over the data for the duration of the study. The patients' personal information was removed to protect their privacy.

2.2. MRI protocol

Patients were examined on a three Tesla (3 T) MRI clinical scanner (GE Medical Systems, Discovery MR750) using a quadrature body coil as the transmitter and receiver to acquire all images. Before scanning, patients were trained to hold their breath for > 20 s at the end of expiration. With the patients lying supine, the shoulders fitted closely against the coil, and the scan center was located at the xiphoid process. The scan included the region from the diaphragm to the belly button, and the patients used a respiratory door control hose.

IDEAL-IQ generated water images, fat images, fat fraction images, R2* relaxation images, in-phase images and out-phase images in one scan (see Fig. 1). The sequence parameters were TR/TE = 3.8/1.724 ms, flip angle = 12°, NEX = 1, FOV = 90 mm, matrix = 512 × 512, slice thickness = 5.6 mm, slice gap = 0, slice number = 76, ETL = 1, band width = 781.25, and a single-breath-hold scan duration = 20 s. The imaging and spectroscopic pulse sequences were performed in a single-breath-hold acquisition.

In this study, SAT and VAT were quantified based on water image sequences and fat image sequences in IDEAL-IQ. The significance of the water images participating in the experiment is that they have clearer peritoneal cavity contours than the fat images. We mapped the peritoneal cavity contour obtained by the automatic segmentation algorithm onto the fat images for the next step of VAT clustering.

2.3. Data augmentation

It is commonly known that the more data a deep learning model has been fed, the more effective it can be. Training deep neural networks for processing medical images is challenging due to the relative lack of large labeled datasets [30,31]. Data augmentation is a way to reduce overfitting. The main methods directly seek to augment the input data to the model, such as cropping, mirroring, color casting, scaling and rotation. In fact, not all data augmentation methods can be casually used. If robustness cannot be guaranteed, these methods will bring uncontrollable and harmful effects that will interfere with the convergence of the model.

We only increased the amount of training data, while simultaneously expanding the image data and the label data; otherwise, the new image would lack corresponding label data. The experimental results show that polar transformation and affine transformation improve the Dice results, but we cannot guarantee that the methods are useful for other data.

2.4. Experimental environment and data labeling

The experiment was performed on a PC with a Windows 10 64-bit operating system with an Intel i7-6800K CPU, 16 GB RAM, and an NVIDIA GeForce GTX 1080 video card. The deep learning network was developed on Keras with a TensorFlow backend. The program design languages are Python 3.6 (deep network design) and MATLAB® 2016b (MathWorks Inc., USA) (including preprocessing, data augmentation, and statistical analysis).

Three experts using in-house developed software took the manual measurements in this paper. The manual results were used as the ground truth data (see Fig. 2) for the comparison with the automated method results. Three radiologists marked the ground truth regions. The average age of the three radiologists is 37.7 years (range, 28–49 years), and the average of number of years of experience is 12.3 years (range, 2–24 years). Radiologist 1 (HMZ, chief physician, 24 years of experience) delineated the VAT. Radiologist 2 (LZ, 11 years of experience) outlined the peritoneal cavity contour. Radiologist 3 (SZ, 2 years of experience) labeled the inner and outer boundary of the SAT. Each doctor was responsible for delineating a type of ground truth, but to ensure the accuracy of the label, the other two radiologists independently checked these labels. Ground truth regions were generated for all 75 subjects.

2.5. Methodology for SAT/VAT segmentation

Fig. 3 shows the flow diagram for automating the quantification of SAT and VAT. The deep network completes the SAT quantification. The VAT quantification system consists of two main parts: the segmentation deep network and AFK-MC^2 clustering [32,33].

2.5.1. U-Net architecture

Deep learning is an important branch of machine learning. Deep neural networks extract patterns from data to solve complex problems, and deep learning shows breakthrough performance in image-related tasks. Radiology obtains information from images, so it is a very natural field for applying deep learning. At present, deep learning has shown promising performances in computer-aided diagnosis (CAD), including medical image detection [34–37], segmentation [38–42], classification

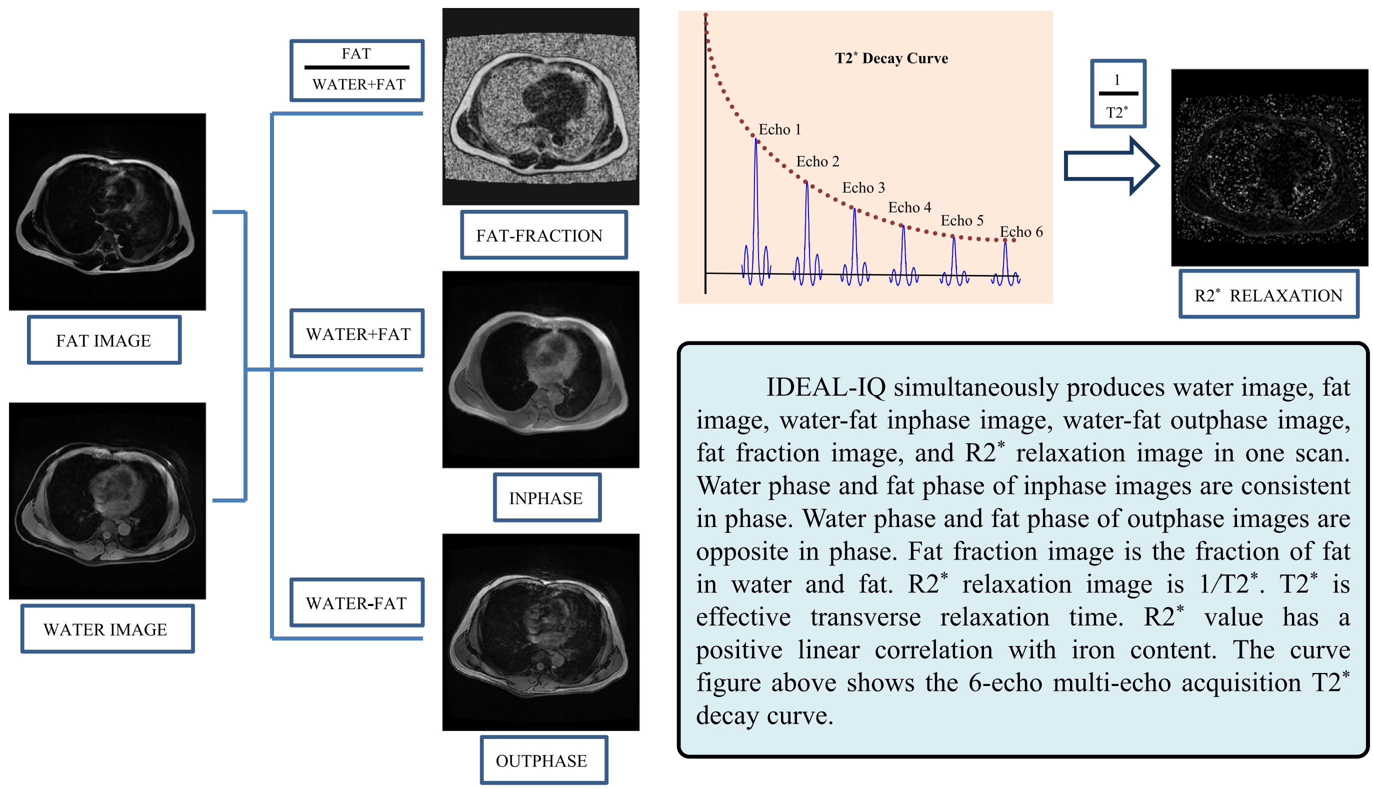

Fig. 1. Six sequences generated by IDEAL-IQ technique.

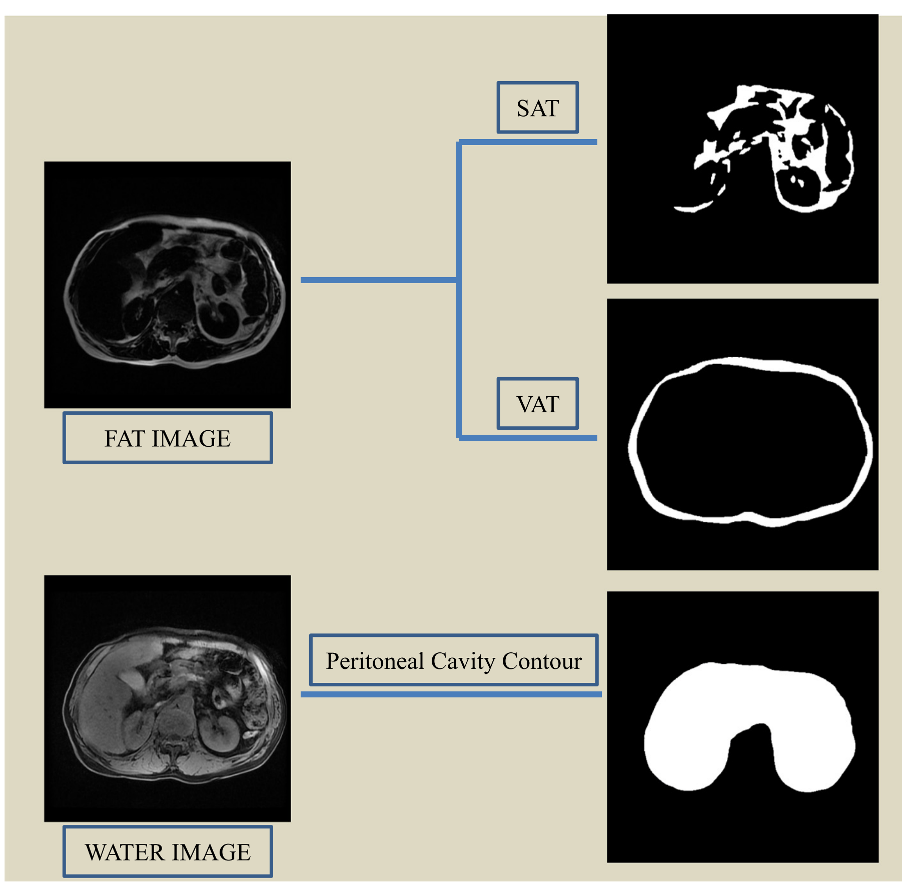

Fig. 2. The ground truth of SAT and VAT are outlined based on the fat images. The ground truth of peritoneal cavity contour are drawn based on the water images.

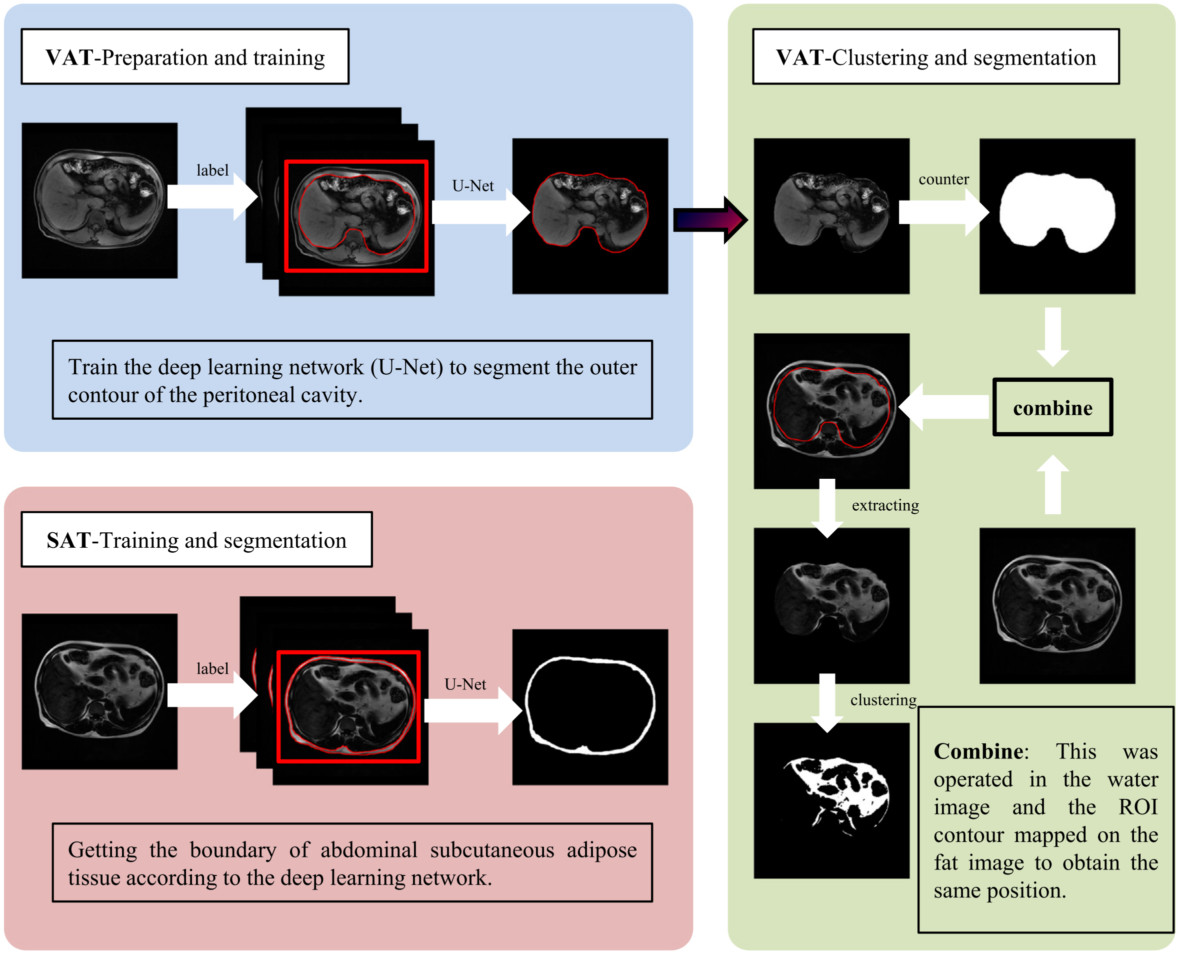

Fig. 3. Flowchart of the segmentation of SAT/VAT. The background of pink is the segmentation process of SAT, and the background of blue and green is the segmentation process of VAT. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

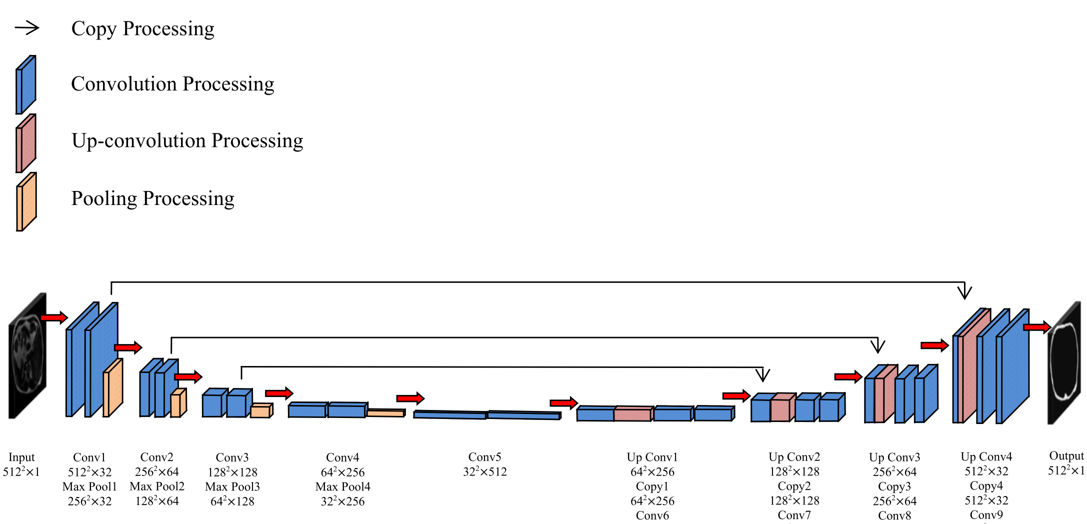

Fig. 4. U-Net architecture in this paper. All convolution processes use a filter size of 3 × 3. All max pooling layers and up-convolution filter size is 2 × 2. The padding for the convolution operation is the same.

[43–47] and others [48]. Segmentation is the most common task in CAD.

Fully convolutional networks (FCNs) are architectures of deep neural networks that promote image semantic segmentation from the image-level to the pixel-level. FCNs are composed of convolution, pooling and upsampling layers, but they do not have a fully connected layer. FCNs can work on images of all sizes. The convolution and pooling layers are used to capture what is in the image, while the up-sampling layer is used to localize where an object is. Skip connection is often used to recover the fine-grained spatial information which lost in the pooling layers. Compared with ordinary images, medical images have a large grayscale and unclear boundaries. The most successful end-to-end segmentation FCN is the U-Net framework [49]. U-Net evolved from FCNs and became a general network structure in biomedical image segmentation. The underlying information combines with high-level information to improve the segmentation effect. Copy processing (see Fig. 4) is often used to fully recover the fine-grained spatial information. Since it is one kind of FCN, the input is flexible in terms of size in both 2D and 3D data.

In this work, a deep network completed two tasks: SAT segmentation in fat images and peritoneal cavity contour segmentation in water images. Fig. 4 demonstrates the training procedure of our method. In the case of a batch size of 10, we used the Adam optimizer with a learning rate of 0.0001 to train the input images and their corresponding segmentation feature maps. We used softmax over the final feature map combined with the Dice loss function to update the weight parameters.

Soft-max is defined as following:

$$p_k(X) = \exp(a_k(X)) / \left(\sum_{k'=1}^{K} \exp(a_{k'}(X)) \right) \tag{1}$$

where $a_k(X)$ represents the activation in feature channel k at the pixel position $X \in \Omega$ with $\Omega \in Z^2$. $p_k(X)$ is the approximated maximum-function. The Dice loss function is:

$$J(\theta) = 1 - \frac{2\sum_{X\in\Omega} p_\theta(X)p(X)}{\sum_{X\in\Omega} p(X)^2 + \sum_{X\in\Omega} p_\theta(X)^2} \tag{2}$$

where p(X) is the label for each pixel; θ is the feature map. The weight parameters are then computed as:

$$\begin{aligned}
g_t &= \nabla_\theta J_t(\theta_{t-1}) \\
m_t &= \beta_1 \bullet m_{t-1} + (1-\beta_1)\bullet g_t \\
v_t &= \beta_2 \bullet v_{t-1} + (1-\beta_2)\bullet g_t^2 \\
\widehat{m}_t &= m_t/(1-\beta_1^t) \\
\hat{v}_t &= v_t/(1-\beta_2^t) \\
\theta_t &= \theta_{t-1} - \alpha \bullet \widehat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)
\end{aligned} \tag{3}$$

where t is the time step. m_t and v_t are the 1st and the 2nd moment vector, and are initialized with a zero vector. The exponential decay rate of 1st and 2nd moment estimation are β1 = 0.9 and β2 = 0.999, respectively. ε = 1 × 10–8 used to prevent the denominator from being zero.

We utilized 65 subjects' IDEAL-IQ sequences (fat images and water images) to construct U-Net architecture for SAT segmentation, VAT segmentation and the peritoneal cavity contour segmentation. Removing the spine and ribs in advance will improve the accuracy of intra-abdominal fat segmentation. A deep learning method was used to accurately locate the peritoneal cavity contour. All slices (*N* = 760) from the remaining 10 subjects were used to test the SAT and VAT segmentation network. The mean Dice coefficients of the SAT and VAT segmentation network were both 0.96.

2.5.2. Clustering based on AFK-MC^2

In this paper, AFK-MC^2-based clustering (preset K = 2, AT and non-AT) automatically performed VAT segmentation inner the peritoneal cavity contour in fat images. The pseudo code of AFK-MC^2 is as follows (see Table 1).

First, AFK-MC^2 was adopted to produce demonstrably good seeding and accelerated the sampling. The key to this approach is that it uses a Markov chain to treat the data points as state points. The first state is a data point that is randomly sampled through a random process to determine whether the state of the chain is transferred to other random data points and whether the state transition is independent of the initial distance of all points (i.e., the stable state of the Markov chain is independent of the initial state). The initial distance is calculated only once as part of the pre-processing phase. Next, the mini-batch K-means algorithm reduced the computation time. This algorithm is an iterative clustering algorithm that uses distance as a similarity index to find K classes in a given data set. AFK-MC^2-based clustering can classify the fat in the peritoneal cavity into AT and non-AT. The mean Dice coefficient of the VAT segmentation based on AFK-MC^2 was 0.97.

3. Results

Using the above model, we can obtain the SAT and VAT measurements by directly entering the fat images and water images, which is necessary to evaluate the accuracy and reproducibility of the results to guarantee the availability of the method.

3.1. MRI evaluation

In our method, the SAT and VAT pixel points obtained by the automated method were defined as foreground objects and were filled in white. The background (nonfat) was filled in black (see Fig. 5). The area of AT on each slice of the fat images was calculated as the number of foreground pixels multiplied by the size of each pixel. The volume of AT was calculated by summing the area of fat in each slice and multiplying by the thickness of the slice. Finally, the total SAT and VAT volumes were calculated by summing the respective volume accumulations.

3.2. Evaluation results

To validate the method, the areas of SAT and VAT were measured using the automated method, and the results were compared with the ground truths. The method was evaluated by the percentage mean error (PMR), the percentage mean false-positive errors (FPE), the percentage mean false-negative errors (FNE), and the Pearson correlation coefficient (PCC), as well as intraclass correlation coefficient (ICC). The statistical results are presented in Tables 2 and 3.

There is an excellent correlation between the proposed method and the ground truths (the PCC of both the SAT and VAT comparisons are > 0.990). Moreover, the ICC between the proposed method and the results of the manual measurements indicate a high reliability ($P < 0.001$). The SAT and VAT content obtained by the automated method are both larger than the ground truths. The FPE for both SAT and VAT is higher than the respective FNE. The FPE and the FNE for SAT are higher than the values for VAT. The mean FPE and the mean FNE for SAT and VAT are < 5%. The mean differences of SAT and VAT between the proposed method and ground truths are < 2%.

The Bland-Altman plot displays the distribution of the differences between two measurements. On the vertical axis, the mean of the difference is the mean reference line. The 95% agreement limits can be calculated from mean difference (the estimated bias) and the standard deviation of the differences (mean ± 1.96SD). Fig. 6a indicates that the mean difference between the proposed method and the ground truth is positive, which means that our method slightly overestimates SAT relative to the ground truth. Fig. 6b shows that the mean difference is also positive, which means our method also slightly overestimates the VAT area relative to the ground truths. The mean difference of the SAT is larger than the VAT, while the standard deviation of the SAT is smaller than that of the VAT. According to the plots, there is no significant correlation between the quantitative results and the quantitative errors.

To compare the data distributions of the test results and the ground truth, we used a box-plot (see Fig. 6c). The maximum value, the minimum value, the upper quartile and the lower quartile of the SAT ground truth are greater than the test. However, the median value of the SAT for the ground truth is less than the test result. Regarding the maximum value, the median value, the upper quartile and the lower

Table 1
The pseudo code of AFK-MC^2.

Algorithm: AFK-MC^2 sampling and mini-batch K-means

Require: Mini-batch size b, iterations t, data set $\mathscr{X}$, centers k, chain length m
// AFK-MC^2 sampling
// Preprocessing step
c_1 sampled from $\mathscr{X}$ at random, $\mathscr{X} = \{x_1, x_2, \cdots, x_n\}$
for all $x \in \mathscr{X}$ do
$q(x \mid c_1) = d(x, c_1)^2 / 2 \sum_{x' \in \mathscr{X}} d(x', c_1)^2 + 1/2\,|\mathscr{X}|$
end for
// Main loop
$C_1 \leftarrow \{c_1\}$
for $i = 2, 3, \cdots, k$ do
x sampled from $\mathscr{X}$ using $q(x)$
$d_x \leftarrow d(x, C_{i-1})^2$
for $j = 2, 3, \cdots, m$ do
y sampled from $\mathscr{X}$ using $q(y)$
$d_y \leftarrow d(y, C_{i-1})^2$
if $\frac{d_y q(x)}{d_x q(y)} >$ Unif [0,1]
then $x \leftarrow y$, $d_x \leftarrow d_y$
else $x \leftarrow x$, $d_x \leftarrow d_x$
end if
end for
$C_i \leftarrow C_{i-1} \cup \{x\}$
end for
return C_k

// Mini-batch K-means
$v \leftarrow 0$
$d[c] \leftarrow \varnothing$
for $i = 2, 3, \cdots, t$ do
$M \leftarrow b$ examples picked from $\mathscr{X}$ at random
for $x \in M$ do
$c \leftarrow f(C_k, x)$
$d[c] \leftarrow d[c] \cup \{x\}$
end for
for $x \in M$ do
$c \leftarrow f(C_k, x)$
$C_k \leftarrow C_k \backslash \{c\}$
$v[c] \leftarrow v[c] + 1$
$\mu \leftarrow 1/v[c]$
$c \leftarrow (1 - \mu)c + \mu x$
$C_k \leftarrow C_k \cup \{c\}$
end for
end for
return $C = \{d[c_1], d[c_2], \cdots, d[c_k]\}$

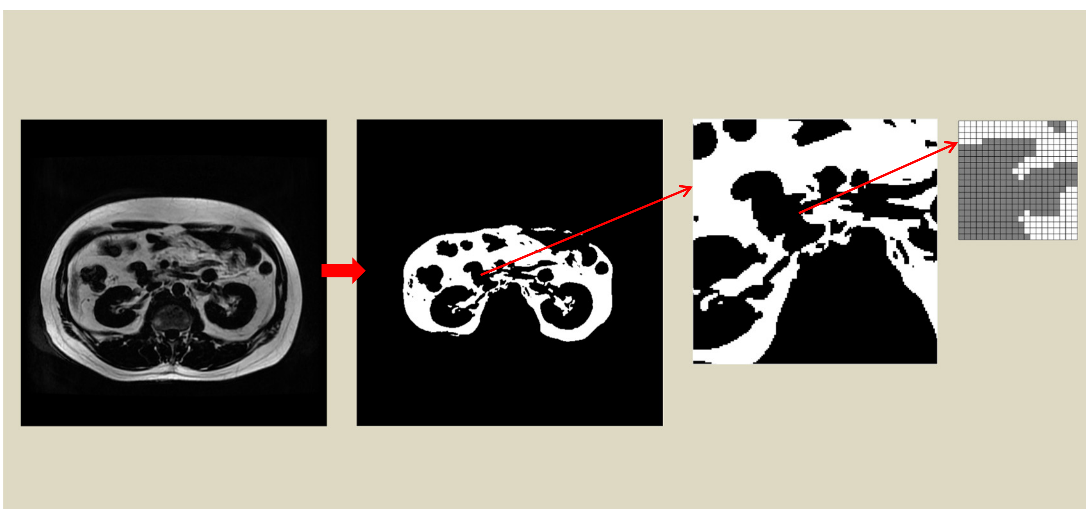

Fig. 5. VAT clustering results. From left to right, the original image, the clustering result, the enlarged image in order and zoom in to the pixel level for the second time. The area of fat is equal to the number of white pixels multiplied by the actual size of each pixel.

quartile of the VAT, the ground truth is greater than the test results. For the minimum value of the VAT, the ground truth is less than the test results. The plot illustrates that the test results and the ground truth distribution are highly similar. This proves that the experimental method is feasible for quantitative SAT and VAT measurements.

4. Discussion

Although various approaches have been proposed, an end-to-end segmentation model for SAT and VAT has yet to be developed that improves the efficiency and accuracy, especially for MRI. To achieve this goal, we attempt to develop and validate a new method inspired by

Table 2
SAT and VAT contents resulting from the proposed method and the ground truth ($N = 760$).

	Measurements area (cm^2)		Proposed method vs ground truth			
	Proposed (cm^2)	Ground truth (cm^2)	PMR (%)[a]	PCC	FPE (%)	FNE (%)
SAT	106.308 ± 38.632	106.613 ± 38.596	−0.346 ± 2.731	0.997	3.802 ± 1.957	4.147 ± 2.074
VAT[b]	90.381 ± 47.841	89.894 ± 48.201	0.715 ± 3.828	0.998	3.145 ± 2.619	2.430 ± 2.795

[a] PMR represents the difference between the proposed method and ground truth.
[b] VAT represents the AFK-MC^2 based VAT segmentation.

Table 3
Comparison and verification between the automated method and the ground truth (N = 760).

	Area (cm^2)	F	P value[a]	ICC	P value[b]
SAT_{auto}	106.308 ± 38.632	0.024	0.878	0.997	< 0.001
$SAT_{ground\ truth}$	106.613 ± 38.596				
VAT^{c}_{auto}	90.381 ± 47.841	0.039	0.843	0.996	< 0.001
$VAT^{c}_{ground\ truth}$	89.894 ± 48.201				

[a] *P* value for ANOVA test.
[b] *P* value for ICC.
[c] VAT represents the AFK-MC^2 based VAT segmentation.

deep learning and clustering. This study used IDEAL-IQ to automatically and accurately quantify abdominal SAT and VAT. The advantages are an increase in the degree of automation and objectivity, a more accurate segmentation result, and a model that requires only a small amount of label data.

In the fat sequence, the spine and ribs and nearby non-VAT areas show a similar signal intensity to VAT. If clustering is directly carried out in fat images, some nonfat tissue will be incorrectly classified, which will lead to an overestimation of the VAT content, and it is the reason why we operate VAT clustering within the peritoneal cavity contour. The experimental results show that we obtain the expected results.

The advantage of the proposed method is that it is unsupervised and does not require labels for VAT segmentation. VAT lacks clearly definable boundaries. Unlike with SAT labeling, it is more difficult to obtain the VAT ground truth regions. Water image sequences and fat image sequences are acquired simultaneously in IDEAL-IQ, and they have an accurate position correspondence. To eliminate the false detection of VAT, the peritoneal contour depicted by U-Net with water images could map onto the fat image to obtain the same position. Therefore, the VAT pattern can be obtained in fat images.

The K-means clustering algorithm can cluster some observations without labels, but there has to be a clear total number of categories, such as those with a total of K categories. The core of the algorithm involves the need to find cluster centers for each category. A random seed point is required to initiate the algorithm. However, it is difficult to obtain an optimal solution using a random seed algorithm. The AFK-MC^2 algorithm can obtain better clustering results without assuming a data distribution. It improves the generation of initial seeds, and its clustering speed is several orders of magnitude higher than the current best K-means+ + clustering method. Between the solution quality and computational cost, the sampling algorithm is an agreeable compromise.

Our approach combines supervised and unsupervised learning. The improvement gained by combining different methods of machine learning is encouraging. The use of machine learning methods for SAT and VAT quantification has broad prospects and far-reaching significance. However, our scheme still faces some problems: 1) although the quantification process is quick, the deep learning network requires a certain amount of training time; 2) multicenter data should be collected; and 3) the scans include the region in the upper abdomen from the diaphragm to the belly button but not to the coccyx.

5. Conclusions

The purpose of this paper is to develop and validate a new method to quantitative abdominal SAT and VAT using IDEAL-IQ images. There is an excellent agreement between our method and the ground truth. The method has the merits of high operation convenience and accuracy. The quantification results would contribute to developing the CAD system and improving the working efficiency of radiologists.

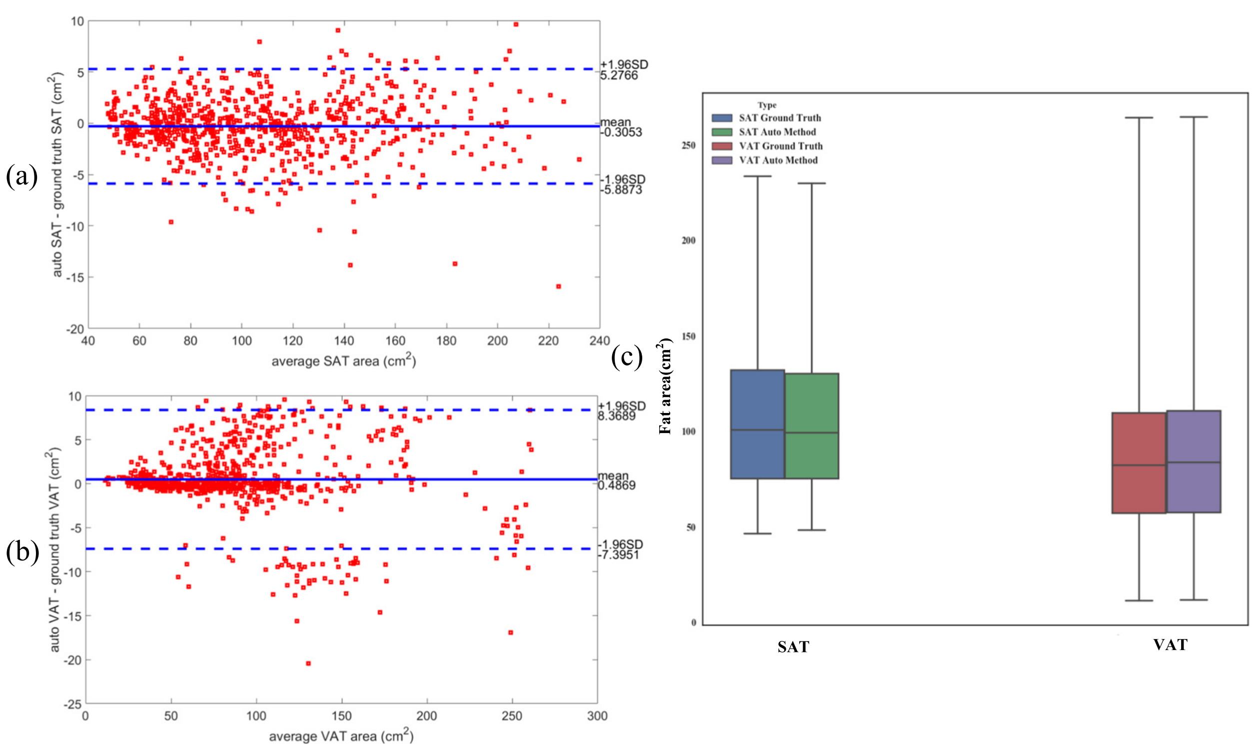

Fig. 6. Bland–Altman plots reflect the difference between the automatic method and the ground truth. The abscissa indicates the average of the area measured by the proposed method and ground truth in each slice. The ordinate represents the area value of each slice automatic method minus the ground truth. The solid line denotes the mean difference. The dashed line denotes the difference that is 1.96 standard deviations away from the mean difference. (a) and (b) represent SAT area and VAT area between the proposed method and the ground truth separately. The box-plot (c) reflects the distribution of test results and ground truths. SAT Auto Method and SAT Ground Truth express the distribution of test result and ground truth of SAT, respectively. VAT Auto Method and VAT Ground Truth express the distribution of test result and ground truth of VAT, respectively.

Acknowledgements

We want to thank the National Health Commission of the People's Republic of China for their support. Funding: This work was supported by the Natural Science Foundation of Jilin Province [grant numbers 20180101038JC]; the Science and Technology Development Plan of Jilin Province [grant numbers 20170622009JC]; and the Provincial and School Joint Construction Project of Jilin University [grant numbers SXGJXX2017-8].

References

[1] WHO. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. World health organ tech rep Ser. vol. 894. 2000. p. i–. 1–253 https://doi.org/10.1002/jps.3080150106.
[2] Jiang Y, Xu Y, Bi Y, Wang L, Zhang M, Zhou M, et al. Prevalence and trends in overweight and obesity among Chinese adults in 2004-10: data from three nationwide surveys in China. The Lancet 2015;386:S77. https://doi.org/10.1016/S0140-6736(15) 00658-3.
[3] Cheung LTF, Chan RSM, Ko GTC, Lau ESH, Chow FCC, Kong APS. Diet quality is inversely associated with obesity in Chinese adults with type 2 diabetes. Nutr J 2018;17(1):63. https://doi.org/10.1186/s12937-018-0374-6.
[4] Lee JJ, Pedley A, Hoffmann U, Massaro JM, Fox CS. Association of changes in abdominal fat quantity and quality with incident cardiovascular disease risk factors. J Am Coll Cardiol 2016;68(14):1509–21. https://doi.org/10.1016/j.jacc.2016.06.067.
[5] Mittendorfer B, Yoshino M, Patterson BW, Klein S. VLDL triglyceride kinetics in lean, overweight, and obese men and women. J Clin Endocrinol Metab 2016;101(11):4151–60. https://doi.org/10.1210/jc.2016-1500.
[6] Björnson E, Adiels M, Taskinen MR, Borén J. Kinetics of plasma triglycerides in abdominal obesity. Curr Opin Lipidol 2017;28(1):11–8. https://doi.org/10.1097/MOL. 0000000000000375.
[7] Bril F, Barb D, Portillo-Sanchez P, Biernacki D, Lomonaco R, Suman A, et al. Metabolic and histological implications of intrahepatic triglyceride content in nonalcoholic fatty liver disease. Hepatology 2017;65(4):1132–44. https://doi.org/10.1002/hep.28985.
[8] Bi X, Loo YT, Henry CJ. Ultrasound measurement of intra-abdominal fat thickness as a predictor of insulin resistance and low HDL cholesterol in Asians. Nutrition 2018;55-6:99–103. https://doi.org/10.1016/j.nut.2018.04.003.
[9] Genske F, Kühn JP, Pietzner M, Homuth G, Rathmann W, Grabe HJ, et al. Abdominal fat deposits determined by magnetic resonance imaging in relation to leptin and vaspin levels as well as insulin resistance in the general adult population. Int J Obes (Lond) 2018;42(2):183–9. https://doi.org/10.1038/ijo.2017.187.
[10] Jackson AS, Pollock ML, Graves JE, Mahar MT. Reliability and validity of bioelectrical impedance in determining body composition. J Appl Physiol 1988;64(2):529–34. https://doi.org/10.1152/jappl.1988.64.2.529.
[11] Ida M, Hirata M, Hosoda K, Nakao K. Abdomen specific bioelectrical impedance analysis (BIA) methods for evaluation of abdominal fat distribution. Nihon Rinsho 2013;71(2):262–5.
[12] Tong Y, Udupa JK, Wu C, Pednekar G, Subramanian JR, Lederer DJ, et al. Fat segmentation on chest CT images via fuzzy models. Medical imaging 2016: Image-guided procedures, robotic interventions, and modeling vol. 9786. 2016. p. 09. https://doi.org/10.1117/12.2217864.
[13] Kim YJ, Park JW, Kim JW, Park CS, Gonzalez JP, Lee SH, et al. Computerized automated quantification of subcutaneous and visceral adipose tissue from computed tomography scans: development and validation study. JMIR Med Inform 2016;4(1):e2https://doi.org/10.2196/medinform.4923.
[14] Wang Y, Qiu Y, Thai T, Moore K, Liu H, Zheng B. A two-step convolutional neural network based computer-aided detection scheme for automatically segmenting adipose tissue volume depicting on CT images. Comput Methods Programs Biomed 2017;144:97–104. https://doi.org/10.1016/j.cmpb.2017.03.017.
[15] Weston AD, Korfiatis P, Kline TL, Philbrick KA, Kostandy P, Sakinis T, et al. Automated abdominal segmentation of CT scans for body composition analysis using deep learning. Radiology 2019;290(3):669–79. https://doi.org/10.1148/radiol. 2018181432.
[16] Berglund J, Johansson L, Ahlström H, Kullberg J. Three-point Dixon method enables whole-body water and fat imaging of obese subjects. Magn Reson Med 2010;63(6):1659–68. https://doi.org/10.1002/mrm.22385.
[17] Szumowski J, Coshow WR, Li F, Quinn SF. Phase unwrapping in the three-point Dixon method for fat suppression MR imaging. Radiology 1994;192(2):555–61. https://doi.org/10.1148/radiology.192.2.8029431.
[18] Shinohara Y, Kato A, Yamashita E, Ogawa T. R2* map by IDEAL IQ for acute cerebral infarction: compared with susceptibility vessel sign on T2*-weighted imaging. Yonago Acta Med 2016;59(3):204–9.
[19] Idilman IS, Tuzun A, Savas B, Elhan AH, Celik A, Idilman R, et al. Quantification of liver, pancreas, kidney, and vertebral body MRI-PDFF in non-alcoholic fatty liver disease. Abdom Imaging 2015;40(6):1512–9. https://doi.org/10.1007/s00261-015-0385-0.
[20] Aoki T, Yamaguchi S, Kinoshita S, Hayashida Y, Korogi Y. Quantification of bone marrow fat content using iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): reproducibility, site variation and correlation with age and menopause. Br J Radiol 2016;89(1065):20150538. https://doi.org/10.1259/bjr.20150538.
[21] Ternifi R, Pouletaut P, Sasso M, Mietteb V, Charleuxc F, Bensamouna SF. Improvements of liver MR imaging clinical protocols to simultaneously quantify steatosis and iron overload. IRBM 2018;39(3):219–25. https://doi.org/10.1016/j.irbm. 2018.04.004.
[22] Joe E, Lee JM, Kim KW, Lee KB, Kim SJ, Baek JH, et al. Quantification of hepatic macrosteatosis in living, related liver donors using T1-independent, T2*-corrected chemical shift MRI. J Magn Reson Imaging 2012;36(5):1124–30. https://doi.org/10.1002/jmri.23738.
[23] Hayashi T, Saitoh S, Takahashi J, Tsuji Y, Ikeda K, Kobayashi M, et al. Hepatic fat quantification using the two-point Dixon method and fat color maps based on nonalcoholic fatty liver disease activity score. Hepatol Res 2017;47(5):455–64. https://doi.org/10.1111/hepr.12767.
[24] Wang D, Shi L, Chu WC, Hu M, Tomlinson B, Huang WH, et al. Fully automatic and nonparametric quantification of adipose tissue in fat-water separation MR imaging. Med Biol Eng Comput 2015;53(11):1247–54. https://doi.org/10.1007/s11517-015-1347-y.
[25] Hui SCN, Zhang T, Shi L, Wang D, Ip CB, Chu WCW. Automated segmentation of abdominal subcutaneous adipose tissue and visceral adipose tissue in obese adolescent in MRI. Magn Reson Imaging 2018;45:97–104. https://doi.org/10.1016/j.mri.2017.09.016.
[26] Sun J, Xu B, Freeland-Graves J. Automated quantification of abdominal adiposity by magnetic resonance imaging. Am J Hum Biol 2016;28(6):757–66. https://doi.org/10.1002/ajhb.22862.
[27] Sadananthan SA, Prakash B, Leow MK, Khoo CM, Chou H, Venkataraman K, et al. Automated segmentation of visceral and subcutaneous (deep and superficial) adipose tissues in normal and overweight men. J Magn Reson Imaging 2015;41(4):924–34. https://doi.org/10.1002/jmri.24655.
[28] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57(1):97–109. https://doi.org/10.2307/2334940.
[29] Bachem O, Lucic M, Hassani H, Krause A. Fast and provably good seedings for K-means. Adv Neural Inf Process Syst 2016;2016:55–63.
[30] Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning requires rethinking generalization. arXiv 2016. [preprint arXiv:1611.03530].
[31] Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning. arXiv 2017;1712:04621.
[32] Lloyd S. Least squares quantization in PCM. IEEE Trans Inf Theory 1982;28(2):129–37. https://doi.org/10.1109/TIT.1982.1056489.
[33] Sculley D. Web-scale K-means clustering. International conference on world wide web vol. 219. 2010. p. 1177–8. https://doi.org/10.1145/1772690.1772862.
[34] Bruijne Marleen de. Machine learning approaches in medical image analysis: from detection to diagnosis. Med Image Anal 2016;33:94–7. https://doi.org/10.1016/j.media.2016.06.032.
[35] Setio AA, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel SJ, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 2016;35(5):1160–9. https://doi.org/10.1109/TMI.2016.2536809.
[36] Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal 2017;35:303–12. https://doi.org/10.1016/j.media.2016.07.007.
[37] Charbonnier JP, Rikxoort EMV, Setio AAA, Schaefer-Prokop CM, Ginneken BV, Ciompi F. Improving airway segmentation in computed tomography using leak detection with convolutional networks. Med Image Anal 2017;36:52–60. https://doi.org/10.1016/j.media.2016.11.001.
[38] Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Med Image Anal 2017;35:159–71. https://doi.org/10.1016/j.media.2016.05. 009.
[39] Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Med Image Anal 2016;30:108–19. https://doi.org/10.1016/j.media.2016. 01.005.
[40] Roth HR, Lu L, Lay N, Harrison AP, Farag A, Sohn A, et al. Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation. Med Image Anal 2018;45:94–107. https://doi.org/10.1016/j.media.2018.01.006.
[41] Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders MJ, Isgum I. Automatic segmentation of MR brain images with a convolutional neural network. IEEE Trans Med Imaging 2016;35(5):1252–61. https://doi.org/10.1109/TMI.2016.2548501.
[42] Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal 2017;36:61–78. https://doi.org/10.1016/j.media.2016. 10.004.
[43] Vu TH, Mousavi HS, Monga V, Rao G, Rao UK. Histopathological image classification using discriminative feature-oriented dictionary learning. IEEE Trans Med Imaging 2016;35(3):738–51. https://doi.org/10.1109/TMI.2015.2493530.
[44] Sirinukunwattana K, Ahmed Raza SE, Tsang Yee-Wah, Snead DR, Cree IA, Rajpoot NM. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 2016;35(5):1196–206. https://doi.org/10.1109/TMI.2016.2525803.
[45] Xie F, Fan H, Li Y, Jiang Z, Meng R, Bovik A. Melanoma classification on dermoscopy images using a neural network ensemble model. IEEE Trans Med Imaging 2017;36(3):849–58. https://doi.org/10.1109/TMI.2016.2633551.
[46] Manescu P, Jong Lee Y, Camp C, Cicerone M, Brady M, Bajcsy P. Accurate and interpretable classification of microspectroscopy pixels using artificial neural networks. Med Image Anal 2017;37:37–45. https://doi.org/10.1016/j.media.2017.01.

001.
[47] Pang S, Du A, Orgun MA, Yu Z. A novel fused convolutional neural network for biomedical image classification. Med Biol Eng Comput 2019;57:107–21. https://doi.org/10.1007/s11517-018-1819-y.
[48] Kuestner T, Gatidis S, Liebgott A, Schwartz M, Mauch L, Martirosian P, et al. A machine-learning framework for automatic reference-free quality assessment in MRI. Magn Reson Imaging 2018;53:134–47. https://doi.org/10.1016/j.mri.2018.07.003.
[49] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Med Image Comput Comput Assist Interv 2015;9351:234–41. https://doi.org/10.1007/978-3-319-24574-4_28.