Computer Methods and Programs in Biomedicine 185 (2020) 105329

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images

Junying Chen[a,*], Haijun You[a], Kai Li[b]

[a] School of Software Engineering, South China University of Technology, Guangzhou, Guangdong 510006, China
[b] Department of Ultrasound, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong 510630, China

ARTICLE INFO

Article history:
Received 31 May 2019
Revised 8 January 2020
Accepted 8 January 2020

Keywords:
Thyroid ultrasound image
Gland segmentation method
Nodule segmentation method
Segmentation performance analysis

ABSTRACT

Background and objective Thyroid image segmentation is an indispensable part in computer-aided diagnosis systems and medical image diagnoses of thyroid diseases. There have been dozens of studies on thyroid gland segmentation and thyroid nodule segmentation in ultrasound images. The aim of this work is to categorize and review the thyroid gland segmentation and thyroid nodule segmentation methods in medical ultrasound.

Methods This work proposes a categorization approach of thyroid gland segmentation and thyroid nodule segmentation methods according to the theoretical bases of segmentation methods. The segmentation methods are categorized into four groups, including contour and shape based methods, region based methods, machine and deep learning methods and hybrid methods. The representative articles are reviewed with detailed descriptions of methods and analyses of correlations between methods. The evaluation metrics for the reviewed segmentation methods are named uniformly in this work. The segmentation performance results using the uniformly named evaluation metrics are compared.

Results After careful investigation, 28 representative papers are selected for comprehensive analyses and comparisons in this review. The dominant thyroid gland segmentation methods are machine and deep learning methods. The training of massive data makes these models have better segmentation performance and robustness. But deep learning models usually require plenty of marked training data and long training time. For thyroid nodule segmentation, the most common methods are contour and shape based methods, which have good segmentation performance. However, most of them are tested on small datasets.

Conclusions Based on the comprehensive consideration of application scenario, image features, method practicability and segmentation performance, the appropriate segmentation method for specific situation can be selected. Furthermore, several limitations of current thyroid ultrasound image segmentation methods are presented, which may be overcome in future studies, such as the segmentation of pathological or abnormal thyroid glands, identification of the specific nodular diseases, and the standard thyroid ultrasound image datasets.

1. Introduction

Thyroid is an important endocrine organ of the human body, which lies in the anterior part of the neck just below the thyroid cartilage [1], secreting thyroid hormone to regulate the body's metabolism [2]. Some thyroid nodules have regular and well-defined margins, but some others are in irregular shapes. Thyroid nodules are usually characterized as solid, cystic, or a combination of solid and cystic nodules [3]. According to epidemiologic studies, palpable thyroid nodules occur in 4%–7% of the population, but nodules found incidentally in ultrasound examinations show a proportion of 19%–67% [4]. Thyroid nodules are classified as hypoechoic, isoechoic, or hyperechoic. Previous studies have shown that hypoechoic nodules with irregular boundaries are more likely to develop into malignant nodules [5]. The incidence of malignant thyroid nodules is 0.1%-0.2% [6].

Different imaging modalities are used for thyroid disease diagnoses, such as computed tomography, magnetic resonance imaging, ultrasound imaging, and radionuclide imaging. Among these imaging methods, ultrasound imaging technology is the most

* Corresponding author.
E-mail addresses: jychense@scut.edu.cn (J. Chen), 201721045312@mail.scut.edu.cn (H. You), likai@mail.sysu.edu.cn (K. Li).

https://doi.org/10.1016/j.cmpb.2020.105329
0169-2607/© 2020 Published by Elsevier B.V.

commonly used method for diagnoses of thyroid diseases, which is real-time, inexpensive, non-invasive and non-radioactive [7]. But the relatively low quality and the speckle noise in ultrasound images make the organ tissues in ultrasound images fuzzy and inhomogeneous [8]. In clinical diagnoses, ultrasound examinations are dependent on the experienced clinicians who are required to interpret the ultrasound images. Moreover, clinicians need long learning curves to be sophisticated in ultrasound examinations and it is hard to remove subjective factors in the diagnostic processes [7]. The important sonographic features for thyroid nodule diagnoses are divided into five categories in thyroid imaging, reporting and data system (TI-RADS) which is released by American College of Radiology [9], including composition, echogenicity, shape, margin and echogenic foci. Shapes and boundaries of thyroid nodules are the key characteristics to distinguish between benign and malignant nodules [10], especially taller-than-wide shape and spiculated margin are two of the major features to determine malignant thyroid nodules [11]. The specificity and the sensitivity of using taller-than-wide shape as the feature to diagnose malignant nodules are 91.4% and 40.0%, and the specificity and the sensitivity of using spiculated margin as the feature are 91.8% and 48.3% [11]. Inaccurate segmentation results may lead to misdiagnoses which are based on boundary features. Hence, accurate thyroid nodule segmentation is the basis of differentiation between benign and malignant nodules in clinical applications. On the other hand, thyroid gland segmentation is fundamental for thyroid volume estimation. Thyroid volume is used to analyze the secretion of thyroid hormone, which is an important characteristic of thyroid abnormality diagnoses. However, it is difficult for clinicians to estimate these characteristics without the help of computers. As a result, thyroid gland segmentation and thyroid nodule segmentation methods are necessary to promote the study of thyroid disease diagnoses, providing valuable information for clinicians to make the best possible diagnostic decisions.

Ultrasound image segmentation has been investigated for decades since ultrasound devices became important diagnostic devices in clinical medicine in the 1980s [12]. Ultrasound image segmentation is a complicated problem, because of the ultrasound image characteristics such as attenuation, speckles, shadows, low contrast, and signal loss [13]. For example, Fig. 1 shows eight ultrasound images with thyroid glands and thyroid nodules. Fig. 1(a) ~ (d) are ultrasound images of thyroid glands with different shapes and sizes from different views, and Fig. 1(e) ~ (h) are ultrasound images with various thyroid nodules, where Fig. 1(b) is from longitudinal view and the others are from transverse view. As shown in Fig. 1, these thyroid glands and nodules differ in echogenicity, shape, size and boundary regularity. In the upper left area of the white rectangular frame in Fig. 1(c), as pointed out by a red arrow, the boundary between the thyroid and the muscle becomes blurred. And in the upper side of the white rectangular frame which is pointed out by another red arrow in Fig. 1(d), the texture of the muscle tissue is similar to that of the thyroid, forming a blurred boundary region. In addition, there are some isolated and highlighted calcifications in the thyroid ultrasound images, as pointed out by yellow arrows in Fig. 1(g) and (h). The inherent inhomogeneity of ultrasound images poses a challenge to develop accurate segmentation methods for ascertaining the shape features of objects in ultrasound images. The advances in ultrasound imaging technologies have constantly prompted the research on ultrasound image segmentation algorithms so as to keep up with the technological growth. Recently, image segmentation technologies have been widely used to segment ultrasound images of different human organs, such as thyroid, breast [14], prostate [15], carotid artery [16], bronchus [17], fetal left ventricle [18], etc.

This work mainly focuses on methods developed for thyroid gland segmentation and thyroid nodule segmentation in medical

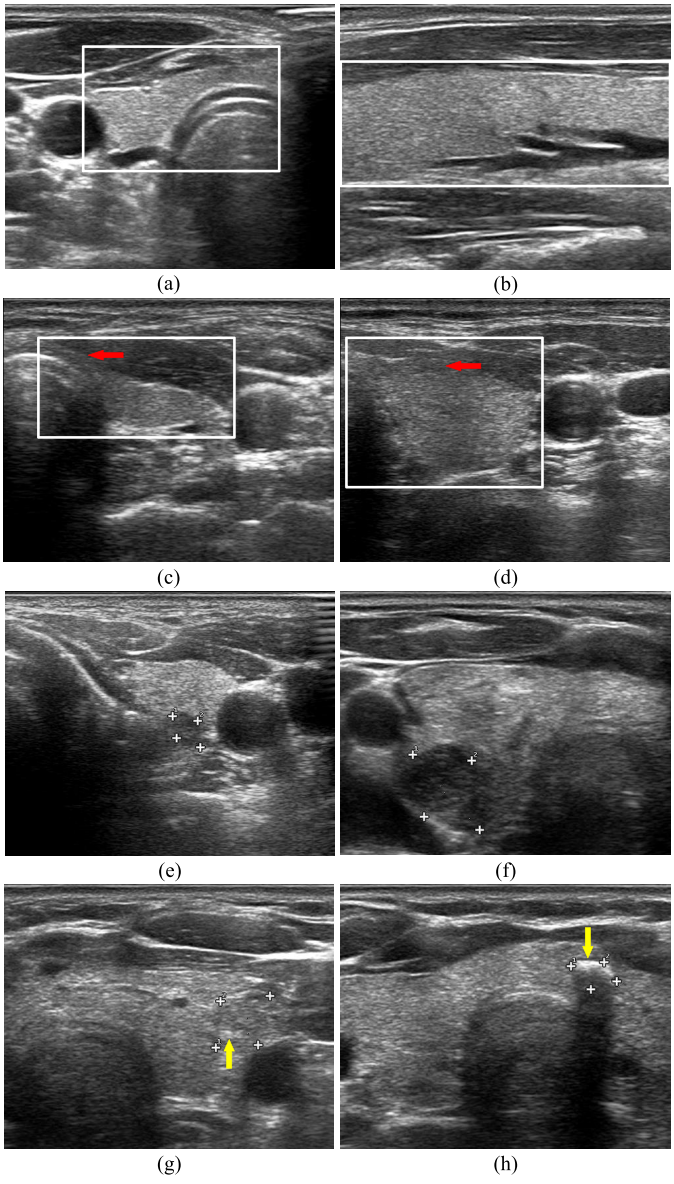

Fig. 1. Ultrasound images with thyroid glands and thyroid nodules. (a) ~ (d) are thyroid ultrasound images of thyroid glands with different shapes and sizes from different views, and the white rectangular frames indicate the thyroid gland regions. (e) ~ (h) are thyroid ultrasound images with various thyroid nodules with four white markers indicating the thyroid nodule regions. (b) is from longitudinal view and the others are from transverse view.

ultrasound images. The two different segmentation scenarios are shown in Fig. 2. Fig. 2(a) is a thyroid ultrasound image of a normal thyroid, Fig. 2(b) is the result of the thyroid gland segmentation, Fig. 2(c) is a thyroid ultrasound image of an abnormal thyroid containing a nodule, and Fig. 2(d) is the result of the thyroid nodule segmentation. Thyroid gland segmentation is usually used to obtain the locations and the sizes of the thyroid glands, which are useful to detect thyroid abnormalities. And thyroid nodule segmentation is usually used to obtain the shapes and boundaries of the lesion areas, which are useful to identify benign or malignant nodules. The classification of thyroid gland and thyroid nodule segmentation methods helps to select the appropriate segmentation methods for different application scenarios. Thyroid gland segmentation and thyroid nodule segmentation are two different applications. As the region of thyroid gland is much larger than the region of thyroid nodule, the region of thyroid gland in ultrasound

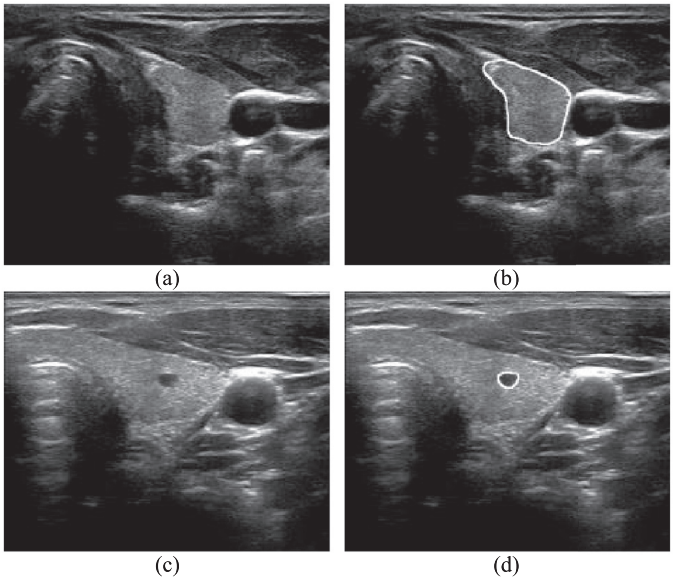

Fig. 2. Two different segmentation scenarios for thyroid gland segmentation and thyroid nodule segmentation. (a) Thyroid ultrasound image of a normal thyroid. (b) Segmentation result of the thyroid gland in (a). (c) Thyroid ultrasound image of an abnormal thyroid containing a nodule. (d) Segmentation result of the thyroid nodule in (c).

image is more susceptible to noise, spots, artifacts, and signal attenuation, usually exhibiting multiple texture properties for different areas within the thyroid gland region. Furthermore, there are other tissues around thyroid gland, such as trachea, blood vessels, muscles and so on. Hence, not well-defined margins and fuzzy boundary regions are formed by thyroid gland and surrounding tissues under the influence of artifacts and speckles. In such complex case, it is difficult to adjust the traditional contour and shape based models and region based models to the applicable scenario. Besides, the machine learning or deep learning models have sufficient ability to identify thyroid gland regions in this complex case. On the other hand, thyroid nodule segmentation is different from thyroid gland segmentation, because thyroid nodules are surrounded by thyroid gland tissues. Thyroid nodules have lower boundary complexity because at least one of the regions within and outside the nodule boundary is homogeneous. So traditional contour and shape based models and region-based models can get competitive segmentation results with smaller amount of computation as compared with machine and deep learning models.

Three words “thyroid, ultrasound, segmentation” are used as key words to search mainly in Web of Science, Engineering Village and PubMed databases, and more than 1000 related papers are obtained under this search strategy. Forty-four papers are found related to ultrasound thyroid gland segmentation and ultrasound thyroid nodule segmentation through lots of reading and investigation. Among these 44 papers, the works with detailed method descriptions and sufficient experimental results are most concerned. But the papers describing feasible and valuable methods are also selected, even though they do not demonstrate quantitative evaluation results. After careful investigation, 28 representative papers are selected for this review.

To the best of our knowledge, this is the first review on thyroid gland and thyroid nodule segmentation methods for ultrasound images. This work categorizes and reviews the existing thyroid segmentation technologies. In order to have a clear assessment of the segmentation performance of different methods, the methods are grouped according to the different theoretical computational approaches. The processes of technology developments are analyzed and the evaluation metrics are presented. Moreover, the evaluation metrics with different names but using the same calculation method are named uniformly for the convenience of analyses and comparisons. In summary, the major contributions of this work are as follows:

A) A review of thyroid gland segmentation and thyroid nodule segmentation methods in medical ultrasound images is first presented in this work, which reviews the development of the segmentation methods, analyzes the similarities and differences among the segmentation methods, and concludes the advantages and disadvantages of each segmentation method.

B) A categorization approach is proposed according to the theoretical bases of related segmentation methods. Such categorization approach is convenient for analyzing the segmentation methods.

C) The evaluation metrics for the reviewed segmentation methods are named uniformly. The segmentation results using the uniformly named evaluation metrics are compared, and selection of appropriate segmentation methods for specific applications are discussed.

The categorization of the segmentation methods will be introduced in Section 2. Thyroid gland segmentation methods will be described in Section 3, and the thyroid nodule segmentation methods will be presented in Section 4. Section 5 will further analyze the correlations between the reviewed thyroid gland and thyroid nodule segmentation methods, and the advantages and disadvantages of various methods. Quantitative evaluation of thyroid gland and thyroid nodule segmentation methods in 2D and 3D thyroid ultrasound images will be provided in Section 6. Method selection for specific applications will be discussed in Section 7. The conclusion and perspective will be given in Section 8.

2. Method categorization

According to different segmentation targets, the studies of thyroid ultrasound image segmentation are divided into two directions, namely, thyroid gland segmentation and thyroid nodule segmentation. Although the targets of the two segmentation applications are different, the method categorizations of the two segmentation applications are roughly the same. As shown in Fig. 3, the segmentation methods are classified into mainly four groups according to the theoretical bases taken to solve the problem, including contour and shape based methods, region based methods, machine and deep learning methods, and hybrid methods. The methods of thyroid gland segmentation contain three of the four groups except the hybrid methods, and the methods of thyroid nodule segmentation contain the four groups. By means of such categorization, the similarities and differences between the segmentation methods are clearly compared, and their advantages and disadvantages are easier to be concluded.

A) Contour and shape based methods: The methods in this category use the boundary or shape information of thyroid gland or thyroid nodule to segment thyroid gland or nodule in thyroid ultrasound images. But the boundaries between thyroid tissue areas and thyroid nodules are sometimes blurred due to low image contrast and image artifacts. Furthermore, thyroid nodules are often irregular. As a result, such methods usually need initial contours and prior shape information to improve the accuracy of boundary segmentation.

B) Region based methods: The methods in this category utilize the statistical properties of gray intensities between regions, such as mean value and variance, to obtain the minimized boundary energy function. These methods assume that the same tissue regions in the thyroid ultrasound images are homogeneous while different tissue regions are inhomogeneous. However, in the actual thyroid ultrasound images, the differences between different tissue regions are sometimes not significant, and the classifications of

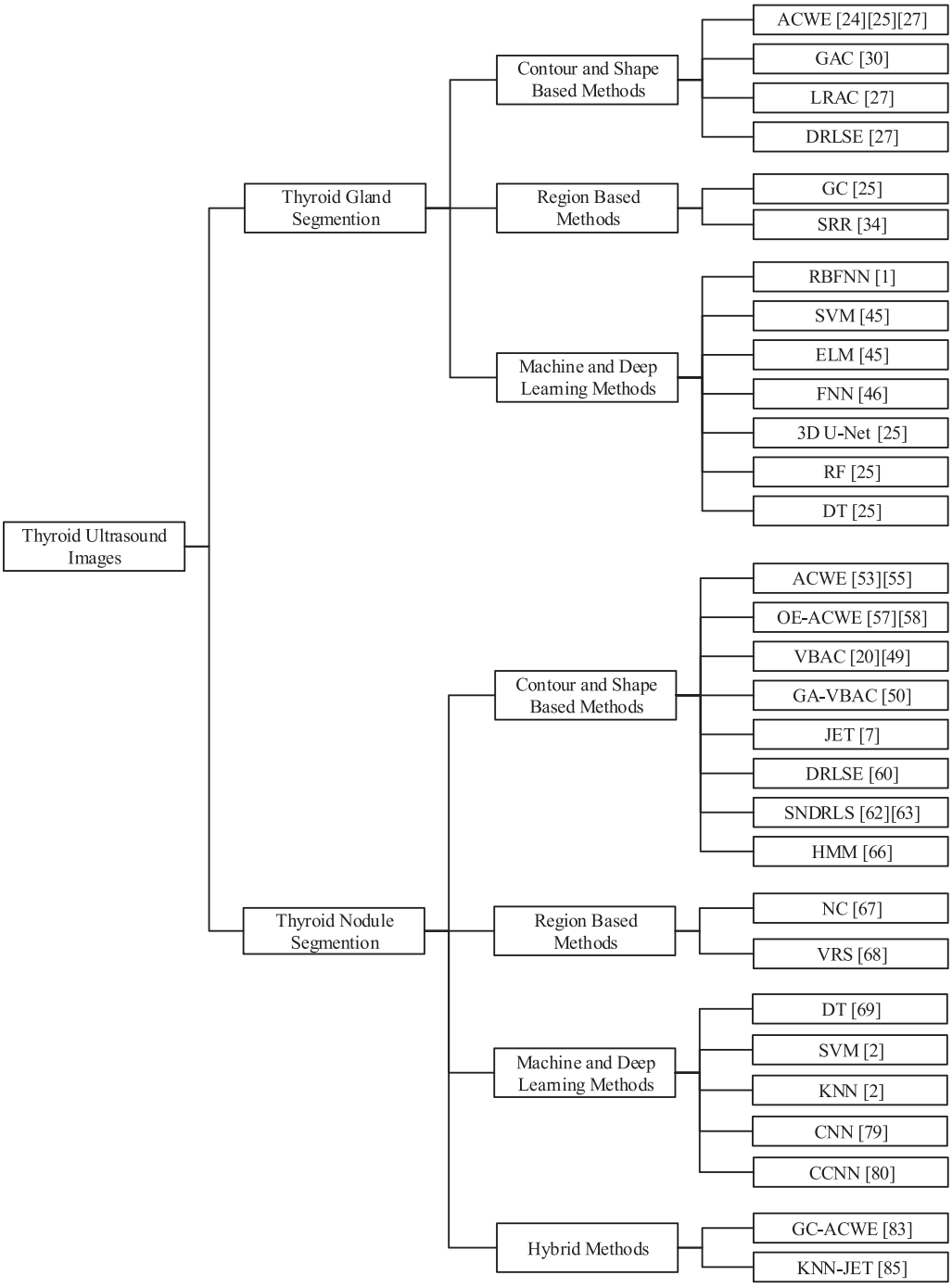

Fig. 3. The categorized methods in this work.

intersecting regions are ambiguous. Hence, such methods need spatial location information and prior shape information to realize the segmentation of target regions.

C) Machine and deep learning methods: The methods in this category construct classifiers to classify the pixels or image blocks in the thyroid ultrasound images, so as to realize the segmentation of target tissue areas. The constructions of classifiers are mostly based on machine and deep learning algorithms, including radial basis function neural network (RBFNN), feedforward neural network (FNN), support vector machine (SVM), extreme learning machine (ELM), random forest (RF), decision tree (DT), convolutional neural network (CNN) and so on. But the training process of the machine and deep learning classifiers is time consuming and requires a lot of training data and tags.

D) Hybrid methods: The methods in this category combine two or more of the above-mentioned methods to improve the process of segmentation, making the segmentation results more accurate or more robust. Given the situation of ensuring segmentation accuracy, automatic segmentation methods are undoubtedly more efficient and more objective than semi-automatic segmentation and manual segmentation methods.

The corresponding techniques involved in each group will be described and discussed in detail in the following sections. Brief introductions to the steps of various segmentation methods and

related technologies will be presented, and the strengths and drawbacks of these methods are also analyzed.

3. Thyroid gland segmentation methods

The thyroid gland segmentation methods are categorized into three groups, which are contour and shape based methods, region based methods, and machine and deep learning methods. Especially, the methods based on machine and deep learning algorithms are the dominant methods in thyroid gland segmentation.

3.1. Contour and shape based methods

The contour and shape based segmentation methods use the boundaries between the target regions and the background regions to accurately get the contour or the shape of target regions. The active contour based methods are commonly used, because these methods can topologically change the contour and have certain ability to distinguish fuzzy boundaries. The shape based methods have poor practicability, because such methods are effective on the premise that targets have uniform shape and regular boundaries. Each of the selected methods in this category will be described in the following.

The active contour model, such as the snake model [19], was a framework for delineating target tissues in ultrasound images. Such model utilized the internal energy and external energy to change the contour. The goal of active contour model was to obtain the corresponding object contour when the energy function was minimal. The internal energy was used to control the elastic deformation of the contour so as to maintain the continuity and smoothness of the contour [19]. The external energy derived from image features was image-dependent, which was used to force the contour towards the desired object boundary. During the deformation process, the force was calculated from internal energy and external energy. According to the representation and implementation characteristics of different active contour models, there were two types of active contour models, i.e., the parametric active contour models and the geometric active contour models [20]. Parametric active contour models cannot handle topological structure changes and need to be reinitialized, while geometric active contour models can capture multiple targets, automatically process geometric topological changes and directly expand to higher dimensional space, which have more common applications in medical image segmentation. In the following, variations of the geometric active contour model used in thyroid gland segmentation will be described.

Active contours without edges (ACWE) model [21] was proposed based on level set following Mumford-Shah segmentation method [22]. ACWE model involved integral operations which provided an inherent noise filtering mechanism [23], so ACWE model was relatively insensitive to noise. The ACWE model assumed that the foreground and the background of the image were homogeneous respectively, which meant that the pixel intensities in the target area and the non-target area were basically consistent with little variation respectively. The ACWE model divided the image domain into three parts: target region Ω_1, background region Ω_2 and target contour C. The energy function of ACWE model was as follows:

$$E(C, c_1, c_2) = \mu \int_C ds + \lambda_1 \iint_{\Omega_1} (I - c_1)^2 dx\, dy + \lambda_2 \iint_{\Omega_2} (I - c_2)^2 dx\, dy, \quad (1)$$

where I was the image intensity, c_1 and c_2 were the mean intensity in target region Ω_1 and the mean intensity in background region Ω_2, respectively. The first item in energy function was internal energy, which limited the size of the contour, so that the contour can be smooth and continuous. The second and third items in energy function were external energy consisted of inside contour energy and outside contour energy, which drove the contour getting close to the boundary of the object region according to the difference of image characteristics between the inside and the outside of the contour. The inside contour energy was expressed as the product of the inside contour energy weight and the variation of pixel intensity inside the contour, and the outside contour energy was expressed as the product of the outside contour energy weight and the variation of pixel intensity outside the contour. When the assumption of respective homogeneity of foreground and background was established, the object boundary was successfully extracted by active contour, with both internal energy and external energy reaching their minima.

Poudel et al. [24,25] used ACWE models to segment thyroid glands in 2D thyroid ultrasound images, and the segmented 2D thyroid images were used to reconstruct 3D thyroids to realize 3D thyroid image segmentation. In the image preprocessing stage, histogram equalization was used to adjust the contrast, and median filter was used to reduce speckle noise. In the step of thyroid segmentation, the thyroid glands in 2D ultrasound images were delineated using ACWE model, and then ImFusion Suite [26] was used to do 3D reconstruction. The test dataset consisted of ten 3D thyroid ultrasound images, and each 3D image contained multiple 2D thyroid ultrasound images. As a result, the test dataset had 1416 2D thyroid ultrasound images in total. ACWE model was used to segment 1416 2D thyroid ultrasound images, and the dice coefficient of its segmentation result was 80%.

Kaur and Jindal [27] compared three contour-based segmentation models, including ACWE, localized region based active contour (LRAC) [28], and distance regularized level set evolution (DRLSE) [29]. LRAC used local region parameters to construct energy function for segmentation segmentation. In this approach, the foreground and background were described in terms of smaller local areas. Each point was considered separately to optimize local energy. DRLSE utilized distance regularization term and external energy term to drive contour to the desired edges. They used these three models to segment scintigraphy images and thyroid ultrasound images. LRAC had the shortest iteration time and the highest accuracy on 10 scintigraphy images. But they did not provide quantitative evaluation results for thyroid gland segmentation in ultrasound images.

Kollorz et al. [30] proposed a semi-automatic method for segmenting thyroid glands in 3D thyroid ultrasound images. Their method was divided into two steps: image preprocessing and image segmentation. In the process of ultrasound imaging, the propagation of ultrasonic waves in human tissues gradually degrades, so that the pixel intensities of deep tissues in thyroid ultrasound images are smaller than their actual values. Therefore, in the image preprocessing step, Kollorz et al. applied a linear ramp for pixel intensity correction, and utilized curvature flow (CF) and gradient anisotropic diffusion (GAD) [31] filters respectively to reduce speckle noise and to preserve object edges. In the image segmentation step, geodesic active contour (GAC) [32] was used to segment the entire 3D volume of the thyroid lobe. GAC was based on level sets, which utilized the gradient of the image to construct the edge stopping function to ensure that the evolution curve stopped at the boundary of the object. Meanwhile, they proposed an improved model, geodesic active contour with mean extension (GACM), which added the mean pixel intensity of the initial contour region to GAC model to suppress leakage at low-contrast edges. In the experiments, different filters and segmentation methods were combined to realize segmentation of 3D thyroid lobes, such as CF-GAC, CF-GACM, GAD-GAC, and GAD-GACM. Ten 3D thyroid lobes were used to test these four combinations. It was observed that all combinations performed similarly in the sensitivity evaluation metric, but the specificity of

combinations using GACM as segmentation method were higher than those of combinations using GAC as segmentation method.

3.2. Region based methods

Region based methods use the statistical properties of regions to distinguish between different tissues. In thyroid ultrasound images, the brightness and the distribution of pixels in thyroid gland region are different from those in other tissue regions (e.g. carotid, trachea, muscles). This makes it feasible to segment the thyroid glands in thyroid ultrasound images using region based methods.

Poudel et al. [25] used graph cut (GC) [33] method to segment the thyroid gland in a 2D thyroid ultrasound image. Graph cut was a very useful and popular energy optimization algorithm, which associated the image segmentation problem with the min cut problem of the graph. But this method was a semi-automatic method requiring manual sketches of the foreground and the background out of the ultrasound images. The graph cut model was used to segment 1416 2D thyroid ultrasound images, and the dice coefficient of its segmentation result was 76.5%.

Narayan et al. [34] proposed a method for automatic segmentation of multiple organs in 2D ultrasound images, which was based on the similarity of the echogenicity of each organ. Their method could segment carotid, trachea, muscles, and thyroid glands from ultrasound images. In ultrasound images, the echogenicity referred to the relative brightness between different tissues [35]. When comparing the echogenicity characteristics of different tissues, such tissues were divided into hyperechoic, isoechoic, and hypoechoic tissues. For example, the thyroid gland was hyperechoic as compared to the muscle, and it was hypoechoic as compared to the artifacts. The same organ tissue contained one or multiple isoechoic regions, so the ultrasound images were divided into multiple isoechoic regions, and each isoechoic region was regarded as a similar reflective region (SRR). For a thyroid ultrasound image $f(x)$, Hessian matrix and second-order partial derivative test [36] were used to get the speckles corresponding to the occurrence of local minima at these points. These speckles were classified into the same or different SRRs by pairwise comparing the difference of $M \times M$ pixel regions around the speckles. After obtaining the SRRs in a thyroid ultrasound image, the pixel intensity in mth SRR was set to the average pixel intensity of the mth SRR to get the quantified image $f_Q(x)$. The SRR with the maximum average pixel intensity corresponded to the enhancement artifacts below the carotid in a thyroid ultrasound image. As a result, the region of carotid artery was located above the enhancement artifacts. The carotid was segmented by the local phase congruency [37,38] and the circular shape constraints of candidate regions. The region of trachea was segmented by the connected component analysis of binary images obtained by the local phase. Muscles were above the carotid and the thyroid gland region, and below the layers of the hyperechoic fatty tissues in a thyroid ultrasound image. The region of muscles was segmented by the connected component analysis of binary images obtained by the local phase, which was similar to the segmentation of the trachea region. At the mean time, the spatial positions of the carotid, trachea, and muscles were used to segment the thyroid gland region. The thyroid gland was between the carotid and the trachea, and below the muscles. A digital line was obtained by the the central point of carotid and the coordinate of the highest point of trachea [39], which determined the edge pixels of the lower boundary of the thyroid gland. The lower boundary of the thyroid gland was obtained by fitting the spline to the determined edge pixels. In the experiments, 52 2D thyroid ultrasound images including transverse scans and longitudinal scans were segmented using the proposed method, and the average dice coefficient obtained by comparing the segmentation results with two ground truth datasets sketched by two experts was 84.47% and 83.23%, respectively.

3.3. Machine and deep learning methods

In the field of ultrasound image segmentation, most of the methods based on machine and deep learning algorithms were more efficient and more accurate than the traditional segmentation methods. Many machine and deep learning algorithms have been applied in the field of thyroid ultrasound image segmentation, including RBFNN, SVM, ELM, FNN, CNN, etc. The methods based on machine and deep learning algorithms performed well in thyroid gland segmentation, but the main limitations and bottlenecks of these methods were that these methods required plenty of marked datasets and long training time. Recently, researchers have applied several supervised learning models to segment thyroid glands in thyroid ultrasound images, which will be described in the following.

Chang et al. [1] proposed an automatic method to segment thyroid glands in 2D thyroid ultrasound images, which was based on RBFNN. The 3D volumes of thyroids were estimated based on the 2D segmented images. The process of this segmentation method consisted of four steps, including locating probable thyroid region and image enhancement, feature extraction, training RBFNN, and thyroid recovery. The image preprocessing step included locating probable thyroid region and image enhancement. The purpose of image preprocessing step was to reduce speckle noise, enhance ultrasound image quality, and further reduce the amount of computation. For a 2D thyroid ultrasound image, the vertical projection was first used to locate the probable thyroid gland region. And then adaptive weighted median filter (AWMF) [40] and morphological operations were used to reduce speckle noise. Subsequently, gray-level compensation was utilized to adjust the contrast between thyroid gland and the background. In the feature extraction step, six discriminative texture features were adopted, including the mean of haar wavelet (HAARM) feature, the variance of haar wavelet (HAARV) feature, coefficient of local variation (CV) feature, histogram feature, block difference of inverse probability (BDIP) feature [41], and normalized multiscale intensity difference (NMSID) feature [42]. The RBFNN consisted of three layers, including input layer, hidden layer, and output layer. The stochastic gradient-based supervised learning algorithm [43] was adopted to construct and train the RBFNN. The purpose of thyroid recovery was to eliminate serrated edges and smooth delineated contours. Multiple structuring elements were used to recover the shape of thyroid gland region. Finally, the particle swarm optimization algorithm [44] was used to estimate the volume of the whole thyroid gland based on the segmented 2D thyroid ultrasound image. In the experiments, 60 training patterns including 30 thyroid tissues and 30 non-thyroid tissues were adopted to train the RBFNN, and five 2D thyroid ultrasound images were used as test images to evaluate the performance of the proposed segmentation method. The average accuracy of the test dataset was 96.52%.

Selvathi and Sharnitha [45] realized the automatic segmentation of thyroid glands in 2D ultrasound images with support vector machine and extreme learning machine methods. The whole segmentation system was divided into two related sub-systems, i.e., thyroid nodule recognition and thyroid gland segmentation sub-systems. When the thyroid nodule recognition sub-system detected no thyroid nodules in the ultrasound image, the thyroid gland segmentation sub-system was then activated. The thyroid gland segmentation sub-system consisted of three parts, including image enhancement, feature extraction, and segmentation model using SVM and ELM methods. The purpose of image enhancement was to reduce speckle noise, enhance the quality of image, and reduce the amount of computation. The image enhancement part

consisted of four steps, including locating the probable thyroid gland region, AWMF, morphological operations, and gray-level compensation. In the feature extraction part, six features were adopted, including HAARM, HAARV, CV, histogram feature, BDIP, and homogeneity feature. SVM and ELM were utilized to construct two separate classifiers, which were adopted respectively to classify $M \times M$ image blocks on thyroid ultrasound images into thyroid region and the background. In the experiments, the segmentation accuracy of methods using SVM and ELM classifiers was 84.78% and 93.56%, respectively.

Garg and Jindal [46] used FNN to automatically segment thyroid glands in 2D thyroid ultrasound images. Their method consisted of four steps, including image enhancement, feature extraction, training of FNN, and classification. Image enhancement in this method had three steps, including locating the probable thyroid gland region, AWMF, and morphological operations. Nine texture features were extracted from selected thyroid regions and non-thyroid regions, which were HAARM, HAARV, histogram mean, histogram standard deviation, histogram variance, histogram entropy, histogram skewness, histogram kurtosis, and histogram energy. They utilized FNN [47] which was trained with stochastic gradient-based supervised learning algorithm to construct the classifier. A total of 45 training patterns including 20 thyroid tissues and 25 non-thyroid tissues were used to train the FNN. On the test dataset of five 2D thyroid ultrasound images, the average accuracy of segmentation results was 96.51%.

Poudel et al. [25] applied a 3D U-Net CNN [48] model to automatically segment the entire thyroid glands in 3D thyroid ultrasound images. The 3D U-Net CNN model was composed of an encoder and a decoder. The encoder was responsible for identifying and classifying the pixels in ultrasound images, and the decoder was used to locate the specific locations of the pixels. The dataset contained ten 3D thyroid ultrasound models which had 1416 2D thyroid ultrasound images in total. The data was augmented through image rotation, scaling, flipping and translation to prevent the 3D U-Net CNN model from overfitting. A dropout of 25% after each pooling layer was added to discard unnecessary neurons and Adam optimizer was utilized to adjust the learning rate to avoid local optimum. Furthermore, four other methods for thyroid gland segmentation were also implemented, which were ACWE, GC, RF and DT. Among these methods, 3D U-Net CNN and RF methods were implemented as automatic segmentation methods, and ACWE, GC and DT methods were implemented as non-automatic segmentation methods. DT model could be implemented as a fully automatic model, but user interactions were added into the model to improve the accuracy of segmentation. The segmentation performance of these methods were analyzed and compared in the experiments, and the 3D U-Net CNN model demonstrated the highest average dice coefficient of the segmentation results, which was 87.6%.

4. Thyroid nodule segmentation methods

The thyroid nodule segmentation methods are categorized into four groups, including contour and shape based methods, region based methods, machine and deep learning methods, and hybrid methods. Especially, most of the thyroid nodule segmentation methods are based on contour and shape.

4.1. Contour and shape based methods

Maroulis et al. [20,49] proposed a variable background active contour (VBAC) model based on ACWE model for the segmentation of nodules in 2D thyroid ultrasound images. The theoretical basis of VBAC model was the same as ACWE model. The limitation of the ACWE model was that it assumed that the foreground and the background in the thyroid ultrasound images are homogeneous respectively. However, the background in the thyroid ultrasound images was inhomogeneous with speckle noise and other tissue inhomogeneities. Therefore, the VBAC model was proposed to overcome the limitation of the ACWE model. The pixels representing speckle noise and other homogeneous tissues were excluded when calculating the outside contour energy of the energy function. In this way, the VBAC model used variable background regions to reduce the impact of inhomogeneous tissues in ultrasound images. In the experiments in [20], the VBAC model was used to segment 71 thyroid nodule ultrasound images, and the mean overlap value of VBAC was 91.1%. The ACWE model was also implemented for the segmentation of thyroid nodules in ultrasound images as comparison, and the mean overlap value of ACWE model for segmenting 71 thyroid nodule ultrasound images was 84.8%.

The ultrasound image segmentation methods based on active contour model are device dependent. Therefore, different parameters are required to construct the active contour based segmentation models for ultrasound images acquired from different devices or from the same devices using different settings (e.g. contrast, brightness). As a result, automatic parameter tuning not only saves the time of manual parameter tuning, but also makes the model more robust and applicable. Iakovidis et al. [50] proposed genetic algorithm-variable background active contour (GA-VBAC) model to segment thyroid nodules in 2D thyroid ultrasound images. GAs were stochastic non-linear optimization algorithms based on the theory of natural selection and evolution [51,52], which were combined with VBAC model to realize automatic parameter tuning so as to enhance the clinical practicability of active contour model. The experimental dataset consisted of 45 thyroid nodule ultrasound images, and 45 experiments were conducted. In each experiment, a single image with ground truth delineated thyroid nodules in the dataset was used to train the GA module, so that the GAs obtained an optimal set of parameters for the VBAC module. And the rest 44 thyroid nodule ultrasound images in the dataset were used for testing. The ground truth images of the experiments were determined by this rule: For each ultrasound image, ground truth delineation was obtained following the rule that each pixel was considered as part of the nodule when it was included in at least two out of three experts' delineations. The ranges of the overlap values obtained by the automated GA-VBAC method and by the experts with respect to the ground truth delineations were evaluated. GA-VBAC led to overlap value ranges which fell within the experts' overlap value ranges in 22 teat cases, led to overlap value ranges which exceeded the experts' overlap value ranges in 17 test cases, and led to overlap value ranges which fell below the experts' overlap value ranges in 6 test cases. The mean overlap value of GA-VBAC was 92.5% and the mean overlap value of experts was 91.8%. The segmentation accuracy obtained was comparable to or even higher than that obtained by individual expert radiologists.

Speckle noise in ultrasound images makes segmentation process more complicated. Filter technique is a common way to reduce noise so as to improve the segmentation results. Although ACWE model is relatively insensitive to noise, high quality ultrasound images can make the segmentation results more accurate. Nugroho et al. [53] utilized bilateral filters [54] to preprocess 2D thyroid ultrasound images, which was a low pass filter that smooths images and preserves the boundaries of objects. ACWE model was then applied to segment thyroid nodules in preprocessed ultrasound images. The experiments showed that bilateral filtering can improve the segmentation accuracy of active contour and make the localization of nodules clearer and unequivocal.

The contour length constraint in the energy function of ACWE model may not maintain enough properties to segment the objects in some complex situations [55]. For example, the object tissues and the surrounding tissues in the images are usually mixed to-

gether due to the low quality and noise of ultrasound imaging. Gui et al. [55] proposed a medical image segmentation method based on level set and isoperimetric constraint, and applied this method to medical images such as ultrasound, computed tomography and magnetic resonance images, including thyroid ultrasound images. In thyroid ultrasound images, the nodules were often compact shaped. Therefore, the contour length constraint in the ACWE model was replaced with a compact shape constraint based on the isoperimetric constraint. This method can deal with boundary ambiguity or partially missing targets. In the experiments, 52 thyroid ultrasound images were used as the test set, and the dice coefficient of this method reached 89.6%.

In ACWE model, the energy function (1) is composed of internal energy and external energy. The internal energy is contour regularization term, and the external energy is data fidelity term. Contour regularization and data fidelity have corresponding weight coefficients. The larger the weight coefficient, the stronger the corresponding force during contour evolution. In structure tensor field [56], multiple orientations of structure tensors appear in non-target, unstructured edge regions, whereas approximately constant orientation of structure tensors appears in target, structured edge regions. Mylona et al. [57] [58] utilized the orientation entropy (OE) to automatically set the weight coefficients of contour regularization and data fidelity items in ACWE model. OE obtained low value in structured regions, in which the orientation variability was low. And OE obtained high value in unstructured regions, which contained edges of multiple orientations. The weight coefficient of contour regularization was inversely proportional to the weight coefficient of data fidelity. When the active contour was far away from the boundary regions, the weight coefficient of the data fidelity was high, which accelerated the evolution of the active contour towards the boundary. When the active contour was close to the boundary region, the weight coefficient of the data fidelity gradually decreased, while the weight coefficient of the contour regularization gradually increased, making the contour close to the target edge and keep the contour continuous and smooth. In the experiment, 20 thyroid ultrasound images were used to test the segmentation performance of this OE-ACWE model. The evolution speed of OE-ACWE model was faster than ACWE and the mean overlap of segmentation result was 83.70%

Joint echogenicity-texture (JET) model was similar to ACWE model and followed Mumford-Shah segmentation method, which was proposed by Savelonas et al. [7]. JET model can accurately segment the nodules in 2D thyroid ultrasound images according to the echogenicity and the texture of thyroid nodules. Incorporating texture information into the active contour model enabled the model to get some findings which cannot be distinguished by the variation of pixel intensity and the average pixel intensity, such as isoechoic thyroid nodules. Local binary pattern (LBP) distribution [59] was efficient and effective to represent the image texture information, and Mumford-Shah function was used to co-evaluate regional image intensity and LBP distribution. Moreover, new log-likelihood goodness-of-fit terms were used to evaluate the differences between the internal and external areas of the active contour model. In the experiments, the mean overlap value of the JET model was 92.9% when JET model was applied to segment 38 thyroid ultrasound images featuring hypoechoic nodules, and the mean overlap value was 91.5% when applied to segment 36 thyroid ultrasound images featuring isoechoic nodules. Hence, the JET model can accurately segment thyroid nodules regardless of their echogenicity, and the mean overlap value of the JET model for the 74 thyroid ultrasound images was 92.21%.

The energy functions of geometric active contour models were derived based on image region characteristics (e.g. ACWE, VBAC, JET) or edge characteristics. The energy function based on edge characteristics utilized speed stopping constraint to stop the active contour at the edge. Du and Sang [60] proposed an improved DRLSE method [29] to segment thyroid nodules in 2D thyroid ultrasound images. The method was divided into three steps, including image preprocessing, extracting boundaries, and image postprocessing. The speckle reducing anisotropic diffusion (SRAD) filter [61] was adopted to enhance the images in the preprocessing step. The speed stopping constraint had some drawbacks in original DRLSE model. Firstly, due to the existence of speckle noise and artifacts in ultrasound images, the edge detection effect of classical speed stopping constraint was poor. Secondly, boundary leakage might occur because the classical speed stopping constraint was never equal to zero. Therefore, Du and Sang improved the calculation method of the speed stopping constraint so that it could be equal to zero so as to avoid boundary leakage in the boundary extraction step. Due to the heterogeneity within the thyroid nodules, the generated segmentation image had some holes in it. Hence, category relationships between adjacent pixels were utilized to eliminate the holes in the postprocessing step. In the experiments, they used 20 thyroid nodule ultrasound images to evaluate the segmentation performance, and the mean overlap value was 79.0%.

The DRLSE model performed poor in segmentation of weak edges and was sensitive to the locations of manually initialized contours. These defects made it difficult for DRLSE to automatically segment thyroid nodule ultrasound images. Koundal et al. [62,63] proposed an automatic segmentation method based on spatial neutrosophic distance regularized level set (SNDRLS) model in 2D thyroid ultrasound images. In their method, thyroid boundaries detection (TBD) algorithm [64] was used to first locate the approximate area of thyroid tissue. The SNDRLS model utilized the neutrosophic L-means (NLM) clustering method [65] to incorporate spatial information into the DRLSE model, which consisted of two steps. Firstly, the contour of the nodule area in the ultrasound image was approximated by the spatial NLM (SNLM) clustering method with spatial information. Subsequently, the SNLM was used to initialize and regularize the level set by estimating the parameters in the evolving equation. SNDRLS model integrated the advantages of SNLM clustering and level set methods, which was capable of delineating multiple nodules and had strong resistance to noise. In the experiments in [62], the SNDRLS model was used to segment 42 thyroid nodule ultrasound images, and the mean overlap value of SNDRLS was 93.15%.

There are few shape based segmentation methods for thyroid ultrasound images, which are limited by prior shape information in practical application scenarios. Tsantis et al. [66] proposed a hybrid multiscale model (HMM) to segment thyroid nodules in 2D thyroid ultrasound images, which combined wavelet-based edge detection method and Hough transform. Dyadic wavelet transform was used to extract edges in ultrasound images to obtain contour representation images. Hough transform was adopted to transform Cartesian image space of contour representation images to Hough parameter space. The points which were collinear in Cartesian image space intersected at a common point in Hough parameter space. Therefore, it was easy to find the points that were collinear in Cartesian image space when viewed in Hough parameter space. The prior circular shape of nodule made the value of the common points formed by the circular boundaries of nodule in Hough parameter space being the maximum. The segmentation performance of this method was quantitatively assessed using 40 thyroid ultrasound images. The segmentation results were compared with the manual segmentation images by two experts, and the sensitivity was 88.83% and 87.58%, respectively.

4.2. Region based methods

Graph theory based methods are common image segmentation technologies, including GC, normalized cut (NC) and so on. The

biggest advantage of NC is that it is not prone to small region segmentation. However, NC model occupies a large amount of computer memory space, whose weight matrix requires a large amount of calculation. And it often generates over segmentation and under segmentation results. Zhao et al. [67] proposed a thyroid nodule segmentation method based on NC. They introduced homomorphic filtering, anisotropic diffusion and fractional differential homomorphic filtering to the NC model. Homomorphic filtering was used for noise reduction, anisotropic diffusion was utilized to preserve important edge details, and fractional differential homomorphic filtering reduced the amount of calculation of the weight matrix. This model could segment the nodules and the trachea areas in the thyroid ultrasound images, but the authors did not provide quantitative evaluation results.

Region based methods use the statistical properties of regions to distinguish between different tissues. Alrubaidi et al. [68] proposed a method based on variance reduction statistic (VRS) to segment thyroid nodules in 2D thyroid ultrasound images. Given regions of interest (ROI) containing nodules, VRS method was utilized to roughly determine the edge points of nodules in radial lines. Subsequently, the nodule edge points were connected to generate the shape of the nodule by selecting the nearest neighbor points. Finally, B-spline method was applied to improve the accuracy of the segmentation curve. The segmentation performance of this method was evaluated for the delineation of thyroid nodules of 26 ultrasound images. The sensitivity of the testing results was 82.49%.

4.3. Machine and deep learning methods

Chang et al. [69] used DT model to segment thyroid nodules in 2D thyroid ultrasound images. The method consisted of two parts, including image preprocessing and image segmentation. In the image preprocessing stage, histogram equalization was first used to enhance the contrast between the nodule and background. The suspicious nodular areas were then determined by horizontal and vertical projections. Subsequently, eleven feature extraction technologies were used to extract 41 features. The feature extraction technologies included co-occurrence matrix [70], statistical feature matrix [71], gray level run-length matrix [72], Laws' texture energy measures, neighboring gray level dependence matrix [73], homogeneity, histogram, BDIP, discrete cosine transform, NMSID, and haar wavelet. In image segmentation stage, decision trees were applied to build the classifier, and the constructed classifier was utilized to classify $M \times M$ image blocks into nodular or background regions. In the experiments, six 2D thyroid ultrasound images were used as the test dataset and the accuracy of segmentation results was 97.5%.

Keramidas et al. [2] put forward a thyroid nodule detection (TND) system, which was applied to analyze 2D thyroid ultrasound images and videos. The detection process of TND system was divided into five steps, which were data preprocessing, ROI definition, feature extraction, classification, and postprocessing. In the step of data preprocessing, pixel values were redistributed to cover the entire display brightness range of each image or video frame [74]. In the ROI definition step, the TBD-2 algorithm [64] was used to obtain the approximate area of thyroid tissue. In the feature extraction step, fuzzy local binary patterns [75] feature and fuzzy grey-level histograms [76] feature were adopted. In the classification step, two machine learning algorithms were utilized to build classifiers, which were SVM algorithm [77] and k-nearest neighbors (KNN) algorithm respectively. There were misclassified pixels and regions in the output binary images of the classifier. Hence, the majority voting decision criterion [78] was adopted to reduce such misclassification effect in the postprocessing step. The accuracy of thyroid nodule segmentation obtained by SVM was 91.3% with the test dataset containing 118 2D thyroid nodule ultrasound images.

In recent years, deep neural networks have performed well in semantic segmentation, object recognition and detection. Ma et al. [79] used deep convolutional neural network to segment the thyroid nodules in 2D thyroid images. The proposed CNN model contained fifteen convolutional layers and two max pooling layers. The input image size of CNN model was 353×353 and the output image size was 44×44. The dataset contained 22123 ultrasound images acquired from different sonographics systems. In the experiments, tenfold cross validation was used to comprehensively and quantitatively evaluate the CNN model. Among these ten folds, eight folds were used as training set, one fold was for validation, and the other fold was for testing. The learning rate was set to 0.0002, and the batch size was 32. 103.86 h were spent on training the CNN model, and the mean overlap value of the trained CNN model which was used to segment the test set was 86.83%.

A phased CNN model called cascaded convolutional neural network (CCNN) model was proposed by Ying et al. [80], which was different from the end-to-end CNN model described in [79]. CCNN model was composed of three phases to segment thyroid nodules. Firstly, U-Net based model was utilized to extract ROIs containing thyroid nodules, which were rectangular regions located near the center of the whole images. Secondly, referring to the ultrasonic examination for patients made by doctors, artificial marks were made through manual sign recognition and boundary adjustment method [81]. These artificial marks were made in the upper, lower, left and right parts of the thyroid nodules according to the ROIs. Finally, a fully convolutional neural network based on VGG-19 [82] was used to segment thyroid nodules in the ROIs. In the experiments, the dataset of thyroid ultrasound images contained 1000 images, among which 800 were used as the training set and 200 as the testing set. The mean overlap value of the segmentation result on testing set was 87.00%.

4.4. Hybrid methods

The purposes of hybrid methods for ultrasonic thyroid nodule segmentation included the optimization of the segmentation model and the realization of automatic segmentation process.

In terms of optimizing the segmentation model, Zhou [83] combined GC and ACWE models to obtain a new model so as to better segment thyroid nodules in 2D thyroid ultrasound images. Firstly, inside contour energy and outside contour energy items of ACWE energy function were discretized by introducing a binary variable to mark the position of each point in the image. The binary variables of the points inside the contour were less than zero, the binary variables of the points outside the contour were greater than zero, and the binary variables of the points belonging to the contour were equal to zero. The energy function of GC model was also discretized by the binary variables. In fact, the binary variables had the same effect as the original Heaviside function in ACWE model. Secondly, the contour length item of ACWE energy function was discretized by geo-cuts [84] algorithm. Finally, the discrete ACWE model and the discrete GC model were combined by comparison to form a new energy function, thus the GC-ACWE model was established. The GC-ACWE model can reduce nodes and edges in the graph structure, improve computing speed and prevent boundary leakage. According to the segmentation images presented in the experiments, GC-ACWE model had better segmentation performance than ACWE model, which could cope with boundary leakage. However, no quantitative performance results were given in the experiments, limiting the comparisons of GC-ACWE model with other models.

In terms of realizing automatic segmentation process, Legakis et al. [85] proposed a method to automatically segment thyroid

nodules in thyroid ultrasound images, which combined the machine learning algorithm and the active contour model. The TBD algorithm was first used to roughly detect the boundaries of thyroid glands. Then the rough regions of thyroid nodules were detected within the boundaries of thyroid glands. The thyroid nodular region detection was done by the classifier constructed by KNN algorithm and a texture feature of LBP distribution. The obtained thyroid nodule regions were used as the initialization of the subsequent active contour model. The active contour model JET was adopted in this method, which combined the texture and the echogenicity of the ultrasound images to segment the nodules. The segmentation performance of this method was evaluated for thyroid nodule segmentation in 142 ultrasound images, and the mean overlap value of the segmentation results was 92.30%.

5. Analysis of reviewed methods

5.1. Thyroid gland segmentation methods

Table 1 summarizes the advantages and disadvantages of the reviewed thyroid gland segmentation methods and their applications in 2D and 3D segmentation scenarios. The analyses of the reviewed methods include the characteristics of the algorithms and their implementation manners. These thyroid gland segmentation methods were all automatic except ACWE, GC and DT methods. The reviewed thyroid gland segmentation methods were used in two application scenarios, including delineating the boundary of 2D thyroid gland and segmenting the entire 3D volume of the thyroid lobe. There were two kinds of 3D thyroid image segmentation. One was directly segmenting 3D thyroid glands in 3D thyroid models, and the other was first conducting 2D thyroid segmentation and then reconstructing 3D segmentation results based on the segmented 2D thyroid images. Among the reviewed methods shown in Table 1, GAC and 3D U-Net methods directly segmented 3D thyroid glands in 3D thyroid ultrasound images. Other 3D thyroid segmentation methods were obtained based on 2D thyroid segmentation results.

As contour evolution could not reach all the regions of thyroid and the initialization of the contour might be wrong, when applying ACWE model to segment 2D thyroid ultrasound images, undersegmented and oversegmented results would appear in some cases [25]. As a result, user interactions were required to stop the iteration or reinitialize the initial mask during contour evolution. Therefore, the segmentation process of ACWE model took a long time due to user interactions. ACWE model was also applied in 3D thyroid segmentation scenario, which was based on 2D thyroid segmentation results.

GAC method was an automatic segmentation method to directly segment 3D thyroid glands in 3D thyroid ultrasound images. Since GAC model was sensitive to noise and boundaries in images, noise reduction and image enhancement were necessary to generate higher quality 3D ultrasound images. Using GAC model to evolve the contours directly in 3D space greatly increased the calculations, which increased the computation time. However, direct segmentation of 3D thyroid glands in 3D space can effectively utilize spatial information to improve the accuracy of 3D thyroid gland segmentation.

The LRAC model used local image statistics to replace the global gray mean in the ACWE model, which improved the ACWE model and was able to segment non-uniform gray level images. In addition, level set regular term was added to the energy function so as to avoid the reinitialization of the level set. In the DRLSE model, distance regularization item was added into the energy function of level set as a penalty item, which allowed more simple initialization of the level set and reduced the dependence on the initial contour in image segmentation. ACWE, GAC, LRAC and DRLSE models were all contour based segmentation models. However, the segmentation results of insufficient segmentation or edge overstep the boundary were prone to occur when the target edge areas were complex.

GC method required more user interactions when it was used to segment 2D thyroid glands [25]. Users could visualize current GC segmentation results in real-time and improve the segmentation results through interactions. Therefore, for 2D thy-

Table 1
Analysis of thyroid gland segmentation methods.

	Method	Advantages	Disadvantages	2D	3D
Contour and Shape Based Methods	ACWE	Insensitive to noise	Require manual initialization of the mask, require user interactions during contour evolution	✓	✓
	GAC	Automatically and directly segment 3D thyroid glands	Require noise reduction and image enhancement, large amount of calculations	–	✓
	LRAC	Segment no-uniform gray images	Require manual initialization of the mask, manual parameter tuning	✓	–
	DRLSE	Low dependence on initialization boundary	Require manual initialization of the mask, manual parameter tuning	✓	–
Region Based Methods	GC	Easy to implement, robust with user interactions	Require user interactions	✓	✓
	SRR	Automatic multi-organ segmentation, combine spatial locations and characteristics of organs, consider noise and echogenicity information	Rough segmentation boundaries, limited by imaging characteristics	✓	–
Machine and Deep Learning Methods	RBFNN	Automatic segmentation	Require marked training data, require image preprocessing and postprocessing, sensitive to selected features	✓	✓
	SVM	Automatic segmentation	Require marked training data, require image preprocessing, sensitive to selected features	✓	–
	ELM	Automatic segmentation	Require marked training data, require image preprocessing, sensitive to selected features	✓	–
	FNN	Automatic segmentation	Require marked training data, require image preprocessing, sensitive to selected features	✓	–
	3D U-Net	Automatic segmentation, able to learn without handcrafted features, directly segment 3D thyroid glands	Require plenty of marked training data, long training time	✓	✓
	RF	Automatic segmentation	Require marked training data, require image preprocessing and postprocessing, sensitive to selected features	✓	✓
	DT	Easy to implement, low computational complexity	Require marked training data, require image preprocessing and postprocessing, sensitive to selected features	✓	✓

roid ultrasound images, GC segmentation results were directly proportional to the number of user interactions. Furthermore, user interactions made GC model more robust. GC method can reconstruct 3D thyroid segmentation results based on 2D thyroid segmentation results.

SRR could automatically segment multiple organs in 2D thyroid ultrasound images, such as carotid, trachea, muscle, and thyroid gland. SRR utilized echogenicity and relative positions between organs to segment each organ. Assuming that the same organ tissue contained one or more isoechoic regions, the 2D thyroid ultrasound images were divided into multiple isoechoic regions. The similar isoechoic regions were merged by the local features of noise points. The corresponding organs were then located based on these similar isoechoic regions. The organ segmentation results of this method were somewhat rough, and were limited by the imaging characteristics such as brightness, contrast, roughness and ultrasound scanning angle of the thyroid ultrasound images.

The characteristics and segmentation procedures of thyroid gland segmentation methods based on machine and deep learning algorithms were basically similar. These methods can automatically segment thyroid glands in 2D or 3D thyroid ultrasound images except DT method. This was because that user interactions were added to obtain better segmentation results in the DT model implementation. Machine learning based methods were divided into traditional learning algorithms (e.g. RBFNN, SVM, ELM, etc.) and deep learning algorithms (e.g. 3D U-Net CNN). Traditional learning algorithms all required feature engineering, whose classification capability depended on data, features and algorithms. RBFNN, SVM, ELM, FNN, RF and DT methods required image preprocessing to enhance the images or image postprocessing to improve the segmentation accuracy. When considering the computational complexity of these models, DT model was relatively the simplest model with low computational complexity. 3D U-Net method had the capability to learn the appropriate features, but its classification capability still depended on data and algorithm. All methods required marked data to train, especially 3D U-Net method required plenty of marked training data. But due to the difficulty of obtaining medical image data and the high cost of image data annotation, the obtained image datasets in these reviewed machine and deep learning based methods were small. This limited the performance of these machine and deep learning algorithms. Furthermore, 3D U-Net method took a long time to train the neural network model.

5.2. Thyroid nodule segmentation methods

Thyroid nodules can be classified as hyperechoic nodules, hypoechoic nodules and isoechoic nodules in thyroid ultrasound images. Malignant thyroid nodules usually appear as hypoechoic nodules with irregular boundaries, whereas the malignancy risk of isoechoic nodules is considerable as well [85]. Among the reviewed thyroid nodule segmentation methods, some methods had good segmentation effect on hyperechoic and hypoechoic nodules, while their segmentation capability of isoechoic nodules was insufficient. As shown in Table 2, ACWE, OE-ACWE, VBAC, GA-VBAC, VRS and GC-ACWE methods cannot segment isoechoic nodules. In addition, as seen from Table 2, all reviewed thyroid nodule segmentation methods were applied in 2D segmentation scenario.

ACWE model was one of the earliest methods used in thyroid nodule segmentation. ACWE model used image pixel intensity information to guide contour evolution, which involved integral operators that provided an inherent noise filter mechanism. Therefore, ACWE model was not sensitive to noise. The image noise reduction can be avoided in the segmentation methods adopting ACWE based model. ACWE model assumed that the target and the background regions were homogeneous respectively. But this hypothesis was violated in the thyroid ultrasound images due to the heterogeneity of thyroid tissue texture and the presence of speckle noise [86]. In addition, manual initial contour was required in ACWE model. Furthermore, most of the thyroid nodule segmentation methods based on active contour model were device dependent, which meant that different parameters were required to construct the active contour models to handle the ultrasound images acquired from different devices or the same devices using different settings. As a result, ACWE, VBAC, JET, DRLSE, SNDRLS, KNN-JET, and GC-ACWE methods required manual parameter tuning. But the OE-ACWE model can set parameters automatically according to the orientation consistency of the structure tensors.

VBAC model was an improvement based on ACWE model and it inherited some of the characteristics of ACWE model. Therefore, VBAC model was insensitive to noise and required manual initial contour. The improvement of VBAC model was that it did not follow the assumption that the target and the background areas were homogeneous respectively. The VBAC model utilized the information from sparse background regions to solve the problem of intensity inhomogeneity in thyroid ultrasound images [85]. VBAC model was device dependent and sensitive to model parameters. In addition, the improvement of GA-VBAC model was to automatically tune model parameters to cope with device dependent. Moreover, GC-ACWE model was proposed combining the characteristics of GC and ACWE models. It reduced nodes and edges in the graph structure, improved computing speed, and prevented boundary leakage.

The major defect of ACWE, OE-ACWE, VBAC, GA-VBAC and GC-ACWE models was the insufficient capability to segment isoechoic thyroid nodules. Both JET model and ACWE model followed Mumford-Shah segmentation method, but the difference was that JET model incorporated the texture information into active contour model. Combining with the texture characteristics represented by LBP distribution, JET model had the capability to segment isoechoic thyroid nodules. The limitation of JET was that it required initial ROI to initialize contour. Thus, KNN-JET model was proposed by adding a KNN classifier before JET model to automatically generate initial contour. Therefore, the method based on KNN-JET model can automatically segment thyroid nodules without manual participation in the segmentation process.

The energy function of DRLSE method was based on edge characteristics, which utilized an improved speed stopping constraint to stop the active contour at the edge. DRLSE model was able to detect the boundaries of nodules regardless of the echogenicity of nodules. Thus it had the capability to segment isoechoic thyroid nodules in thyroid ultrasound images. The improved speed stopping constraint of DRLSE model can also prevent boundary leakage in the process of contour evolution. But DRLSE model was poor in segmentation of weak edges and sensitive to the locations of manually initialized contours. These defects made it difficult for DRLSE model to automatically segment thyroid nodule ultrasound images. SNDRLS model can make up for the deficiency of DRLSE model, which integrated the advantages of SNLM clustering and level set, including delineation of multiple nodules and strong resistance to noise. But SNLM clustering required a large amount of calculations, which made SNDRLS method spent a relatively longer time on segmenting thyroid nodules.

HMM method was a segmentation method based on edges and shapes. Firstly, the edge map of the image was obtained, and then the location of the nodule was found by detecting the shape of the edge. HMM method was effective only if the image edges were clear and completely extracted, and the shapes of nodules were elliptic or circular with regular boundaries. The advantage of this method was that it was insensitive to echogenicity of nodules and thyroid nodules were automatically segmented.

NC method was one of the region-based segmentation methods, whose advantage was that it was not easy to generate small region segmentation and suitable for medical image segmentation. But

Table 2
Analysis of thyroid nodule segmentation methods.

	Method	Advantages	Disadvantages	Hyperechoic	Hypoechoic	Isoechoic	Dimension
Contour and Shape Based Methods	ACWE	Insensitive to noise	Require manual initial contour, manual parameter tuning	✓	✓	–	2D
	OE-ACWE	Automatic parameter setting, fast contour evolution speed, insensitive to noise	Require manual initial contour, assume respective homogeneity of target and background regions	✓	✓	–	2D
	VBAC	Insensitive to noise, solve intensity inhomogeneity	Require manual initial contour, manual parameter tuning	✓	✓	–	2D
	GA-VBAC	Insensitive to noise, solve intensity inhomogeneity, automatic parameter setting	Require manual initial contour	✓	✓	–	2D
	JET	Insensitive to nodule echogenicity	Require manual initial contour, manual parameter tuning	✓	✓	✓	2D
	DRLSE	Insensitive to nodule echogenicity, insensitive to noise, prevent boundary leakage	Sensitive to manual initial contour, manual parameter tuning	✓	✓	✓	2D
	SNDRLS	Insensitive to nodule echogenicity, strong resistance to noise, automatic segmentation, multiple nodule delineation	Large amount of calculations, manual parameter tuning	✓	✓	✓	2D
	HMM	Insensitive to nodule echogenicity, automatic segmentation	Require prior shape information, sensitive to quality of ultrasound images	✓	✓	✓	2D
Region Based Methods	NC	Less computation of weight matrix	Coarse segmentation results	✓	✓	–	2D
	VRS	Easy to implement, fast segmentation speed	Sensitive to manual initial ROI, sensitive to noise and blur nodule boundaries	✓	✓	–	2D
Machine and Deep Learning Methods	DT	Insensitive to nodule echogenicity, automatic segmentation	Require marked training data, require image preprocessing, sensitive to selected features	✓	✓	✓	2D
	SVM	Insensitive to nodule echogenicity, automatic segmentation	Require marked training data, require image processing, sensitive to selected features	✓	✓	✓	2D
	KNN	Insensitive to nodule echogenicity, automatic segmentation	Require marked training data, require image processing, sensitive to selected features	✓	✓	✓	2D
	CNN	Insensitive to nodule echogenicity, automatic segmentation, able to learn without handcrafted features	Require plenty of marked training data, long training time	✓	✓	✓	2D
	CCNN	More accurate segmentation results, insensitive to nodule echogenicity	Manually make artificial marks, require plenty of marked training data, long training time	✓	✓	✓	2D
Hybrid Methods	GC-ACWE	Insensitive to noise, prevent boundary leakage	Manual parameter tuning	✓	✓	–	2D
	KNN-JET	Insensitive to nodule echogenicity, automatic segmentation	Manual parameter tuning	✓	✓	✓	2D

the NC model needed large computer memory for the calculation of weight matrix, and over segmentation and under segmentation results were often generated by the model. Fractional differentiation can be introduced into the NC model to reduce the computational resource consumption on the weight matrix calculation.

VRS method was another region-based segmentation method. Its main advantage was that it was easy to implement and the segmentation speed of an thyroid ultrasound image was fast. VRS method utilized the variance difference of pixel intensity between the nodule region and the background region in an initial ROI to obtain boundary points. Therefore, VRS was very sensitive to manual initial ROI, and it was difficult to get the correct boundary points when the nodule boundary was blur or the nodule was isoechoic.

The advantages and disadvantages of thyroid nodule segmentation methods based on machine and deep learning algorithms were basically similar. These methods were insensitive to the echogenicity of nodules, and able to automatically segment thyroid nodules in thyroid ultrasound images. Similar to the machine learning based methods for thyroid gland segmentation which were analyzed in Section 5.1, the methods based on traditional machine learning algorithms (e.g. DT, SVM and KNN) for thyroid nodule segmentation also required feature engineering, image preprocessing and image postprocessing to improve the segmentation performance. The method based on CNN model for thyroid nodule segmentation also had the capability to learn appropriate features for CNN model as the method based on 3D U-Net model for thyroid gland segmentation did. CCNN was a phased model, which firstly extracted ROIs containing thyroid nodules, and then further segmented the thyroid nodules in ROIs. The advantage was that better segmentation results can be obtained through the phased approach. However, CCNN model was not automatic, which required manual marking in ROIs to segment nodules. All learning methods required marked training data, but the obtained image datasets in these methods were relatively small. Moreover, CNN method took a long time to train the neural network model.

6. Quantitative performance evaluation

The segmentation performance of thyroid gland segmentation methods and thyroid nodule segmentation methods was evaluated by quantitative analysis of evaluation metrics. Through the analysis of literatures, it was found that although some evaluation metrics were named differently, their calculation methods were the

same. For example, Sensitivity [1], Nodule Area [66], True Positive Ratio [60], and True Positive Fraction [68] were calculated in the same way. Therefore, we named these metrics with a uniform name which was Sensitivity, for the convenience of analyses and comparisons. The reviewed thyroid gland segmentation methods were applied to 2D and 3D thyroid ultrasound images, while the reviewed thyroid nodule segmentation methods were applied to 2D thyroid ultrasound images. Thus, the evaluation metrics were categorized into 2D metrics and 3D metrics. The 3D evaluation metrics were only used for 3D thyroid gland segmentation. These evaluation metrics will be described in detail in the following.

6.1. 2D metrics

2D metrics were used to evaluate the performance of 2D thyroid gland segmentation and 2D thyroid nodule segmentation, which contained area based and contour based metrics. The area based metrics included Sensitivity, Specificity, False Positive Rate (FPR), False Negative Rate (FNR), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Accuracy, Mean Overlap (MO), and Dice Coefficient (DC). Contour based metrics included Concavity, Mean Absolute Distance (MAD), and Hausdorff Distance (HD).

Area based metrics were achieved by comparing the similarity between segmented area and ground truth. The segmented area referred to thyroid gland area or thyroid nodule area segmented by a certain segmentation method, and ground truth referred to thyroid gland area or thyroid nodule area delineated by experts. Fig. 4 demonstrates the fundamental statistical definitions used for area based evaluation metrics. As shown in Fig. 4, True Positive (TP) denotes the number of pixels of thyroid gland or nodule region which are exactly segmented as pixels of thyroid gland or nodule region, True Negative (TN) denotes the number of pixels of non-thyroid gland or nodule region which are exactly segmented as pixels of non-thyroid gland or nodule region, False Positive (FP) denotes the number of pixels of non-thyroid gland or nodule region which are falsely segmented as pixels of thyroid gland or nodule region, False Negative (FN) denotes the number of pixels of thyroid gland or nodule region which are falsely segmented as pixels of non-thyroid gland or nodule region. Based on the fundamental statistical definitions, Sensitivity, Specificity, FNR, FPR, NPV, PPV, Accuracy, MO, and DC metrics are defined as follows:

$$Sensitivity = \frac{TP}{TP+FN}, \tag{2}$$

$$Specificity = \frac{TN}{TN+FP}, \tag{3}$$

$$FPR = \frac{FP}{TN+FP}, \tag{4}$$

$$FNR = \frac{FN}{TP+FN}, \tag{5}$$

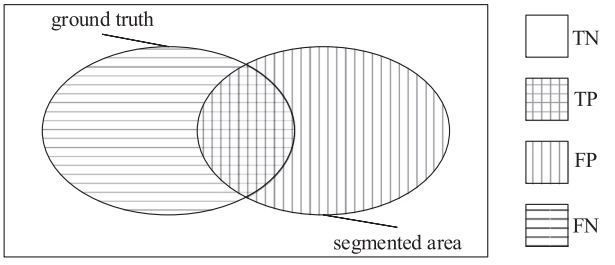

Fig. 4. Fundamental statistical definitions used for area based evaluation metrics.

$$PPV = \frac{TP}{TP+FP}, \tag{6}$$

$$NPV = \frac{TN}{TN+FN}, \tag{7}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}, \tag{8}$$

$$MO = \frac{TP}{FN+TP+FP}, \tag{9}$$

$$DC = \frac{2TP}{FN+2TP+FP}. \tag{10}$$

The Concavity is a contour feature indicating the existence of concave regions. The MAD metric describes the shape difference between two contours [87]. The HD metric is the maximum of the shortest distances between the segmented and the ground truth boundaries. The point set of the ground truth boundary is $A = \{a_1, a_2, ..., a_M\}$, and the point set of the segmented boundary is $B = \{b_1, b_2, ..., b_N\}$, where M is the number of boundary points of ground truth boundary, and N is the number of boundary points of the segmented boundary. Concavity, MAD, and HD metrics are defined as follows:

$$Concavity = \frac{\frac{1}{N}\sum_{i=1}^{N}\sqrt{(X_{cen}-X_{b_i})^2+(Y_{cen}-Y_{b_i})^2}}{\frac{1}{M}\sum_{j=1}^{M}\sqrt{(X_{cen}-X_{a_j})^2+(Y_{cen}-Y_{a_j})^2}}, \tag{11}$$

$$MAD = \frac{1}{2}\{\frac{1}{M}\sum_{j=1}^{M} d(A,B) + \frac{1}{N}\sum_{i=1}^{N} d(B,A)\}, \tag{12}$$

$$HD = max(d(A,B), d(B,A)), \tag{13}$$

where X_{cen} and Y_{cen} are x-coordinate and y-coordinate of the center point of the segmented area respectively, X_{a_j} and Y_{a_j} are x-coordinate and y-coordinate of the jth point within A, and X_{b_i} and Y_{b_i} are x-coordinate and y-coordinate of the ith point within B. Furthermore, $d(A,B) = max_{a\in A} min_{b\in B}||a-b||$, and $d(B,A) = max_{b\in B} min_{a\in A}||b-a||$.

6.2. 3D metrics

Sensitivity and Specificity metrics could also be used as 3D evaluation metrics, when the segmented area and the ground truth were considered as the regions in 3D space. In addition, there was a volume-based evaluation metric for 3D evaluation, which was Mean Square Error (MSE) of the volume. In a 3D thyroid ultrasound image I_k, let V_{sk} be the volume of the segmented area and V_{gk} be the volume of the ground truth. The MSE metric is define as:

$$MSE = \frac{1}{K}\sum_{k=1}^{K}(|V_{sk}-V_{gk}|)^2, \tag{14}$$

where K is the number of the test images.

The overall quantitative performance analyses of the reviewed methods are presented inTables 3–5. The performance measurements shown in Tables 3 and 4 are for 2D thyroid gland segmentation and 3D thyroid gland segmentation, respectively. And the performance of 2D thyroid nodule segmentation methods is presented in Table 5. The methods without 2D or 3D quantitative evaluation results are not included in the corresponding tables. In these tables, the Automatic column indicates whether the corresponding method is fully automated without user interactions in the segmentation process. The Test Dataset column specifies the number of 2D or 3D thyroid ultrasound images used to test the developed methods. The Evaluation Metric column presents the metrics used to evaluate the reviewed methods, and

Table 3
Performance measurements of thyroid gland segmentation methods for 2D thyroid ultrasound images.

	Reference	Year	Method	Automatic	Test Dataset	Evaluation Metric	Value
Contour and Shape Based Methods	[25]	2018	ACWE	No	1416 images	DC	80.00%
						HD	8.10 mm
Region Based Methods	[34]	2017	SRR	Yes	52 images	Sensitivity	85.87%(GT-expert 1)
							92.03%(GT-expert 2)
						Specificity	96.27%(GT-expert 1)
							94.80%(GT-expert 2)
						PPV	84.10%(GT-expert 1)
							77.13%(GT-expert 2)
						DC	84.47%(GT-expert 1)
							83.23%(GT-expert 2)
	[25]	2018	GC	No	1416 images	DC	76.50%
						HD	8.30 mm
Machine and Deep Learning Methods	[1]	2010	RBFNN	Yes	5 images	Sensitivity	91.58%
						Specificity	97.61%
						PPV	89.14%
						NPV	98.04%
						Accuracy	96.52%
	[45]	2011	SVM	Yes	3 images	Accuracy	84.78%
	[45]	2011	ELM	Yes	3 images	Accuracy	93.56%
	[46]	2013	FNN	Yes	5 images	Sensitivity	89.06%
						Specificity	98.90%
						FPR	1.09%
						FNR	10.93%
						Accuracy	96.51%
	[25]	2018	3D U-Net	Yes	1416 images	DC	87.60%
						HD	8.10 mm
	[25]	2018	RF	Yes	1416 images	DC	86.20%
						HD	7.50 mm
	[25]	2018	DT	No	1416 images	DC	67.00%
						HD	9.50 mm

Table 4
Performance measurements of thyroid gland segmentation methods for 3D thyroid ultrasound images.

	Reference	Year	Method	Automatic	Test Dataset	Evaluation Metric	Value
Contour and Shape Based Methods	[30]	2008	GAC	Yes	10 images	Sensitivity	75.00%
						Specificity	97.00%
Machine and Deep Learning Methods	[1]	2010	RBFNN	Yes	5 images	MSE	0.582

the Value column shows the corresponding values of the evaluated metrics. The larger the value of the evaluation metrics such as Sensitivity, Specificity, PPV, NPV, Accuracy, MO, DC and Concavity, the better the performance. But the smaller the value of FNR, FPR, MAD, HD and MSE, the better the performance. In the Value column of Table 5, there are some additional notations, where GT-expert 1 and GT-expert 2 mean that the test images were manually sketched by two experts to obtain two ground truth datasets. The metric value with GT-expert 1 notation was obtained by comparing the segmentation results of the corresponding method and the ground truth dataset by expert 1. The metric value with GT-expert 2 notation was obtained in a similar manner using the ground truth dataset by expert 2. The additional notation Manual Delineation means that the segmentation results were obtained by manual delineation instead of generated by the corresponding method. The metric value with Manual Delineation notation was obtained by comparing the manual delineation results with the ground truth, and the metric value without Manual Delineation notation was obtained by comparing the segmentation results of corresponding method with the ground truth.

7. Discussion

Thyroid ultrasound image segmentation has been developed for years, during which new methods have been proposed constantly. Therefore, there are various options for thyroid gland segmentation and thyroid nodule segmentation methods. In order to select the most appropriate segmentation technology for a particular segmentation application, the application scenario, image features and segmentation performance should be considered together and comprehensively. Furthermore, user interactions and additional information (e.g. manual initial contours or ROIs, prior shapes) should also be considered. Combining the advantages and disadvantages of the reviewed segmentation methods listed in Tables 1 and 2, and the automatic segmentation capability and the segmentation performance for 2D and 3D segmentation scenarios illustrated in Tables 3–5, the selection of appropriate segmentation methods for specific applications are discussed in the following.

In thyroid gland segmentation methods, machine and deep learning based methods are prior choices. The strength of machine and deep learning based methods is that they can distinguish complex patterns. More comprehensive image characteristics, such as texture, shape, boundary, position, etc., can be obtained through feature engineering or feature learning. Compared with segmentation based on single image characteristic, segmentation based on comprehensive image characteristics has higher accuracy. And traditional machine learning based methods (e.g. RBFNN, SVM, etc.) obtain the complete segmentation image by classifying image blocks of certain size using a sliding window, this results in redundant computation and misclassification, as well as jagged segmentation boundaries. On the other hand, deep learning based method (e.g. 3D U-Net CNN) performs faster and gains better segmentation performance than traditional machine learning based methods. Deep learning based method can obtain more effective image features through feature learning, hence it has better discriminating capability in pixel classification. However, plenty of

Table 5
Performance measurements of thyroid nodule segmentation methods for 2D thyroid ultrasound images.

	Reference	Year	Method	Automatic	Test Dataset	Evaluation Metric	Value
Contour and Shape Based Methods	[49]	2005	ACWE	No	35 images	MO	82.60%
	[49]	2005	VBAC	No	35 images	MO	88.80%
	[66]	2006	HMM	Yes	40 images	Sensitivity	88.83% (GT-expert 1)
							87.58% (GT-expert 2)
						Concavity	89.21% (GT-expert 1)
							89.08% (GT-expert 2)
						MAD	2.54 pixels (GT-expert 1)
							2.16 pixels (GT-expert 2)
	[20]	2007	ACWE	No	71 images	MO	84.80%
	[20]	2007	VBAC	No	71 images	MO	91.10%
	[50]	2007	GA-VBAC	No	45 images	MO	92.50%
							91.80% (Manual Delineation)
	[7]	2009	JET	No	74 images	MO	92.21%
						DC	95.91%
						MAD	1.30 pixels
	[57,58]	2014	OE-ACWE	No	20 images	MO	83.70%
	[60]	2015	DRLSE	No	20 images	Sensitivity	85.00%
						MO	79.00%
	[62]	2016	SNDRLS	Yes	42 images	Sensitivity	95.45%
						MO	93.15%
						DC	94.25%
						MAD	1.80 pixels
						HD	0.70 pixels
	[55]	2017	ACWE	No	52 images	PPV	96.00%
						DC	89.60%
	[63]	2018	SNDRLS	Yes	138 images	Sensitivity	94.69%
						MO	91.99%
						DC	93.84%
						HD	0.38 pixels
Region Based Methods	[68]	2016	VRS	No	26 images	Sensitivity	82.49%
						Specificity	99.90%
						FPR	0.10%
						FNR	17.51%
						Accuracy	99.63%
Machine and Deep Learning Methods	[69]	2009	DT	Yes	6 images	Sensitivity	89.60%
						Specificity	98.50%
						PPV	89.10%
						NPV	98.30%
						Accuracy	97.50%
	[2]	2010	SVM	Yes	118 images	Accuracy	91.30%
	[79]	2017	CNN	Yes	22123 images	Sensitivity	91.50%
						MO	86.83%
						DC	92.24%
						HD	0.62 pixels
	[80]	2018	CCNN	No	200 images	MO	87.00%
Hybrid Methods	[85]	2011	KNN-JET	Yes	142 images	MO	92.30%
						MAD	1.40 pixels

marked training data are demanded in these machine learning based methods. The methods based on active contour or region generally face the problems of large amount of computation and insufficient segmentation capability in large scale thyroid ultrasound images. For example, ACWE method needs a lot of time for iterations because of the large image size and large initial contour size, and user interactions are required to stop the iterations or reinitialize the initial contour. Moreover, GC method requires more user interactions as compared to ACWE method. Another region based method, SRR method, is capable of segmenting transverse and longitudinal scans of thyroid ultrasound images, but the model design based on similar reflective regions and positional relationships between tissues limits its ability to segment thyroid ultrasound images with other scanning angles.

Active contour based methods are good choices for thyroid nodule segmentation. Active contour models are insensitive to noise, which effectively reduces the influence of speckle noise in ultrasound images. However, the echogenicity of nodules should be taken into account when selecting specific active contour based methods for thyroid nodule segmentation. Most of the active contour models with energy functions based on image region characteristics (e.g. ACWE, VBAC, etc.) are incapable to segment isoechoic nodules, but the active contour models with energy functions based on image edge characteristics (e.g. DRLSE) are able to segment hyperechoic, hypoechoic, and isoechoic nodules. The main limitation of active contour models is that they require manual initial contours, and the locations of the initial contours are important for the segmentation. Manually initializing contours limits the automation of the methods and has subjective influence on the segmentation results. Thus, automatic initial contour extraction is a good solution. For example, SNDRLS and KNN-JET methods utilize proper algorithms to obtain initial contours before applying active contour models to thyroid nodule segmentation. Another limitation of active contour models is that most models are sensitive to the parameters in energy functions and require manual parameter tuning. GA algorithms can be used to automatically tune the parameters so as to enhance the robustness of the model, such as GA-VBAC. The shape based method (e.g. HMM) requires prior shape information, which limits its practicality. Machine learning based methods are preferred for automatic thyroid nodule segmentation. Traditional machine learning based methods (e.g. DT, SVM, etc.) require image preprocessing and feature en-

gineering, but deep learning based method can learn appropriate features through data driven algorithms. The main limitation of machine and deep learning based methods is the small amount of available marked training datasets. If a large amount of marked training data is obtained, the machine learning based methods, especially deep learning based methods, can probably highly extend their segmentation performance, which will make them the preferable choices for thyroid nodule segmentation.

8. Conclusion and perspective

Thyroid image segmentation is an indispensable procedure in medical image diagnoses and computer-aided diagnosis systems. Thyroid gland segmentation and thyroid nodule segmentation in ultrasound images promote the study of thyroid disease diagnoses and provide valuable information for clinicians to make the best diagnostic decisions. This work categorizes and reviews the existing thyroid gland segmentation and thyroid nodule segmentation methods in ultrasound images. These methods are categorized into four groups according to the theoretical bases utilized to solve the problem. Twenty-eight representative papers are selected for comprehensive analyses and comparisons in this review. The similarities and differences of these methods are analyzed, and the strengths and limitations of the methods are discussed along with quantitative performance evaluations using uniformly named metrics. The dominant thyroid gland segmentation methods are machine and deep learning methods. The training of massive data makes these models have better segmentation performance and robustness. But deep learning models usually require plenty of marked training data and long training time. For thyroid nodule segmentation, the most common methods are contour and shape based methods, which have good segmentation performance. However, most of them are tested on small datasets. Accordingly, the appropriate segmentation method could be selected under the comprehensive consideration of application scenario, image features, method practicability and segmentation performance.

In general, there are many methods for thyroid ultrasound image segmentation, but there is no one-size-fits-all model. Hence, it is important to choose the suitable approach for a particular scenario and condition. There are many application scenarios and conditions, such as sample size, time cost, accuracy, automation, model size, etc. For example, in the case with a small amount of samples, the contour based segmentation method is preferred. But the active contour based models need manual parameter tuning according to personal experience. Accordingly, GA algorithm or OE algorithm can be used to set parameters automatically, and the drawback of requiring manual annotation can be solved by other methods such as GC and TBD algorithms. Machine and deep learning methods are the best choices for automatic and accurate segmentation if there are a lot of samples. But data augmentations can be done to generate more training data. Active contour models usually need a long time to evolve the contours, while small deep learning model can achieve real-time segmentation with the assistance of parallel computing techniques. The strong fitting ability of the deep learning models enables the models to achieve the best results in thyroid ultrasound image segmentation. Large medical platforms and large amount of medical data promote the development of intelligent medicine, which make the machine and deep learning methods become the mainstream in thyroid ultrasound image segmentation, but the combination of deep learning methods and conventional methods such as contour based methods are emerging.

There are some limitations in current thyroid ultrasound image segmentation studies. Firstly, current thyroid gland segmentation studies focus on the segmentation of normal thyroid glands, and there is no published research on the segmentation of pathological or abnormal thyroid glands. The locations, sizes and textures of pathological or abnormal tissues may bring new challenges to the segmentation of thyroid glands. Secondly, the purpose of current thyroid nodule segmentation studies is to analyze the benign and malignant nodules, not to identify the specific nodular diseases. Moreover, there is a lack of standard thyroid ultrasound image datasets and evaluation metrics, which also limits the research on thyroid disease diagnoses. Future studies may develop to fill in the mentioned blanks in current thyroid ultrasound image segmentation research.

Declaration of Competing Interest

The authors of this manuscript are employed by South China University of Technology and The Third Affiliated Hospital of Sun Yat-Sen University, respectively.

Acknowledgments

This work was supported in part by “National Natural Science Foundation of China” (no. 61802130), “Guangdong Natural Science Foundation” (no. 2018A030310355, 2019A1515012152), “Guangzhou Science and Technology Program” (no. 201707010223), and “the Fundamental Research Funds for the Central Universities (no. D2182350).

References

[1] C.-Y. Chang, Y.-F. Lei, C.-H. Tseng, S.-R. Shih, Thyroid segmentation and volume estimation in ultrasound images, IEEE Trans. Biomed. Eng. 57 (6) (2010) 1348–1357.
[2] E.G. Keramidas, D. Maroulis, D.K. Iakovidis, TND: a thyroid nodule detection system for analysis of ultrasound images and videos, J. Med. Syst. 36 (3) (2012) 1271–1281.
[3] A. Gursoy, F. Erdogan, Ultrasonographic approach to thyroid nodules: State of art, Thyroid Int. (2012) 1–16.
[4] M.J. Welker, D. Orlov, Thyroid nodules, Am. Fam. Physician 67 (3) (2003) 559–566.
[5] K. Ain, M.S. Rosenthal, The Complete Thyroid Book, Second Ed., McGraw-Hill Education, 2010.
[6] A.G. Unnikrishnan, U.V. Menon, Thyroid disorders in India: an epidemiological perspective, Ind. J. Endocrinol. Metab. 15 (2011) 78–81.
[7] M.A. Savelonas, D.K. Iakovidis, I. Legakis, D. Maroulis, Active contours guided by echogenicity and texture for delineation of thyroid nodules in ultrasound images, IEEE Trans. Inform. Technol. Biomed. 13 (4) (2009) 519–527.
[8] S. Wu, Q. Zhu, Y. Xie, Evaluation of various speckle reduction filters on medical ultrasound images, in: Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2013, pp. 1148–1151.
[9] F.N. Tessler, W.D. Middleton, E.G. Grant, J.K. Hoang, L.L. Berland, S.A. Teefey, J.J. Cronan, M.D. Beland, T.S. Desser, M.C. Frates, L.W. Hammers, U.M. Hamper, J.E. Langer, C.C. Reading, L.M. Scoutt, A.T. Stavros, ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee, J. Am. Coll. Radiol. 14 (5) (2017) 587–595.
[10] Y.H. Lee, D.W. Kim, H.S. In, J.S. Park, S.H. Kim, J.W. Eom, B. Kim, E.J. Lee, M.H. Rho, Differentiation between benign and malignant solid thyroid nodules using an us classification system, Korean J. Radiol. 12 (5) (2011) 559–567.
[11] W.-J. Moon, S.L. Jung, J.H. Lee, D.G. Na, J.-H. Baek, Y. Lee, J. Kim, J.S. Byun, D.H. Lee, Benign and malignant thyroid nodules: US differentiation multicenter retrospective study, Radiology 247 (3) (2008) 762–770.
[12] G.E. Mailloux, M. Bertrand, R. Stampfler, S. Ethier, Texture analysis of ultrasound B-mode images by segmentation, Ultrason. Imag. 6 (3) (1984) 262–277.
[13] S.V.B. Jardim, M.A.T. Figueiredo, Automatic contour estimation in fetal ultrasound images, in: Proceedings of International Conference on Image Processing, 2003, pp. 1065–1068.
[14] D.R. Chen, R.F. Chang, W.J. Kuo, M.C. Chen, Y.L. Huang, Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks, Ultrasound Med. Biol. 28 (10) (2002) 1301–1310.
[15] N. Hu, D.B. Downey, A. Fenster, H.M. Ladak, Prostate boundary segmentation from 3D ultrasound images, Med. Phys. 30 (7) (2003) 1648–1659.
[16] S. Petroudi, C. Loizou, M. Pantziaris, C. Pattichis, Segmentation of the common carotid intima-media complex in ultrasound images using active contours, IEEE Trans. Biomed. Eng. 59 (11) (2012) 3060–3069.
[17] X. Zang, R. Bascom, C. Gilbert, J. Toth, W. Higgins, Methods for 2D and 3D endobronchial ultrasound image segmentation, IEEE Trans. Biomed. Eng. 63 (7) (2016) 1426–1439.
[18] L. Yu, Y. Guo, Y. Wang, J. Yu, P. Chen, Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks, IEEE Trans. Biomed. Eng. 64 (8) (2017) 1886–1895.

[19] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vis. 1 (4) (1988) 321331.
[20] D.E. Maroulis, M.A. Savelonas, D.K. Iakovidis, S.A. Karkanis, N. Dimitropoulos, Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images, IEEE Trans. Inform. Technol. Biomed. 11 (5) (2007) 537–543.
[21] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. 10 (2) (2001) 266–277.
[22] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, Commun. Pure Appl. Math. 42 (1989) 577–685.
[23] C.M. Chen, H.H.S. Lu, A.T. Hsiao, A dual-snake model of high penetrability for ultrasound image boundary extraction, Ultrasound Med. Biol. 27 (2) (2001) 1651–1665.
[24] P. Poudel, A. Illanes, C. Hansen, M. Friebe, Active contours extension and similarity indicators for improved 3D segmentation of thyroid ultrasound images, in: Proceedings of SPIE, Medical Imaging: Imaging Informatics for Healthcare, Research, and Applications, 2017, pp. 1–8.
[25] P. Poudel, A. Illanes, C. Hansen, M. Friebe, Evaluation of commonly used algorithms for thyroid ultrasound images segmentation and improvement using machine learning approaches, J. Healthcare Eng. 2018 (2018) 1–13.
[26] ImFusion Suite, (2020), https://www.imfusion.com/products/imfusion-suite, Accessed 8 January 2020.
[27] J. Kaur, A. Jindal, Comparison of thyroid segmentation algorithms in ultrasound and scintigraphy images, Int. J. Comput. Appl. 50 (23) (2012) 24–27.
[28] S. Lankton, A. Tannenbaum, Localizing region based active contours, IEEE Trans. Image Process. 17 (11) (2008) 2029–2039.
[29] C. Li, C. Xu, C. Gui, M.D. Fox, Distance regularized level set evolution and its application to image segmentation, IEEE Trans. Image Process. 19 (12) (2010) 3243–3254.
[30] E.N.K. Kollorz, D.A. Hahn, R. Linke, T.W. Goecke, J. Hornegger, T. Kuwert, Quantification of thyroid volume using 3D ultrasound imaging, IEEE Trans. Med. Imag. 27 (4) (2008) 457–466.
[31] J. Weickert, Applications of nonlinear diffusion in image processing and computer vision, Acta Mathematica Universitatis Comenianae 70 (1) (2001) 3350.
[32] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, in: Proceedings of IEEE International Conference on Computer Vision, 1995, pp. 694–699.
[33] C. Rother, V. Kolmogorov, A. Blake, GrabCut: interactive foreground extraction using iterated graph cuts, ACM Trans. Graph. 23 (3) (2004) 309–314.
[34] N.S. Narayan, P. Marziliano, J. Kanagalingam, C.G.L. Hobbs, Speckle patch similarity for echogenicity-based multiorgan segmentation in ultrasound images of the thyroid gland, IEEE J. Biomed. Health Inform. 21 (1) (2017) 172–183.
[35] B.B. Tempkin, Ultrasound Scanning: Principles and Protocols, Third Edition, Elsevier Health Sciences, 2009.
[36] G.B. Thomas, R.L. Finney, Calculus and Analytic Geometry, 9th Ed., Addison Wesley, 1995.
[37] B. Rahmatullah, A.T. Papageorghiou, J.A. Noble, Integration of local and global features for anatomical object detection in ultrasound, in: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention, 2012, pp. 402–409.
[38] M.T. Rackham, S. Rueda, C.L. Knight, J.A. Noble, Ultrasound image segmentation using feature asymmetry and shape guided live wire, in: Proceedings of SPIE, Medical Imaging: Image Processing, 2013, pp. 1–9.
[39] D. Jayachandra, A. Makur, Directional variance: a measure to find the directionality in a given image segment, in: Proceedings of IEEE International Symposium on Circuits and Systems, 2010, pp. 1551–1554.
[40] T. Loupas, W.N. McDicken, P.L. Allan, An adaptive weighted median filter for speckle suppression in medical ultrasonic images, IEEE Trans. Circuit Syst. 36 (1) (1989) 129–135.
[41] Y.D. Chun, S.Y. Seo, N.C. Kim, Image retrieval using BDIP and BVLC moments, IEEE Trans. Circuit. Syst. Video Technol. 13 (9) (2003) 951–957.
[42] E.-L. Chen, P.-C. Chung, C.-L. Chen, H.-M. Tsai, C.-I. Chang, An automatic diagnostic system for CT liver image classification, IEEE Trans. Biomed. Eng. 45 (6) (1998) 783–794.
[43] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back--propagating errors, Nature 323 (1986) 533536.
[44] R. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: Proceedings of International Symposium on Micro Machine and Human Science, 1995, pp. 39–43.
[45] D. Selvathi, V.S. Sharnitha, Thyroid classification and segmentation in ultrasound images using machine learning algorithms, in: Proceedings of International Conference on Signal Processing, Communication, Computing and Networking Technologies, 2011, pp. 836–841.
[46] H. Garg, A. Jindal, Segmentation of thyroid gland in ultrasound image using neural network, in: Proceedings of International Conference on Computing, Communications and Networking Technologies, 2013, pp. 1–5.
[47] M.T. Hagan, M.B. Menhaj, Training feedforward networks with the marquardt algorithm, IEEE Trans. Neural Netw. 5 (6) (1994) 989–993.
[48] O. Cicek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention, 2016, pp. 424–432.
[49] D.E. Maroulis, M.A. Savelonas, S.A. Karkanis, D.K. Iakovidis, N. Dimitropoulos, Computer-aided thyroid nodule detection in ultrasound images, in: Proceedings of IEEE Symposium on Computer-Based Medical Systems, 2005, pp. 271–276.
[50] D.K. Iakovidis, M.A. Savelonas, S.A. Karkanis, S.A. Karkanis, D.E. Maroulis, A genetically optimized level set approach to segmentation of thyroid ultrasound images, Appl. Intell. 27 (3) (2007) 193–203.
[51] D.E. Goldberg, Genetic Algorithm in Search, Optimization, and Machine Learning, Addison-Wesley Professional, 1989.
[52] J.J. Grefenstette, Optimization of control parameters for genetic algorithms, IEEE Trans. Syst. Man Cybern. 16 (1) (1986) 122–128.
[53] H.A. Nugroho, A. Nugroho, L. Choridah, Thyroid nodule segmentation using active contour bilateral filtering on ultrasound images, in: Proceedings of International Conference on Quality in Research, 2015, pp. 43–46.
[54] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Proceedings of International Conference on Computer Vision, 2017, pp. 839–846.
[55] L. Gui, C. Li, X. Yang, Medical image segmentation based on level set and isoperimetric constraint, Physica Medica 42 (2017) 162–173.
[56] D. Tschumperle, R. Deriche, Vector-valued image regularization with PDEs : acommon framework for different applications, IEEE Trans. Pattern Analy. Mach. Intell. 27 (4) (2005) 506–517.
[57] E.A. Mylona, M.A. Savelonas, D. Maroulis, Automated adjustment of region-based active contour parameters using local image geometry, IEEE Trans. Cybern. 44 (12) (2014) 2757–2770.
[58] E.A. Mylona, M.A. Savelonas, D. Maroulis, Self-parameterized active contours based on regional edge structure for medical image segmentation, SpringerPlus 3 (424) (2014) 1–9.
[59] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Analy. Mach. Intell. 24 (2002) 971–987.
[60] W. Du, N. Sang, An effective method for ultrasound thyroid nodules segmentation, in: Proceedings of International Symposium on Bioelectronics and Bioinformatics, 2015, pp. 207–210.
[61] Y. Yu, S.T. Acton, Speckle reducing anisotropic diffusion, IEEE Trans. Image Process. 11 (11) (2002) 1260–1270.
[62] D. Koundal, S. Gupta, S. Singh, Automated delineation of thyroid nodules in ultrasound images using spatial neutrosophic clustering and level set, Appl. Soft Comput. 40 (2016) 86–97.
[63] D. Koundal, S. Gupta, S. Singh, Computer aided thyroid nodule detection system using medical ultrasound images, Biomed. Signal Process. Control 40 (2018) 117–130.
[64] E.G. Keramidas, D.K. Iakovidis, D.E. Maroulis, S.A. Karkanis, Efficient and effective ultrasound image analysis scheme for thyroid nodule detection, in: Proceedings of International Conference on Image Analysis and Recognition, 2007, pp. 1052–1060.
[65] F. Smarandache, A Unifying Field in Logics: Neutrosophic Logic, American Research Press, 2005.
[66] S. Tsantis, N. Dimitropoulos, D. Cavouras, G. Nikiforidis, A hybrid multi-scale model for thyroid nodule boundary detection on ultrasound images, Comput. Method. Program. Biomed. 84 (2) (2006) 86–98.
[67] J. Zhao, W. Zheng, L. Zhang, H. Tian, Segmentation of ultrasound images of thyroid nodule for assisting fine needle aspiration cytology, Health Inform. Sci. Syst. 1 (2013) 1–12.
[68] W.M.H. Alrubaidi, B. Peng, Y. Yang, Q. Chen, An interactive segmentation algorithm for thyroid nodules in ultrasound images, in: Proceedings of International Conference on Intelligent Computing, 2016, p. 107115.
[69] C.-Y. Chang, H.-C. Huang, S.-J. Chen, Thyroid nodule segmentation and component analysis in ultrasound images, in: Proceedings of APSIPA Annual Summit and Conference, 2009, pp. 910–917.
[70] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cybern. 3 (6) (1973) 610–621.
[71] C.-M. Wu, Y.-C. Chen, Statistical feature matrix for texture analysis, Graph. Model. Image Process. 54 (5) (1992) 407–419.
[72] M.M. Galloway, Texture analysis using gray level run lengths, Computer Graphics and Image Processing 4 (1975) 172–179.
[73] C. Sun, W.G. Wee, Neighboring gray level dependence matrix for texture classification, Comput. Vis. Graph. Image Process. 23 (3) (1983) 341–352.
[74] J.E. Wilhjelm, M.-L.M. Gronholdt, B. Wiebe, S.K. Jespersen, L.K. Hansen, H. Sillesen, Quantitative analysis of ultrasound B-mode images of carotid atherosclerotic plaque: correlation with visual classification and histological examination, IEEE Trans. Med. Imag. 17 (6) (1998) 910–922.
[75] D.K. Iakovidis, E.G. Keramidas, D.E. Maroulis, Fuzzy local binary patterns for ultrasound texture characterization, in: Proceedings of International Conference of Image Analysis and Recognition, 2008, pp. 750–759.
[76] C.V. Jawahar, A.K. Ray, Fuzzy statistics of digital images, IEEE Signal Process. Lett. 3 (8) (1996) 225–227.
[77] O. Chapelle, P. Haffner, V.N. Vapnik, Support vector machines for histogram-based image classification, IEEE Trans. Neural Netw. 10 (5) (1999) 1055–1064.
[78] D.E. Maroulis, D.K. Iakovidis, S.A. Karkains, D.A. Karras, CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames, Comput. Method. Program. Biomed. 70 (2) (2003) 151–166.
[79] J. Ma, F. Wu, T. Jiang, Q. Zhao, D. Kong, Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks, Int. J. Comput. Assist. Radiol. Surg. 12 (11) (2017) 18951910.
[80] X. Ying, Z. Yu, R. Yu, X. Li, M. Yu, M. Zhao, K. Liu, Thyroid nodule segmentation in ultrasound images based on cascaded convolutional neural network, in: Proceedings of In International Conference on Neural Information Processing, 2018, pp. 373–384.

[81] R. Yu, K. Liu, X. Wei, J. Zhu, X. Li, J. Wang, X. Ying, Z. Yu, Localization of thyroid nodules in ultrasonic images, in: Proceedings of International Conference on Wireless Algorithms, Systems, and Applications, 2018, pp. 635–646.
[82] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: Proceedings of International Conference on Learning Representations, 2015, pp. 1–14.
[83] J. Zhou, Thyroid tumor ultrasound image segmentation based on improved graph cut, in: Proceedings of International Conference on Intelligent Transportation, Big Data and Smart City, 2016, pp. 130–133.
[84] Y. Boykov, V. Kolmogorov, Computing geodesics and minimal surfaces via graph cuts, in: Proceedings of IEEE International Conference on Computer Vision, 2003, pp. 26–33.
[85] L. Legakis, M.A. Savelonas, D.E. Maroulis, D.K. Lakovidis, Computer-based nodule malignancy risk assessment in thyroid ultrasound images, Int. J. Comput. Applications 33 (1) (2011) 2935.
[86] L. Grady, Random walks for image segmentation, IEEE Trans. Pattern Analy. Mach. Intell. 28 (11) (2006) 1768–1783.
[87] V. Chalana, D.T. Linker, D.R. Haynor, Y. Kim, A multiple active contour model for cardiac boundary detection on echocardiographic sequences, IEEE Trans. Med. Imag. 15 (3) (1996) 290–298.