

RESEARCH

Open Access



# Segmentation of pulmonary nodules using adaptive local region energy with probability density function-based similarity distance and multi-features clustering

Bin Li<sup>1\*</sup>, QingLin Chen<sup>1</sup>, Guangming Peng<sup>2</sup>, Yuanxing Guo<sup>2</sup>, Kan Chen<sup>1</sup>, LianFang Tian<sup>1</sup>, Shanxing Ou<sup>2</sup> and Lifei Wang<sup>3</sup>

\*Correspondence:

binlee@scut.edu.cn

<sup>1</sup> School of Automation  
Science and Engineering,  
South China University  
of Technology,  
Guangzhou 510640,  
Guangdong, China  
Full list of author information  
is available at the end of the  
article

## Abstract

**Background:** Pulmonary nodules in computerized tomography (CT) images are potential manifestations of lung cancer. Segmentation of potential nodule objects is the first necessary and crucial step in computer-aided detection system of pulmonary nodules. The segmentation of various types of nodules, especially for ground-glass opacity (GGO) nodules and juxta-vascular nodules, present various challenges. The nodule with GGO characteristic possesses typical intensity inhomogeneity and weak edges, which is difficult to define the boundary; the juxta-vascular nodule is connected to a vessel, and they have very similar intensities. Traditional segmentation methods may result in the problems of boundary leakage and a small volume over-segmentation. This paper deals with the above mentioned problems.

**Methods:** A novel segmentation method for pulmonary nodules is proposed, which uses an adaptive local region energy model with probability density function (PDF)-based similarity distance and multi-features dynamic clustering refinement method. Our approach has several novel aspects: (1) in the proposed adaptive local region energy model, the local domain for local energy model is selected adaptively based on k-nearest-neighbour (KNN) estimate method, and measurable distances between probability density functions of multi-dimension features with high class separability are used to build the cost function. (2) A multi-features dynamic clustering method is used for the segmentation refinement of juxta-vascular nodules, which is based on the nodule segmentation using active contour model (ACM) with adaptive local region energy and vessel segmentation using flow direction feature (FDF)-based region growing method. (3) it handles various types of nodules under a united framework.

**Results:** The proposed method has been validated on a clinical dataset of 113 chest CT scans that contain 157 nodules determined by a ground truth reading process, and evaluating the algorithm on the provided data leads to an average Tanimoto/Jaccard error of 0.17, 0.20 and 0.24 for GGO, juxta-vascular and GGO juxta-vascular nodules, respectively.

**Conclusions:** Experimental results show desirable performances of the proposed method. The proposed segmentation method outperforms the traditional methods.

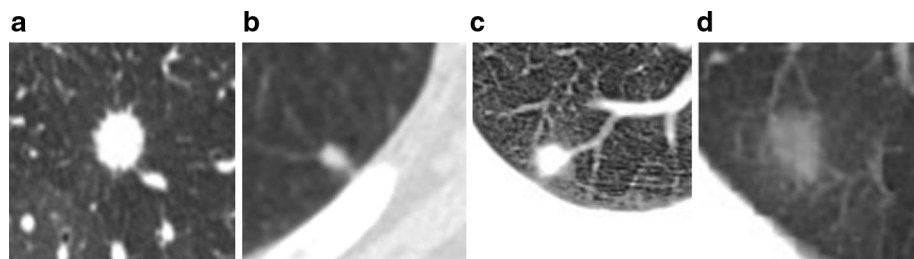
**Keywords:** Image segmentation, Pulmonary nodules, Active contour model, Feature extraction, Computer tomography image

## Background

Pulmonary nodules in high resolution computerized tomography (CT) images are potential manifestations of lung cancer. As pointed by literatures [1, 2, 3, 4], despite much effort being devoted to the nodule segmentation problem, segmentation for various types of pulmonary nodules remains an ongoing research topic [4]. Some classical pulmonary nodules are shown in Fig. 1. One of the major difficulties is the task of segmentation of non-solid and part-solid ground-glass opacity (GGO) nodules with faint contrast and fuzzy margins (shown as Fig. 1d). In particular, non-solid nodules are extremely subtle with fuzzy boundaries, and part-solid nodules exhibit highly irregular intensity variations (called intensity inhomogeneity) and boundary shapes. Studies have shown that nodules of non-solid and part-solid nature are frequent and have higher risks of being malignant than solid ones [1]. Additionally, there is difficulty associated with the segmentation of nodules that are adjacent to vessels when they have very similar intensities (shown as Fig. 1c Juxta-vascular nodule); and are nonspherical in shape. Juxta-vascular nodules account for the largest typology of lung nodules [2]. Thus, handling them under a united framework poses a great challenge to the task of segmentation of pulmonary nodules. Although various algorithms have been reported in literatures [1, 2, 3, 4, 5] for tackling these problems, the technical issues of segmentation still remain. Traditional segmentation methods, such as purely intensity thresholding or model-based segmentation methods, may fail to segment various types of nodules, especially GGO and juxta-vascular nodules, leading to boundary leakage or over-segmentation. All these factors lead to the belief that the field is relatively new and requires further investigation. This paper deals with the above mentioned problems. In this paper, a novel segmentation method is proposed for various types of pulmonary nodules in CT images, especially for GGO nodules (part-solid and nonsolid) and juxta-vascular nodules.

## Previous work on segmentation of pulmonary nodules

The segmentation under a united framework of kinds of pulmonary nodules, especially for GGO nodules and juxta-vascular nodules, is a very difficult task. Active contour models (ACMs) have been one of the most successful methods for image segmentation [6, 7, 8, 9, 10], even in the segmentation of pulmonary nodules [11]. However, GGO nodules possess weak edges, intensity inhomogeneity and irregular shape, so when they are segmented by using traditional edge-based ACMs [12], region-based ACMs [6, 7], or even some complex integrated ACMs combine edge and region energy [13, 14], they



**Fig. 1** Some classical pulmonary nodules. **a** Well-circumscribed nodule; **b** Juxta-pleural nodule; **c** Juxta-vascular nodule; **d** GGO nodule

often occur the problem of boundary leakage. Generally, the model using local information commonly can obtain better performance than that of global statistic information in solving the segmentation problems of intensity inhomogeneity and zigzag edge [7, 15, 16, 17]. The idea of defined local neighbor is more reasonable, especially for segmenting the zigzag and inhomogeneous edge. Li et al. [7] presented a region-based active contour model (ACM) model, in which a data fitting energy is introduced to solve the problem of intensity inhomogeneity. Lankton et al. [16] summarized and proposed different segmentation models based on local information. But typically, the segmentation curve cannot obtain exact edge or deviates from the objects if the local domain is too large or too small [17]. Further, an active contour model based on nonparametric independent and identically distributed statistics of the image may segment an image according to the particular global to local strategy. The local histogram method using Wasserstein distance to measure distribution distance has a good performance in segmenting cluttered scenes [15]. However, the estimation effect in image segmentation is severely influenced by the small cabin volume and the sample distribution, when the pixel density functions are estimated by histogram method. It is difficult to acquire an ideal segmentation effect only by relying on general intensity-data-driven segmentation methods. In order to overcome the limitation, Krinidis et al. [18] used fuzzy energy to solve the problem of “weak” local minima. Also Assen et al. [19] presented a 3D ACM driven by fuzzy inference for cardiac CT and MR images. Zhang et al. [20] added the Bayesian error of edge direction and region statistical information into the ACM model, to improve the convergence speed.

As mentioned above, juxta-vascular nodules account for the largest typology of lung nodules [2]. So besides handling GGO nodules with intensity inhomogeneity and weak edges, it is also important for a segmentation algorithm to be able to treat juxta-vascular nodules. In clinic application, even some nodules are not only GGO but also juxta-vascular nodules. How to handle various types of pulmonary nodules, including GGO nodules and juxta-vascular nodules, under a united framework presents a great challenge [2, 3]. Since vessels can be characterized by the tubular models and a pulmonary nodule is a small round or oval-shaped growth in the lung, so many approaches based on morphological operators [2, 21, 22] have been proposed to segment the juxta-vascular nodules. However the sizes and shapes of vessels as well as those of nodules are irregular, it may lead to the problem of a small volume overestimation if only morphological correction is relied upon. Hence a better segmentation refinement method should be taken into consideration furtherly. Besides intensity feature, the analysis of the shape of pulmonary structures has often been adopted to recognize small lung nodules from the background anatomy [2]. However, approaches utilizing simple criteria like shape rule or gray value evidence are typically not suitable to differentiate between different tubular tree structures and nodules. Lung nodules are embodied in a complex and structured background. Their identification and segmentation is usually affected by surrounding anatomical objects [2]. So, in a broad sense, the feature space for the recognition of nodules should be embed more prior information, including the target structures [23, 24].

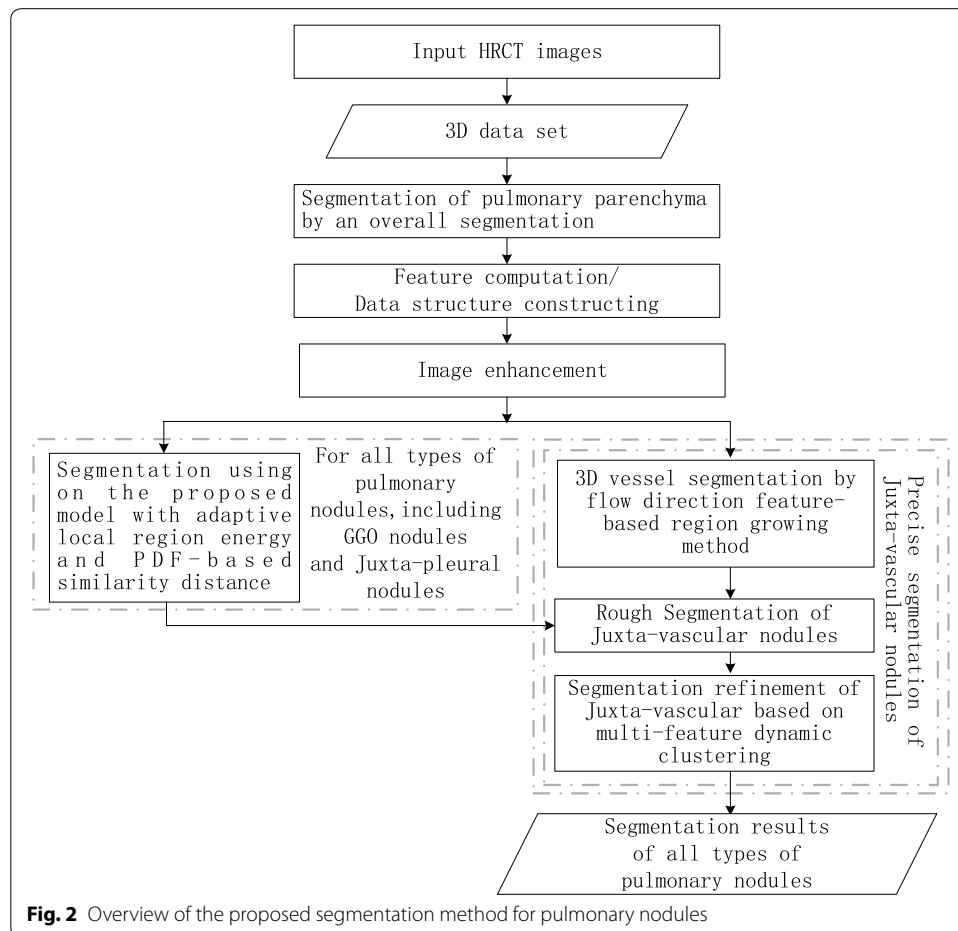
To our knowledge, there are few literatures aimed at handling GGO and juxta-vascular nodules under a united framework and multi-features classification space. In our previous work, we have built a very preliminary fuzzy integrated ACM incorporated

multi-features analysis to realize segmentation of GGO and juxta-vascular nodules [14, 25, 26]. This paper deals with the above mentioned problems further. In our present study, the segmentation problem is converted into the optimization problem of measurable distance between probability density functions of multi-features. A multi-features dynamic clustering method is used for the segmentation refinement of juxta-vascular nodules.

**Our approach**

In this paper, a novel segmentation method for pulmonary nodules, especially for GGO nodules and juxta-vascular nodules in CT images is proposed, which uses an adaptive local region energy model with probability density function (PDF)-based similarity distance and multi-features dynamic clustering refinement method. The flowchart of the proposed segmentation algorithm for pulmonary nodules under a united framework is shown as Fig. 2.

Compared with existing traditional methods, our approach has several novel aspects: (1) in the proposed adaptive local region energy model, the local domain for local energy model is selected adaptively based on k-nearest-neighbour (KNN) estimate method, and measurable distances between probability density functions of multi-dimension features



with high class separability are used to build the cost function. (2) A multi-features dynamic clustering method is used for the segmentation refinement of juxta-vascular nodules, which is based on the nodule segmentation using active contour model with adaptive local region energy and vessel segmentation using flow direction feature (FDF)-based region growing method. (3) It handles various types of nodules under a united framework.

The remainder of this paper is organized as follows. In “[The proposed integrated ACM with adaptive local region energy and PDF-based similarity distance](#)” section and “[Segmentation refinement of juxta-vascular nodules based on multi-features dynamic clustering](#)” section, the proposed segmentation methods of pulmonary nodules are introduced. The experimental results of our method are given in “[Experimental results](#)” section, followed by some discussions in “[Discussion](#)” section. This paper is summarized in “[Conclusions](#)” section.

### **The proposed integrated ACM with adaptive local region energy and PDF-based similarity distance**

We will approach the segmentation problem as a typical optimization task. Thus we need to adopt an appropriate model of energy function as the cost function. The proposed integrated ACM with adaptive local region energy and PDF-based similarity distance differs from the model used in literatures [7, 13, 14, 18, 19], which has the following characteristics.

- (1) The KNN-based adaptive local energy function model is proposed. The nodule with GGO characteristic is either part-solid or nonsolid, in which case it possesses typical weak edges and intensity inhomogeneity. The model using local information commonly can obtain better performance than that of global statistic information in solving the segmentation problem [7, 15, 16]. The local domain for local energy model is selected adaptively, which is approached as non-supervised recognition problem and realized by a KNN estimate method based on medical prior knowledge. It will be explained in detail in “[The proposed KNN-based adaptive local energy function](#)” section.
- (2) The segmentation problem is converted into the optimization problem of measurable distance between PDF of multi-features. The Bhattacharyya distance function is applied to measure the distance of PDF between the foreground and background, and the PDF in the local region are measured by Wasserstein distance. It will be explained in detail in “[The similarity distance based on adaptive local region probability density](#)” section and “[The PDF-based Bhattacharyya similarity distance for global energy](#)” section.
- (3) Multi-features information with high class separability is reflected and used in the proposed integrated ACM model. It will be explained in detail in “[Generation of multi-dimension feature with high class separability](#)” section.

Taking both the edge, local and global region information into consideration, our proposed energy function of integrated active contour model  $E$  is given as Eq. (1).

$$E = E_{edge} + E_{local} + E_{global} \quad (1)$$

where  $E$  is the proposed energy function model;  $E_{edge}$  is the edge-driven energy term, which is used for curved surface to improve the evolved ability in concave regions;  $E_{local}$  is the local-region-driven energy term;  $E_{global}$  is the global-region-driven energy term.  $E_{local}$  and  $E_{global}$  are used to control the image force based on statistical multi-features information in the region and move the curved surface in the decrease direction of feature variance.

Let  $\Omega \subset \mathbb{R}^3$  be the image domain, and  $D : \Omega \rightarrow R$  be the given medical CT image sequence or 3D data set. The segmentation result of the images or data set (for 3D data set)  $D$  is achieved by finding a surface  $\phi$ , which separates  $\Omega$  into disjoint regions.  $\Omega_1$  and  $\Omega_2$  represent the inside regions and outside regions of  $\phi$ , respectively. Besides intensity, more features are used in our active contour model. Our proposed detailed energy function model is given as Eq. (2) based on Eq. (1). Why and how to build the energy term model as Eq. (2) will be explained in detail in the following sections (from “The proposed KNN-based adaptive local energy function”, “The similarity distance based on adaptive local region probability density”, “Generation of multi-dimension feature with high class separability”, “The PDF-based Bhattacharyya similarity distance for global energy” section).

$$\begin{cases} E(\phi) = E_{edge} + E_{local} + E_{global} \\ E_{edge} = \mu \int_{\phi} \delta\phi(v)g|\nabla\phi|dv \\ E_{local} = \int_{\Omega_+} \lambda_1[1 - u(v)]^m H(\phi(v)) \int_{L_{\tau}} \|P(v) - P_1(\tau)\|d\tau dv \\ \quad + \int_{\Omega_-} \lambda_2[1 - u(v)]^m (1 - H(\phi(v))) \int_{L_{\tau}} \|P(v) - P_1(\tau)\|d\tau dv \\ E_{global} = \int_{\Omega} \sqrt{P_-(\tau)P_+(\tau)}d\tau \end{cases} \quad (2)$$

where  $E(\phi)$  is the proposed energy function model;  $v(x, y, z) \in \Omega$  is a given pixel/voxel. In term  $E_{local}$  of Eq. (2),  $\lambda_1$  and  $\lambda_2$  are the weights for region-driven energy term. The membership function  $u(v) \in [0, 1]$  is the degree of membership of  $D$ , and  $m$  is a weighting exponent on each fuzzy membership. The degree of membership is decided by multi-features value  $L$ .  $L$  is the value domain of feature space.  $H\phi$  is the smoothed Heaviside function.  $P(v)$  is the probability density of multi-features vector  $O$  of  $v$ ;  $P_1(\tau)$  and  $P_2(\tau)$  are the means of probability density in local region  $\Omega_{\tau}$  selected adaptively. In term  $E_{global}$  of Eq. (2),  $P_-(\tau)$  and  $P_+(\tau)$  are the kernel-based estimates of the image features observed over the sub-domains  $\Omega_-$  and  $\Omega_+$ .  $E_{edge}$  is similar to our previous work [14], which is used for curved surface to improve the evolved ability in concave regions. In term  $E_{edge}$  of Eq. (2),  $\mu$  is the weight of edge-driven energy term;  $\delta\phi(v)$  is the smoothed version of the Dirac delta;  $g$  is the stop function.

### The proposed KNN-based adaptive local energy function

The model using local information commonly can obtain better performance. The proposed local energy function model is inspired initially by literature [7, 15]. But in our proposed local energy function model, the local domain is not fixed, but flexible and adaptive, which is realized by a KNN estimate method here. In order to segment

elaborately for the image with zigzag edge and noise interference, we construct  $k$ -nearest neighbors and estimate the corresponding probability density functions with Parzen window method in each pixel/voxel.

As we have known, lung nodules, especially malignant nodules, sometimes show spiculation and lobulation signs in CT images. Typically, the segmentation curve deviates from the objects if the local neighbor radius is too large or too small. Consider the synthetic image in Fig. 3. The segmentation results are very different with various radii with the implementation of local histogram-based segmentation method using the Wasserstein Distance [15]. For instance, conventional local segmentation algorithms often suffer from disturbances induced by the mass of irrelevant information when the radius of the neighbor increases. To improve the ability to segment precisely the object where is intensity inhomogeneity, zigzag and difficult to define the boundary, the local region  $\Omega_r$  of  $\int_{L_r} \|P(v) - P_1\| d\tau$  and  $\int_{L_r} \|P(v) - P_2\| d\tau$  in term  $E_{local}$  of Eq. (2), should be selected adaptively. So the KNN-based adaptive local energy function model is proposed. The local domain for local energy model is selected adaptively, which is approached as non-supervised recognition problem and realized by a KNN estimate method based on medical prior knowledge.

Let  $x, y, z$  be the location variables of  $v$ , respectively. For each current voxel  $v$ , selecting a circular (sphere) neighbor domain  $R_L$  with radius  $r$ ; selecting  $k$  nearest neighbors of  $O(v)$  to construct the local domain  $N_k$ , we have the complementary set  $N_c$  of  $N_k$  in the  $R_L$ . If each voxel of  $R_L$  is met the condition as Eq. (3), select  $R_L$  as the local region for computing similarity in Eq. (3).

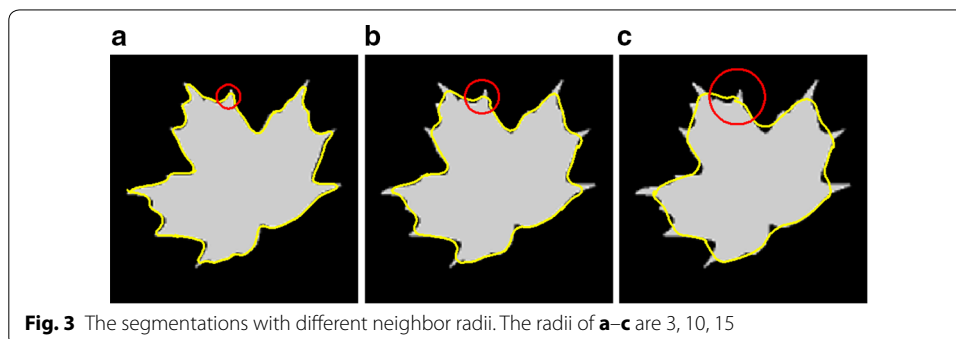
$$d(O(x), O(y)) \leq d(O(x), O(z)) \quad (3)$$

where  $y \in N_k, z \in N_c$ ;  $d(O(x), O(y))$  is the similarity distance, as Eq. (4).

$$d(O(x), O(y)) = \|J(O(x)) - J(O(y))\| \quad (4)$$

where  $J(O)$  is the normalized multi-features vector  $O$ .

In the KNN-based adaptive local energy function model, the parameters  $r$  and  $k$  are determined by medical prior knowledge and the physical resolution of the pixel in CT imaging. As shown in Fig. 4, the local region  $\Omega_r$  in 2-dimensional space is determined by the parameters  $r$  and  $k$ . Rules for determining the parameters  $r$  and  $k$ , as well as their reasons, are as follows.



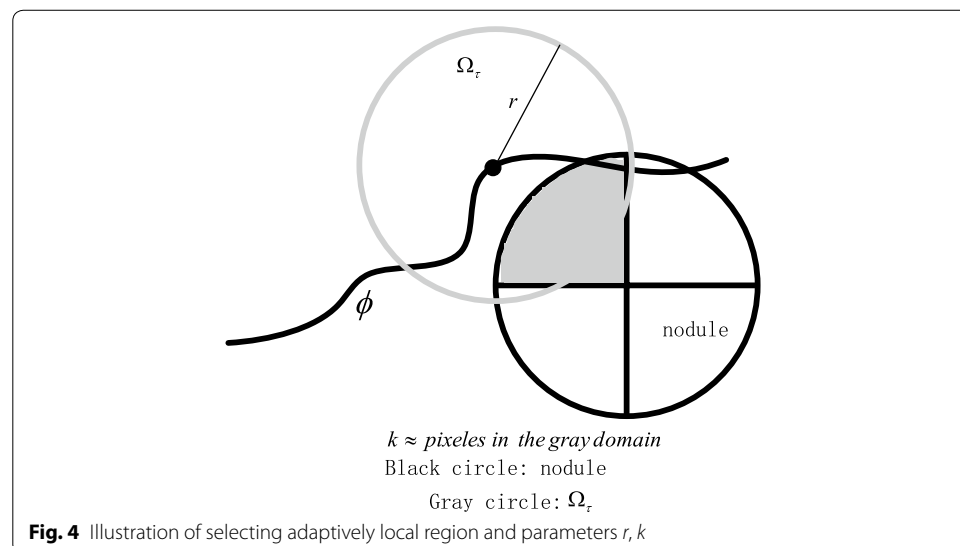
Rule 1:  $r$  and  $k$  are selected as large as possible

Rule 2:  $r \leq (8 \text{ mm}/2)/(\text{pixel spacing})$

Reasons are below: (1) The best choice of  $k$  depends upon the data. Generally, larger values of  $k$  reduce the effect of noise on the classification, but make boundaries between classes less distinct. (2) A pulmonary nodule is a small round or oval-shaped growth in the lung. It is sometimes also called a spot on the lung or a coin lesion. So the local region  $\Omega_r$  is selected by a circle (a sphere for 3-dimensional space). (3)  $k \leq \pi r^2$  (for  $N_k \subset \Omega_r$ , in 2-dimensional space). So  $r$  is selected as large as possible. (4) Pulmonary nodules are generally smaller than 3 cm (about  $\leq 3$  cm) in diameter. If the growth is larger than that, it is known as a pulmonary mass. (5) The nodule may have first been identified by a CT scan. CT scans can give information about the specific features of the nodule, including its shape, size, location and internal density. A CT scan can find very small nodules, as small as 1–2 mm in diameter. Most benign (not cancerous) nodules are small (less than 5 mm) in size. Most nodules between 5 and 10 mm will need additional imaging unless they are unlikely to be cancer based on the way they look. Larger nodules require more careful evaluation and examination including additional imaging tests and possibly a biopsy. So we can determine the parameters  $r$  and  $k$  according to the above medical prior knowledge. (6) According to medical knowledge, nodules less than 8 mm are usually too small for a biopsy. In other word, the nodule more than 8 mm should be evaluated carefully. In other words, we should make sure of the distinct in the local region of the nodule which is more than 8 mm.

Rule 3:  $k \leq \pi(8 \text{ mm}/2(\text{pixel spacing}))^2/K_0$ ,  $K_0 = 4$

Reasons are below: (1) The local region  $\Omega_r$  is selected by a circle (for 2-dimensional space). (2)  $N_k \subset \Omega_r$ . (3) Value of  $k$  should make boundaries between classes more distinct in the local region. (4) The following experiment of a dataset (the phantom nodules) was performed to determine and test the parameters.





In the experiment, some nodule models (5 mm in diameter, from  $-406$  to  $-477$  HU, shown as Fig. 5c) in a phantom (shown as Fig. 5a) are used to determine and test the parameters. A quantitative analysis was performed to determine the parameters. The well known Tanimoto/Jaccard error  $A(C_m, C_o)$  is used as the validation metrics, which refers to volume overlaps between the gold standard and the proposed segmentation method with different  $K_0$ . In the experiment, the gold standard is the boundary of the phantom nodule, which is shown as Fig. 5d.  $A(C_m, C_o)$  is defined as Eq. (5).

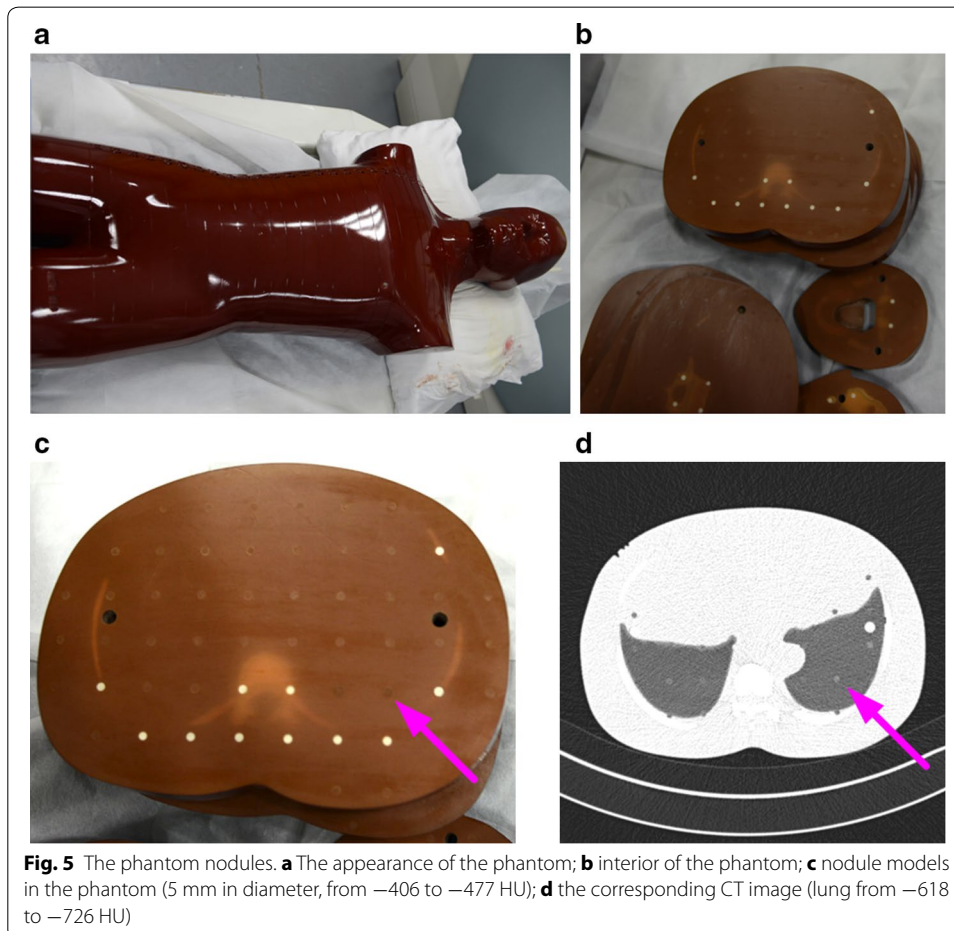
$$A(C_m, C_o) = 1 \frac{\int_{C_m \cap C_o} dx dy dz}{\int_{C_m \cup C_o} dx dy dz} \quad (5)$$

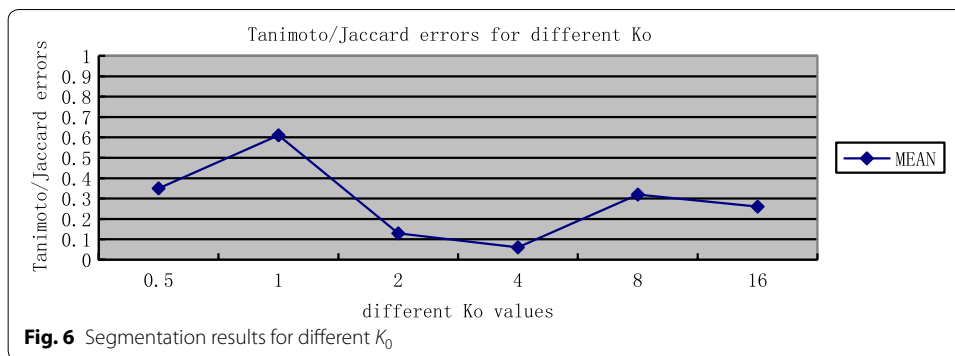
where  $C_m$  and  $C_o$  are the extracted and the desired contours, respectively.

Here,  $k = \pi r^2 / K_0 \cdot r$  is selected as  $5 \text{ mm} / (2 \text{ pixel spacing})$ , which is corresponded to the size of the phantom nodule (5 mm in diameter). Segmentation results for different  $k_0$  are shown in Fig. 6. Figure 6 shows that the segmentation result with  $K_0 = 4$  is the best.

#### The similarity distance based on adaptive local region probability density

The segmentation problem is converted into the optimization problem of measurable distance between probability density functions. In this paper the probability density  $P(v)$





**Fig. 6** Segmentation results for different  $K_0$

of multi-features vector  $O$  for the pixel/voxel  $v$  is estimated by the Parzen window estimation method. Gaussian function is used as the kernel function. The probability density function of neighbor pixel is represented as Eq. (6).

$$P(v) = \frac{1}{k} \int_{N_x} K(\|O(v)\| - \|O(x)\|) dx \tag{6}$$

where  $K(\tau) = (\sqrt{2\pi}\sigma)^{-1} \exp\{-\tau^2/2\sigma^2\}$  is a kernel function, and generally as Gaussian density kernel function, and  $\int K(\tau) d\tau = 1$ .

The similarity distance  $W(P, P_i)$  between  $P(v)$  and  $P_i$  is computed by Wasserstein distance function, as Eq. (7).

$$W(P, P_i) = \int_{L_\tau} \|P(v) - P_i(\tau)\| d\tau |_{i=1,2} \tag{7}$$

where  $L_\tau$  is the value domain of feature space,  $P(v)$  is the probability density of multi-features vector  $O$  of  $v$ ;  $P_i(\tau)$  are the means of probability density in local region  $\Omega_\tau$  selected adaptively, which is as Eqs. (8) and (9).

$$P_1(\tau) = \overline{P(\tau)}, \quad \text{where } \{v : \phi(v) \geq 0\} \tag{8}$$

$$P_2(\tau) = \overline{P(\tau)}, \quad \text{where } \{v : \phi(v) < 0\} \tag{9}$$

**Generation of multi-dimension feature with high class separability**

For one of the major difficults is the task of segmentation of non-solid and part-solid GGO nodules with faint contrast and fuzzy margins. In particular, non-solid nodules are extremely subtle with fuzzy boundaries, and part-solid nodules exhibit highly irregular intensity variations (intensity inhomogeneity) and boundary shapes. GGO pulmonary nodules are hard to distinguish if merely the intensity feature is utilized. Multi-features information with high class separability is reflected and used in the proposed integrated ACM model.

Here, in the feature space, a sample is  $O_i = (x_p, y_p, z_p, I_p, \text{level of } ASM_i)$ ;  $x_p, y_p$  and  $z_p$  is the position feature;  $I_p$  is the intensity feature and  $ASM_i$  is the angular second moment (texture feature) which is as Eq. (10).

$$ASM = \sum_{m=0}^{N_g-1} \sum_{n=0}^{N_g-1} (P(m, n))^2 \tag{10}$$

where  $P(m, n)$  is the matrix element of the normalized gray-level co-occurrence matrix.  $N_g$  is the number of possible gray levels. A gray level co-occurrence matrix is a matrix where the number of rows and columns is equal to the number of gray levels. The matrix element  $P(m, n | \Delta x, \Delta y)$  is the relative frequency with which two pixel, separated by a pixel distance  $(\Delta x, \Delta y)$ , occur within a given neighborhood, one with intensity  $m$  and the other with intensity  $n$ . Given an  $M \times N$  neighborhood, let  $f(m_p, n_l)$  be the intensity at sample  $s$ , line  $l$  of the neighborhood.  $P(m, n | \Delta x, \Delta y)$  is defined as Eq. (11).

$$P(m, n | \Delta x, \Delta y) = WQ(m, n | \Delta x, \Delta y) \quad (11)$$

where

$$W = \frac{1}{(M - \Delta x)(N - \Delta y)}$$

$$Q(m, n | \Delta x, \Delta y) = \sum_{l=1}^{N-\Delta y} \sum_{s=1}^{M-\Delta x} A$$

and

$$A = \begin{cases} 1 & \text{if } f(s, l) = m \text{ and } f(s + \Delta x, l + \Delta y) = n \\ 0 & \text{elsewhere} \end{cases}$$

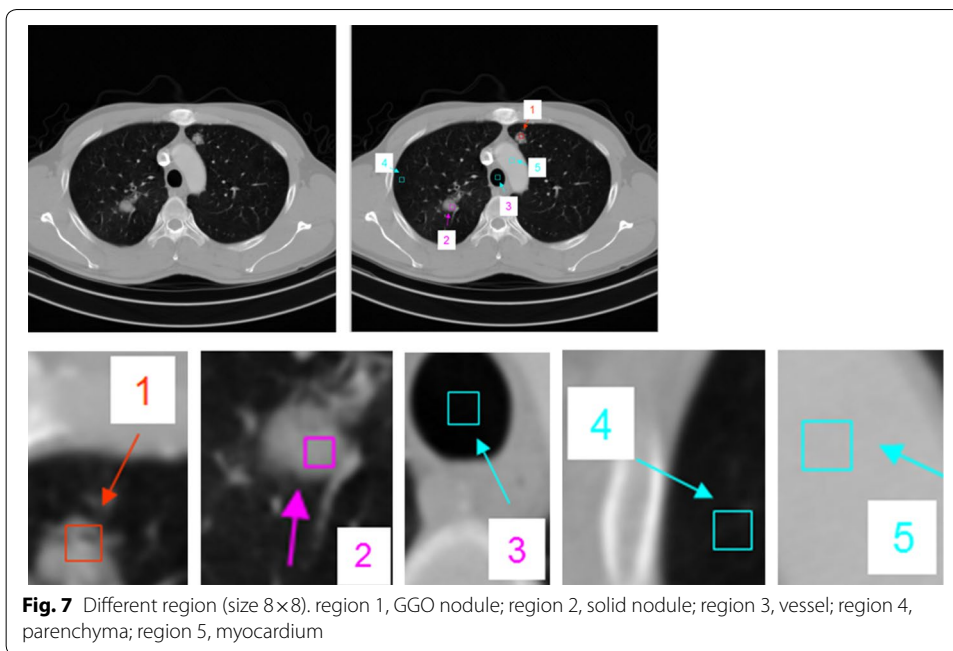
This feature  $ASM_i$  is a measure of the smoothness of the image. Indeed, if all pixels are of the same gray-level  $I = k_i$ , then  $P(k_i, k_i) = 1$  and  $P(m, n) = 0$ ,  $m \neq k_i$  or  $n \neq k_i$ , and  $ASM = 1$ . At the other extreme, if we could have all possible pairs of gray levels with equal probability  $\frac{1}{R}$ , then  $ASM = \frac{1}{R}$ . The less smooth the region is, the more uniformly distributed  $P(i, j)$  and the lower the  $ASM$ .

An example is illustrated in Fig. 7 and Table 1. It implies that the  $ASM$  value is an important feature for segmentation of GGO nodules, solid nodules or other regions.

The segmentation problem is converted into the optimization problem of measurable distance between PDF  $P(v)$  of multi-features  $O$ , shown in Eq. (2). In proposed integrated ACM model,  $P(v)$  is the probability density of multi-features vector  $O_i = (x_p, y_p, z_p, I_p, \text{level of } ASM_i)$  of  $v$ . In “The similarity distance based on adaptive local region probability density” section,  $P(v)$  is estimated by the Parzen window estimation method. In order to reduce the computation cost and save the memory,  $ASM_i$  is sampled as level 0 (from 0.0000 to 0.1500), level 1 (from 0.1501 to 0.2300), level 2 (from 0.2301 to 0.3500), level 3 (from 0.3501 to 0.5000), level 4 (from 0.5001 to 0.8000) and level 5 (from 0.8001 to 1.0000).

### The PDF-based Bhattacharyya similarity distance for global energy

In order to obtain more stable segmentation and achieve the global optimal segmenting, the global energy term  $E_{global}$  shown as Eq. (2), is incorporated into the proposed integrated ACM model.  $E_{global}$  is global energy term which maximizes the distance of probability density function in different regions. Bhattacharyya distance is applied to measure the distance of probability density distributions between the foreground and background. Bhattacharyya coefficient [27, 28] is defined as Eq. (12).



**Fig. 7** Different region (size 8×8). region 1, GGO nodule; region 2, solid nodule; region 3, vessel; region 4, parenchyma; region 5, myocardium

**Table 1** Second order histogram (angular second moment, ASM) feature value for different region

	1 GGO nodule	2 Solid nodule	3 Vessel	4 Parenchyma	5 Myocardium
ASM	0.2212	0.2448	0.4828	0.4817	0.6403

$$E_{global} = \int_{\Omega} \sqrt{P_-(\tau)P_+(\tau)}d\tau \tag{12}$$

where  $E_{global}$  can be thought of as a direction cosine between two points on the sphere, so its value domain is [0,1]. This measure varies between 0 and 1, where 0 indicates complete mismatch, and 1 indicates a complete match.

In global energy term  $E_{global} = \int_{\Omega} \sqrt{P_-(\tau)P_+(\tau)}d\tau$ ,  $P_-(\tau|\phi(x))$  and  $P_+(\tau|\phi(x))$  are kernel-based estimates of the image features observed over the subdomains  $\Omega_-$  and  $\Omega_+$ . The kernel-based estimates are given by Eq. (13).

$$\begin{cases} P_-(\tau|\phi(x)) = \frac{\int_{\Omega} K_-(\tau - \|O(x)\|)H(-\phi(x))dx}{\int_{\Omega} H(-\phi(x))dx} \\ P_+(\tau|\phi(x)) = \frac{\int_{\Omega} K_+(\tau - \|O(x)\|)H(\phi(x))dx}{\int_{\Omega} H(\phi(x))dx} \end{cases} \tag{13}$$

where  $\tau \in R^N$ ,  $K_-(\tau)$  and  $K_+(\tau)$  are two Gaussian density functions that are normalized to have unit integrals with respect to the feature vector  $\tau$ , viz.  $\int_{R^N} K_-(\tau)d\tau = \int_{R^N} K_+(\tau)d\tau = 1$ .

### The proposed model and its numerical implementation

The proposed integrated ACM model is shown as Eq. (2). The zero level set that makes the energy function Eq. (2) minimum is the optimal segmentation result. In the simple case, it is obvious that the boundary of the object  $\phi_0$  is the minimizer of the energy functional. The energy function model as Eq. (2) is solved by using variational level set approach. The numerical implementations of energy function as Eq. (14).

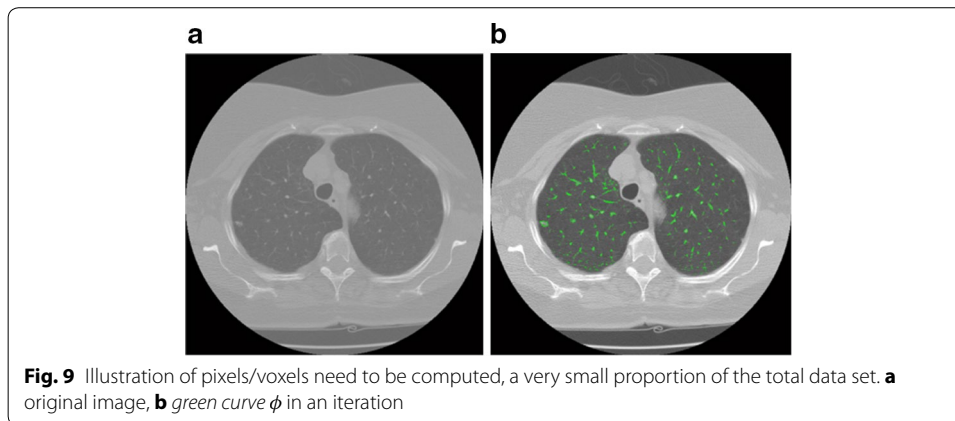
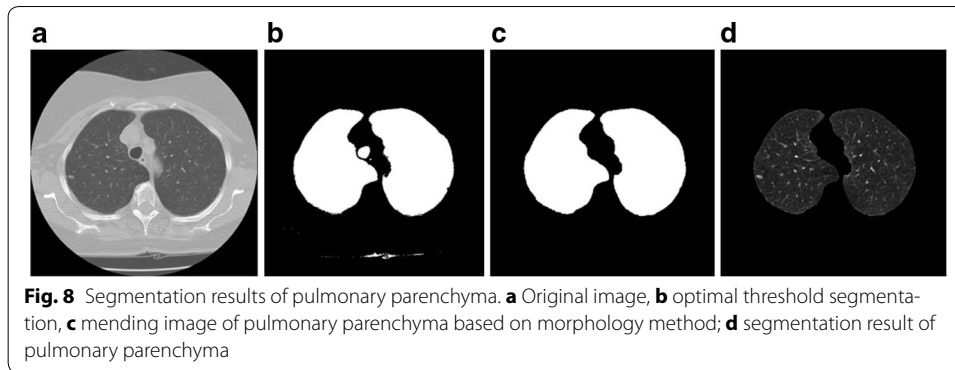
$$\begin{aligned}
 \frac{\partial \phi(v)}{\partial t} = & \delta_{\in} \{ u \operatorname{div} \left( g \frac{\nabla \phi(v)}{|\nabla \phi(v)|} \right) - \lambda_1 [u(v)]^m \int_{L_{\tau}} \|P(\tau) - P_1(\tau)\| d\tau \\
 & + \lambda_2 [1 - u(v)]^m \int_{L_{\tau}} \|P(\tau) - P_2(\tau)\| d\tau \} \\
 & + \frac{1}{2} (A_-^{-1} - A_+^{-1}) \int_{L_{\tau}} \sqrt{P_-(\tau|\phi(v))P_+(\tau|\phi(v))} d\tau \\
 & + \int_{\Omega_z} \frac{1}{A_+} K_+(\tau - \|O(x)\|) \sqrt{\frac{P_-(\tau|\phi(v))}{P_+(\tau|\phi(v))}} d\tau \\
 & + \int_{\Omega_z} \frac{1}{A_-} K_-(\tau - \|O(x)\|) \sqrt{\frac{P_+(\tau|\phi(v))}{P_-(\tau|\phi(v))}} d\tau
 \end{aligned} \tag{14}$$

where  $A_- = \int_{\Omega} H(-\phi(x)) dx$  and  $A_+ = \int_{\Omega} H(\phi(x)) dx$  are the size of the image sub-domains  $\Omega_-$  and  $\Omega_+$ .

### Implementation of potential pulmonary nodule segmentation based on the proposed integrated ACM model

The implementation algorithm for the proposed integrated ACM model is as follows.

- (1) The pulmonary parenchyma is segmented by an overall segmentation method combining thresholding and morphology, which is shown in Fig. 8.
- (2) Multi Feature computation and data structures for data set are constructed.
- (3) Compute the degree of membership  $u(v)$ . In the proposed model, the fuzzy energy is used as the model motivation power evolving the active contour. As shown in Eq. (2),  $u(v): X \rightarrow [0, 1]$  defines the membership degree of a voxel  $v$  in data set  $D$  to the nodule class cluster center. It is different with our previous work [14], the sample in the AFIACM model is  $X_i = (x_i, y_i, z_i, I_i, ASM_i)$ . Thus, the degree of membership for each sample  $X_i = (x_i, y_i, z_i, I_i, ASM_i)$  in our model is calculated by using the fuzzy clustering algorithm based on intensity and angular second moment features.
- (4) Compute the scale parameters  $r$  and  $k$  according to “The proposed KNN-based adaptive local energy function” section, and select adaptively local region  $R_L$  and  $B(v)$ .
- (5) Specify the stop function term using posterior probability.
- (6) Implement the numerical algorithm of the proposed model according to Eq. (14). Here  $P(v)$ ,  $P_1(\tau)$  and  $P_2(\tau)$  in the local region  $\Omega_{\tau}$  are computed according to “The similarity distance based on adaptive local region probability density” section and “Generation of multi-dimension feature with high class separability” section. In each iteration, only  $P(v)$  in the local region  $\Omega_{\tau}$  near to the curve/surface  $\phi$  need to be computed, so pixels/voxels need to be computed are a very small proportion of



the total data set. An example is illustrated in Fig. 9. Only  $P(v)$  in the local region  $\Omega_\tau$  near to the green curve  $\phi$  need to be computed. The numbers are 1591, which is 0.6 % of the whole data set ( $512 \times 512$ ).

### Segmentation refinement of juxta-vascular nodules based on multi-features dynamic clustering

In order to overcome the problem of a small volume over-segmentation in the adhesion region between the juxta-vascular nodule and its attached vessel, and obtain a better segmentation result, a multi-features dynamic clustering method is used for the segmentation refinement of juxta-vascular nodules, which is based on the nodule segmentation using the proposed integrated ACM model and vessel segmentation using FDF-based region growing method. Various types of pulmonary nodules, including GGO nodules (part solid and nonsolid) and juxta-vascular nodules, are segmented under a united segmentation framework. The refinement procession is just used for the pixels/voxels in some boundary regions between the juxta-vascular nodules and blood vessels, without modifying the nodule boundary elsewhere.

#### The refinement region selection using the proposed nodule and vessel segmentation method

##### 3D potential vessel segmentation by FDF-based region growing method

In this paper, the 3D vessel segmentation is used as the rough segmentation process of segmentation refinement of juxta-vascular nodules. 3D vessel segmentation can be

accomplished by some traditional methods, such as region growing [23] and vessel enhancement filter [29, 30]. However shapes and appearances of vessels are irregular, a satisfactory result often cannot be achieved if only these traditional methods are relied upon. According to the medical prior knowledge, vessels are characterized by a tubular model, the 3D gradient vectors in a vessel can be used to extract a vector in the direction of the vessel by identifying a vector that is approximately orthogonal to the gradients in a local neighborhood [21, 22, 31]. Moreover, an ideal vascular structure not only keeps a certain elongation, but also has no holes. Simply connected domain is a particular connected domain, in which every loop can be continuously pulled to a point without leaving the space [24]. So we can use these shapes and appearances as the constraints of the segmentation method, which are as follows. (1) Since vascular structures should have no holes, the segmented objects must be removed if they are not simply connected. (2) the flow direction can be used as the growing direction constraint condition, which is one of the constraint conditions of region growing method for segmentation of blood vessels and attached nodules.

Here a FDF-based region growing method is proposed for 3D vessel segmentation. In the FDF-based region growing method, the growing constraint conditions include the intensity threshold and flow direction defined by flow direction vector  $l_d$ . Assume  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  ( $\lambda_1 \leq \lambda_2 \leq \lambda_3$ ) are eigenvalues of the structure tensor defined by the statistical information of arithmetic average in local region. The structure tensors are computed by directly using the gradient information of each voxel [14, 29]. Let  $e_1$  be the unit length eigenvector belonging to the eigenvalue  $\lambda_1$ . The flow direction vector  $l_d$  of the vessel is set as Eq. (15).

$$l_d = \sqrt{\lambda_2 + \lambda_3} \cdot e_1 \quad (15)$$

#### **Refinement region selection**

In this paper, the intersection between potential vessel region and potential nodules is selected as the refinement region of juxta-vascular nodules based on multi-features dynamic clustering. The processes are as follows.

- (1) The rough 3D potential vessel segmentation result  $S_{c1}$  is gotten by using the flow direction feature-based region growing method which is described in “3D potential vessel segmentation by FDF-based region growing method” section.
- (2) The segmentation result  $S_{c2}$  is gotten by using the segmentation method based on the proposed integrated ACM model described in “The proposed integrated ACM with adaptive local region energy and PDF-based similarity distance” section.
- (3) The potential object  $S_r$  for segmentation refinement is  $S_r = S_{c1} \cap S_{c2}$ .

#### **Generation and construction of multi-features vector in clustering space**

Our proposed multi-features vector takes both the appearance and geometric information into consideration. Because CT values of pulmonary nodules and that of blood vessels are almost uniform, it is difficult to segment exactly the blood vessels and the attached nodules if merely the intensity feature is utilized. So the constructed feature space takes both the appearance and geometric multi-features information into consideration. We build a multi-features space for dynamic clustering in segmentation

refinement region. An extended observation vector  $X_i$  in the clustering space is defined as Eq. (16).

$$X_i = (x_i, y_i, z_i, I_i, SI_i) \quad (16)$$

where  $x_i, y_i$  and  $z_i$  are the position features;  $I_i$  is the intensity feature;  $SI_i$  is the volumetric shape index, reflecting the geometric shape, which is a measure of local shape characteristics [5].

#### Implementation of segmentation refinement based on multi-features dynamic clustering

The algorithm for segmentation refinement based on multi-features dynamic clustering is carried out as the following steps.

- (1) The segmentation refinement region  $S_r$  is selected according to “[The refinement region selection using the proposed nodule and vessel segmentation method](#)” section.
- (2) The feature vector  $X_i = (x_i, y_i, z_i, I_i, SI_i)$  is constructed and computed for the points in the region  $S_r$ , according to “[Generation and construction of multi-features vector in clustering space](#)” section.
- (3) The refinement region  $S_r$  is segmented by dynamic clustering method (K-Means clustering method [32]), then the juxta-vascular nodules is segmented exactly.

#### Experimental results

In this paper, a clinical dataset of 113 chest CT scans was used to evaluate the proposed method, which consisted of 60 thoracic CT scans obtained from LIDC databases [33] and 53 thoracic CT scans obtained from several big hospitals in Guangzhou City and Shenzhen City, Guangdong province, China. The list of these 60 CT scans from LIDC-IDRI databases is provided in Table 2. The used medical CT slices were data sets with an intensity value of 16bits and a resolution of  $512 \times 512$ . Slice thickness varied from 0.5 to 2.5 mm and the total slice number for each scan varied from 49 to 397 with an average of 136/scan. The X-ray tube current ranged from 30 to 280 mA, and the pixel size ranged from 0.5–0.75 mm/pixel. Pulmonary nodules in CT images are solid or GGO (part solid or nonsolid), whose sizes are from 3 to 30 mm. Locations of nodules are uncertain, some are isolated (Solitary Pulmonary Nodule, SPN), others are adhered to blood vessels or lung wall. Each scan was read individually by members of a qualified panel and then a consensual gold standard was defined by the panel. The panel members assigned each nodule to be either GGO, juxta-vascular, GGO juxta-vascular, or others. These nodules are manually segmented. This process defined ground truth of 157 nodules, and the number of different kinds of nodules is shown as Table 3.

Moreover, The different nodule size groups for the GGO nodules, juxta-vascular nodules, GGO juxta-vascular nodules and others is illustrated in Table 4. In our experiment, there are four types of nodules, and they are registered respectively. The performance of our segmentation fusion algorithm was evaluated using both quantitative and qualitative methods. Experimental results in each research step as well as some discussions are presented below.



**Table 2 The list of 60 CT scans from LIDC-IDRI databases**

	CT scans	Nodule	Type		CT scans	Nodule	Type
1	LIDC-IDRI-0003	Nodule 1	GGO	27	LIDC-IDRI-0160	Nodule 1	JV
		Nodule 2	GGO-JV			Nodule 2	Others
		Nodule 3	Others			Nodule 3	Others 10
		Nodule 4	GGO			Nodule 4	JV
2	LIDC-IDRI-0007	Nodule 1	Others	28	LIDC-IDRI-0162	Nodule 1	Others
3	LIDC-IDRI-0008	Nodule 1	GGO-JV			Nodule 2	JV
		Nodule 2	GGO			Nodule 3	JV
		Nodule 3	Others			Nodule 4	Others 6
4	LIDC-IDRI-0011	None				Nodule 5	Others 7
5	LIDC-IDRI-0015	Nodule 1	Others	29	LIDC-IDRI-0167	Nodule 1	Others
6	LIDC-IDRI-0017	Nodule 1	JV	30	LIDC-IDRI-0168	Nodule 1	JV
7	LIDC-IDRI-0018	None				Nodule 2	Others 12
8	LIDC-IDRI-0019	None		31	LIDC-IDRI-0173	Nodule 1	Others
9	LIDC-IDRI-0021	Nodule 1	GGO	32	LIDC-IDRI-0175	Nodule 1	JV
		Nodule 2	JV	33	LIDC-IDRI-0177	Nodule 1	JV
10	LIDC-IDRI-0025	None		34	LIDC-IDRI-0252	Nodule 1	JV
11	LIDC-IDRI-0032	Nodule 1	Others 2			Nodule 2	GGO-JVvascul
12	LIDC-IDRI-0037	Nodule 1	GGO	35	LIDC-IDRI-0273	nodule 1	JV
		Nodule 2	GGO	36	LIDC-IDRI-0350	Nodule 1	Others 14
		Nodule 3	Others	37	LIDC-IDRI-0477	Nodule 1	JV
13	LIDC-IDRI-0044	Nodule 1	GGO			Nodule 2	JV
		Nodule 2	JV	38	LIDC-IDRI-0580	Nodule 1	JV
		Nodule 3	JV	39	LIDC-IDRI-0626	Nodule 1	GGO
		Nodule 4	GGO	40	LIDC-IDRI-0645	Nodule 1	Others
		Nodule 5	Others	41	LIDC-IDRI-0652	Nodule 1	Others
		Nodule 6	GGO	42	LIDC-IDRI-0681	Nodule 1	GGO
14	LIDC-IDRI-0046	None		43	LIDC-IDRI-0684	Nodule 1	Others
15	LIDC-IDRI-0047	Nodule 1	GGO	44	LIDC-IDRI-0703	Nodule 1	Others
		Nodule 2	JV	45	LIDC-IDRI-0723	Nodule 1	Others
16	LIDC-IDRI-0050	Nodule 1	GGO	46	LIDC-IDRI-0796	Nodule 1	GGO
17	LIDC-IDRI-0051	None		47	LIDC-IDRI-0803	Nodule 1	GGO
18	LIDC-IDRI-0052	Nodule 1	GGO	48	LIDC-IDRI-0818	Nodule 1	Others
		Nodule 2	Others	49	LIDC-IDRI-0828	Nodule 1	GGO
19	LIDC-IDRI-0082	Nodule 1	Others 5	50	LIDC-IDRI-0840	nodule 1	Others
20	LIDC-IDRI-0114	Nodule 1	JV	51	LIDC-IDRI-0865	Nodule 1	GGO
21	LIDC-IDRI-0131	Nodule 1	GGO-JV	52	LIDC-IDRI-0882	Nodule 1	Others
		Nodule 2	JV	53	LIDC-IDRI-0914	Nodule 1	Others
22	LIDC-IDRI-0133	Nodule 1	Others	54	LIDC-IDRI-0928	Nodule 1	GGO
23	LIDC-IDRI-0141	Nodule 1	Others 7			Nodule 2	JV
		Nodule 2	GGO	55	LIDC-IDRI-0938	Nodule 1	GGO
		Nodule 3	Others	56	LIDC-IDRI-0986	Nodule 1	JV
		Nodule 4	JV	57	LIDC-IDRI-0915	Nodule 1	JV
		Nodule 5	GGO	58	LIDC-IDRI-0941	Nodule 1	JV
		Nodule 6	JV			Nodule 2	JV
24	LIDC-IDRI-0146	Nodule 1	JV			Nodule 3	JV
25	LIDC-IDRI-0152	Nodule 1	JV	59	LIDC-IDRI-0953	Nodule 1	JV
26	LIDC-IDRI-0159	Nodule 1	JV			Nodule 2	JV
				60	LIDC-IDRI-1012	Nodule 1	GGO

*JV* juxta-pleural nodule; *GGO* ground-glass opacity nodule; *GGO-JV* GGO juxta-vascular nodule

**Table 3 The number of GGO, juxta-vascular, GGO juxta-vascular and other nodules**

Nodule type	GGO pulmonary nodule	Juxta-vascular pulmonary nodule	GGO Juxta-vascular pulmonary nodule	Others	Total
Number	42	73	7	35	157

"GGO nodule" and "Juxta-vascular nodule" do not include GGO Juxta-vascular pulmonary nodule in the table

**Table 4 The different nodule sizes on testing data**

Nodule type	≤ 5 mm	5–10 mm	10–20 mm	Total
GGO pulmonary nodule	7	18	17	42
Juxta-vascular pulmonary nodule	5	49	19	73
GGO Juxta-vascular pulmonary nodule	1	3	3	7
Others	7	15	13	35

"GGO nodule" and "Juxta-vascular nodule" do not include GGO Juxta-vascular pulmonary nodule in the table

**Qualitative validation**

**Validation of the proposed integrated ACM model**

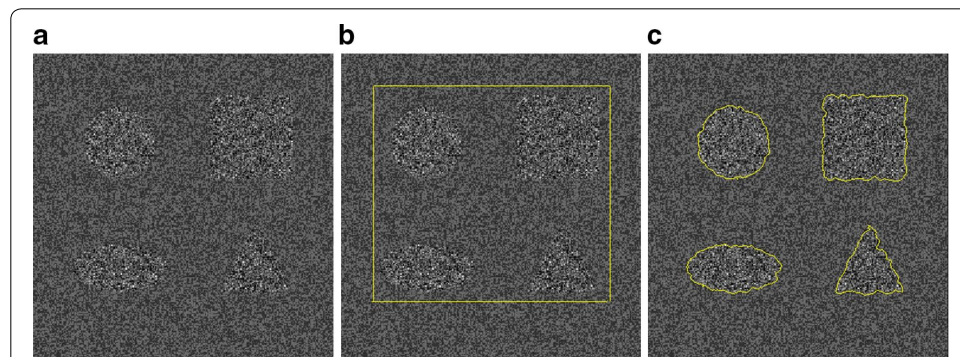
In order to validate the effect of the proposed segmentation method, the synthetic image is used to segmented. As shown in Fig. 10, the problem of boundary leakage at boundaries of objects with intensity inhomogeneity and typical weak edges, whose intensity feature is similar to that of GGO nodules, is solved by using the proposed segmentation method.

**Validation of 3D potential vessel segmentation by FDF-based region growing method**

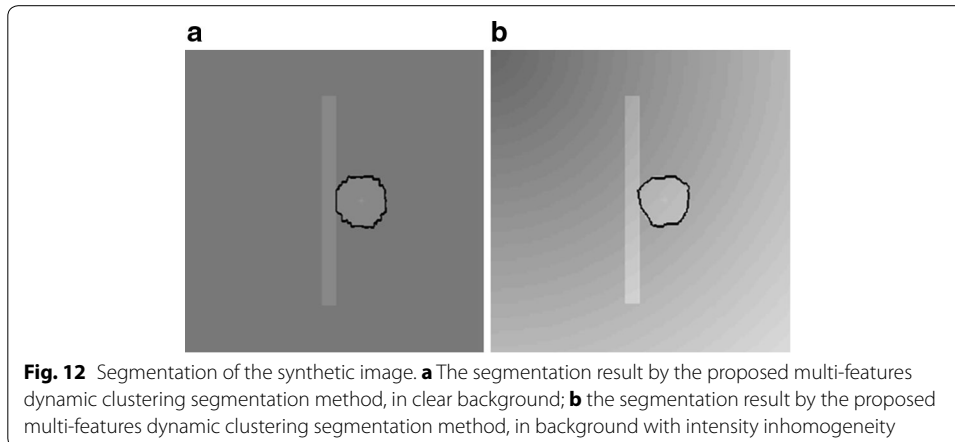
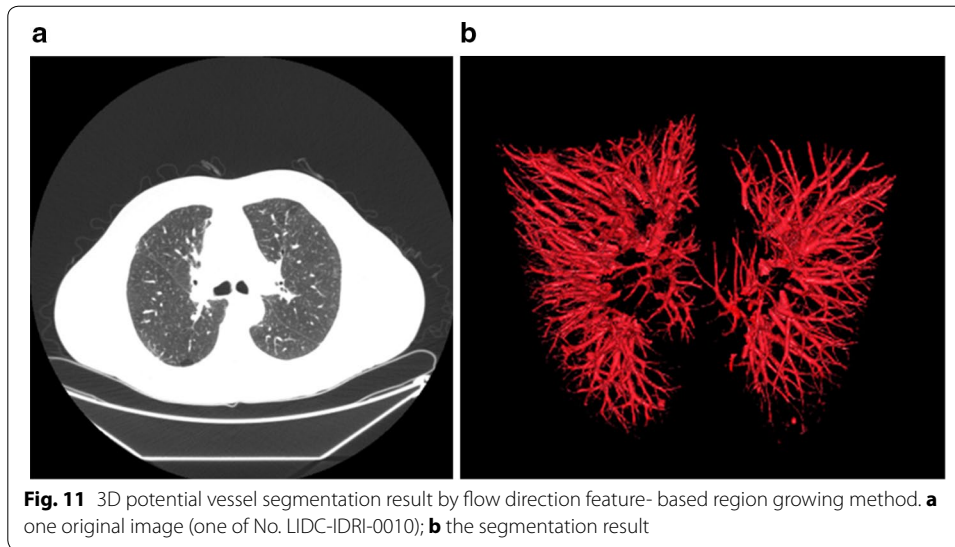
In order to validate the effect of the proposed flow direction feature -based region growing method for 3D potential vessel segmentation, No. LIDC-IDRI-0010 (from LIDC database [33]) is used to segmented. As shown in Fig. 11, 3D potential vessels are segmented successfully by the proposed flow direction feature- based region growing method. It is validated that the method can be used as the segmentation of 3D potential vessels and rough segmentation of juxta-vascular pulmonary nodules.

**Validation of multi-features dynamic clustering method**

In order to validate the effect of the proposed multi-features dynamic clustering segmentation method, the synthetic image is used to segment. As shown in Fig. 12, the tubular



**Fig. 10** Segmentation of the synthetic image. **a** The original image; **b** the initial contour; **c** the final segmentation



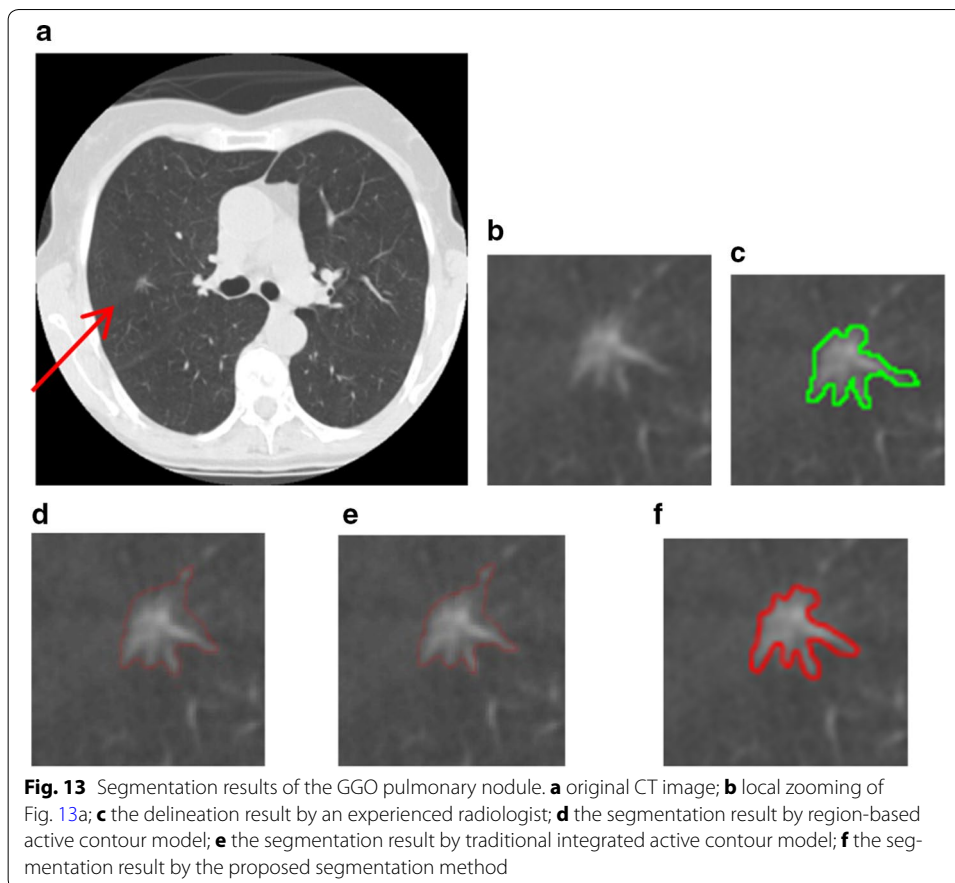
model and round or oval-shaped model can be segmented successfully. This implies the multi-features dynamic clustering method can be used for segmentation of juxta-vascular nodules, since vessels can be characterized by the tubular models and a pulmonary nodule is a small round or oval-shaped growth in the lung.

#### ***Validation of the proposed segmentation method of nodules***

In order to validate the effect of the proposed segmentation method, the clinical data with GGO nodules and juxta-vascular nodules should be segmented and explored.

A comparison for segmentation of GGO nodule between the proposed segmentation method and the traditional approach [34], e.g. region-based active contour model [6, 7] and integrated active contour model [13], is shown in Fig. 13. As shown in Fig. 13d, e, the problem of boundary leakage at boundaries of a GGO pulmonary nodule occurs, while the problem is solved in Fig. 13f.

In this paper, we handle various types of nodules under a united framework. Firstly the image sequences are segmented by using the proposed segmentation method. Then

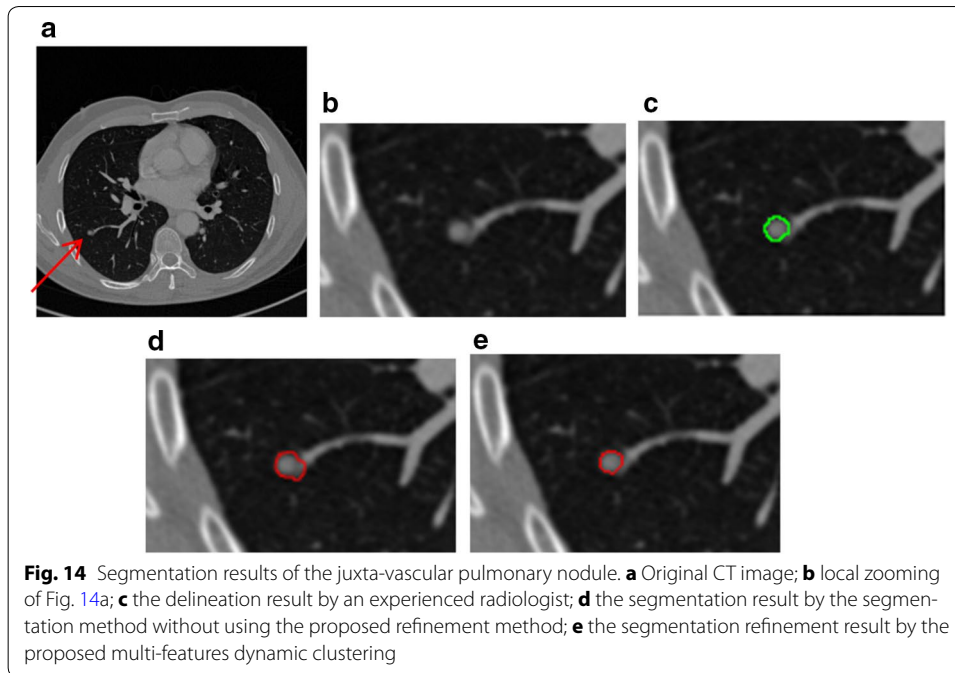


the juxta-vascular nodules are segmented finely by using segmentation refinement based multi-features dynamic clustering method. Some experimental results for segmentation of juxta-vascular pulmonary nodule are shown in Figs. 14 and 15. From Fig. 15, we can see that the refinement procession is only applied to voxels in some regions  $S_r$  containing the blood vessels and juxta-vascular nodules. The refinement method will not modify the nodule boundary elsewhere, so the problem of over-segmentation does not occur. It is validated that the method can be used as the segmentation of juxta-vascular pulmonary nodules.

Figure 16 shows the segmentation results of a GGO juxta-vascular pulmonary nodule using the proposed method, region-based ACM [6, 7] and the integrated ACM [13, 34], respectively. As shown in Fig. 16d, e, the problem of boundary leakage and over segmentation occurs in the adhesion place between the juxta-vascular nodule and its attached vessel; while as shown in Fig. 16f, the problem of boundary leakage and over-segmentation is solved by the proposed segmentation method.

From Figs. 13, 14, 15 and 16, the described segmentation method outperforms the traditional methods.

Other classical experimental results for segmentation of juxta-vascular pulmonary nodule are shown in Figs. 17 and 18. It is validated that the method can be used as the segmentation of various types of pulmonary nodules.



### Quantitative validation

Beyond the visual inspection, a quantitative analysis is necessary to ascertain the accuracy of the proposed segmentation method. Here, the well known Tanimoto/Jaccard error  $A(C_m, C_o)$  as Eq. (5) is used as the validation metrics, which refers to distances between segmentation results or to volume overlaps between the gold standard and the proposed segmentation method. The gold standard typically is a high-quality reference segmentation carried out by experts.

Segmentation measure results are shown in Table 5. Table 5 shows that the errors of the proposed method are less than the other three traditional methods, the edge-based active contour model, the traditional region-based active contour model and the integrated active contour model [6, 7, 13, 34].

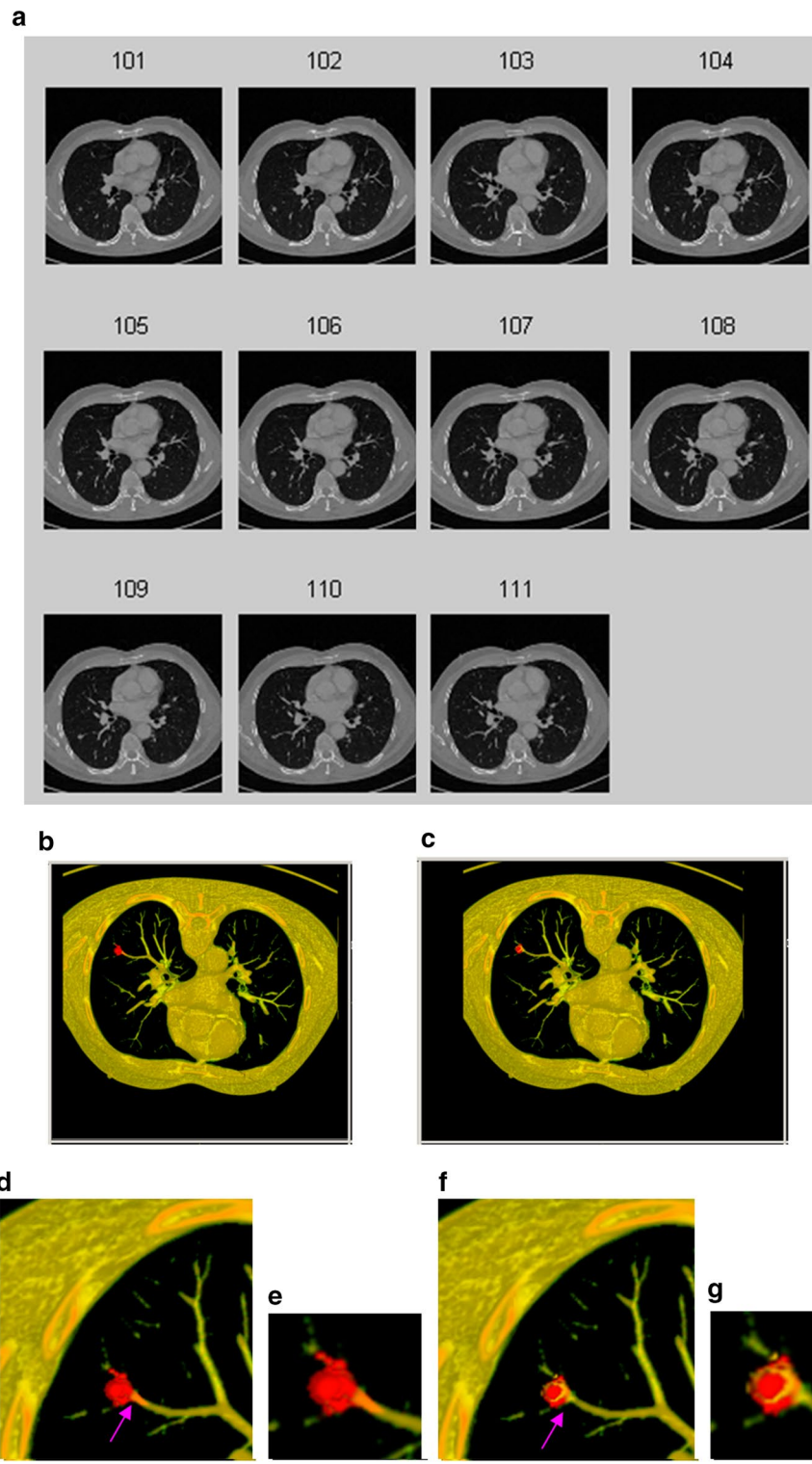
In our experiment, there are 3 nodules missed in the experimental test set of 157 nodules in total. Two of them are small GGO nodules close to 3 mm, and one of missing nodules is a small juxta-pleural nodule with very low contrast and close to 3 mm.

The proposed segmentation algorithm was implemented and tested on the computer with  $3.46 \times 2$  GHz CPU, 192 GB Memory and Graphic Card (GPU memory 12 GB GDDRS, 317 GB/s). On average, it takes about 3.14 min/scan (about 1.37 s/each image), which does not include the cost for data pre-processing.

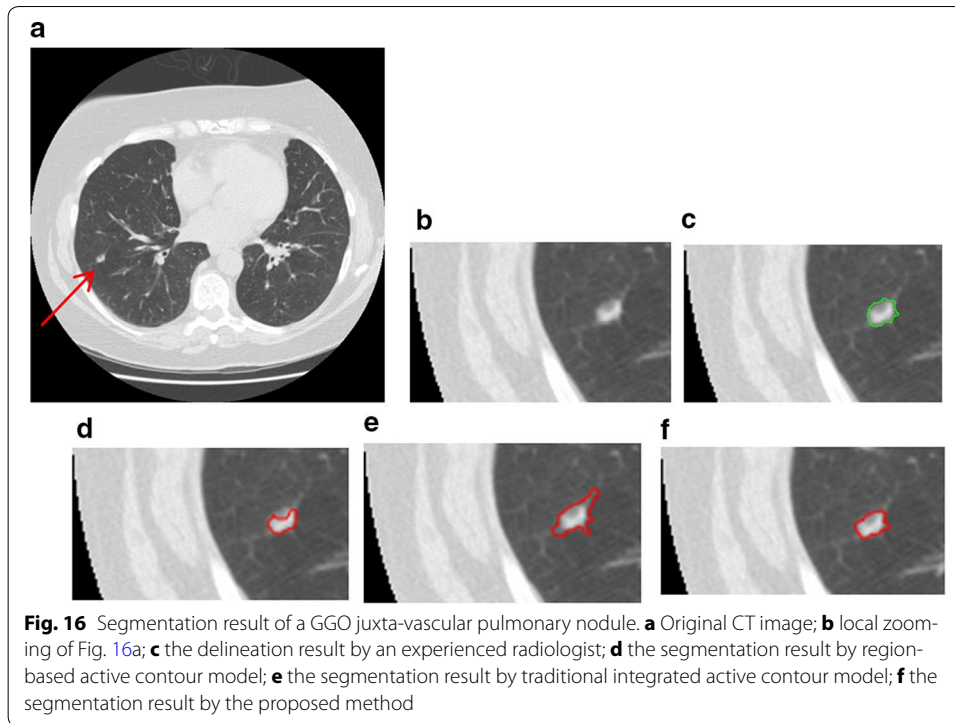
### Discussion

Experimental results of segmentation for pulmonary nodules show desirable performances of the proposed segmentation method using the test dataset. The segmentation performance for GGO, juxta-vascular and GGO juxta-vascular nodules was an average Tanimoto/Jaccard error of 0.17, 0.20 and 0.24, respectively.

We attempt a comparison with the results reported by other research groups. Kubota et al. [1] proposed a segmentation method based on morphological approaches and

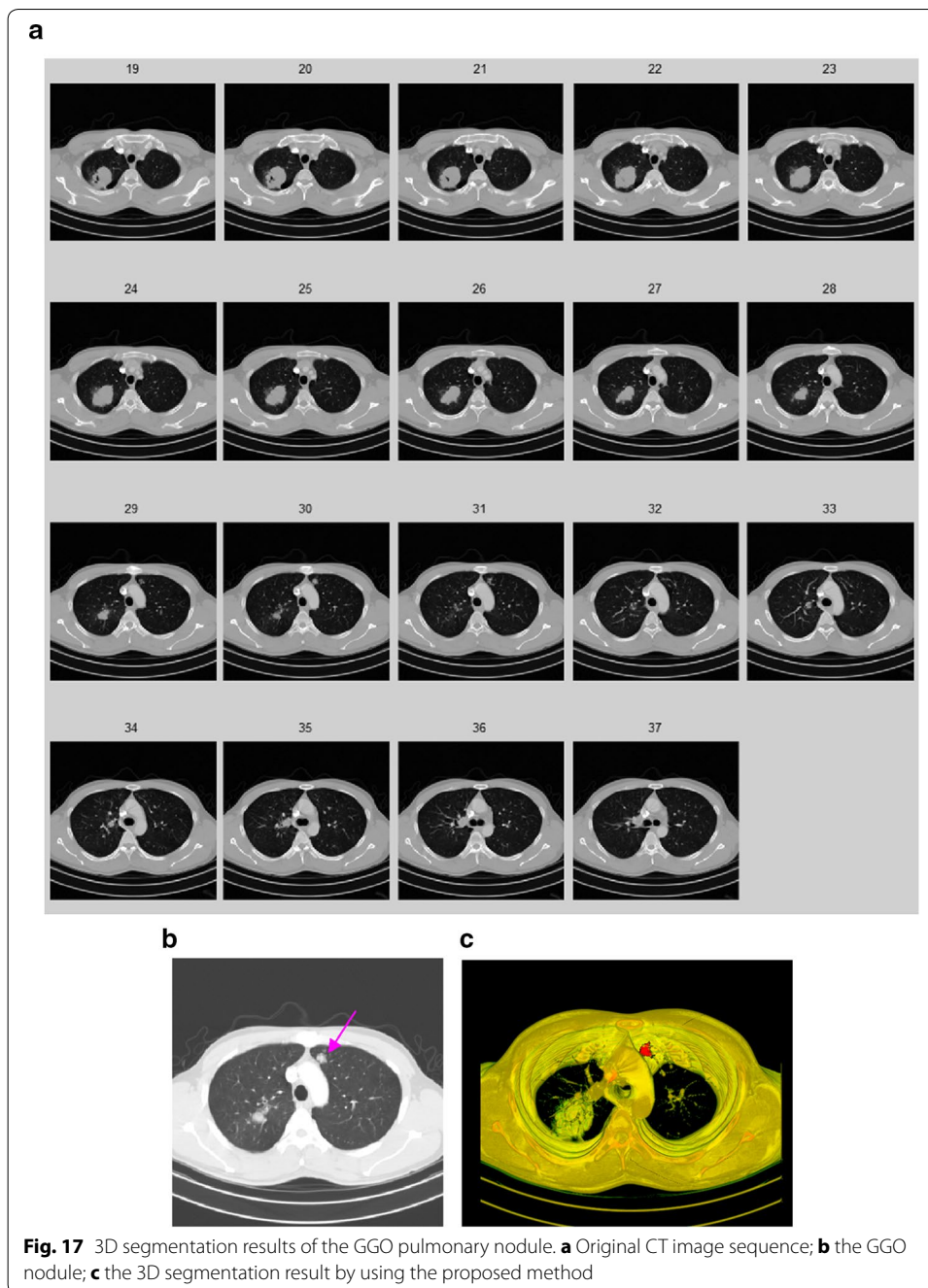


**Fig. 15** 3D segmentation results of the juxta-vascular pulmonary nodule. **a** original CT image sequence; **b** the 3D segmentation result without using the proposed refinement method; **c** the 3D segmentation result using the proposed segmentation method with multi-features dynamic clustering refinement; **d** local zooming of Fig. 15b; **e** local zooming of Fig. 15d; **f** local zooming of Fig. 15c; **g** local zooming of Fig. 15f



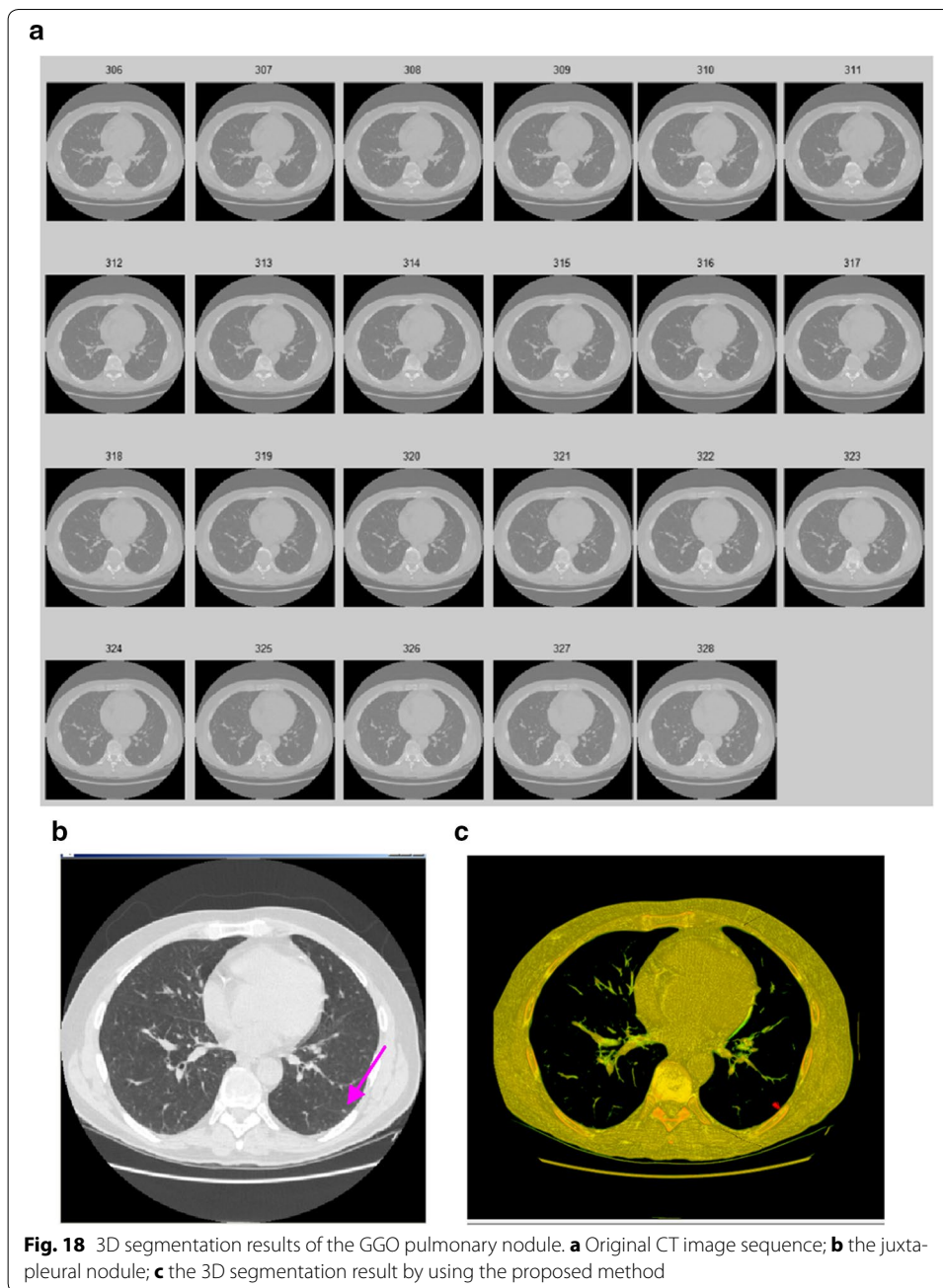
convexity models for segmenting the pulmonary nodules of various densities. Results on 21 LIDC cases were reported with segmentation overlap measures of mean 0.69 and standard deviation 0.18. Diciotti et al. [2], presented a refinement method for the segmentation of juxta-vascular nodules, which was based on a local shape analysis of the initial segmentation making use of geodesic distance map representations. They observed a percentage of successful segmentations of 84.8 % in fully automated mode and of 91.0 % by using an additional interactive mode for improving the segmentation quality of juxta-vascular nodules. However, GGO juxta-vascular nodules were not reported in their work. Kostis et al. [21] collected 21 juxta-vascular nodules, and observed an 80 % successful rate. Okada et al. [35] reported an 81.2 % estimation rate on a dataset of 1310 various types of nodules (3–30 mm in diameter). Unfortunately, though most of algorithms have been developed for lung nodules, most authors did not report quantitative results for various types of nodules. Ye et al. [5] proposed a shape-based SVM method for detecting nodules. The 3-D local geometric and statistical intensity features were used to detect potential solid and GGO nodule. But the segmentation results were not been reported. Murphy [32] used the local image features of shape index and curvedness to detect candidate structures in the lung volume, but the segmentation results were not been reported yet.

Comparing with different segmentation methods covered in literature [3] and other reported literatures above, it seems that the proposed method's relatively precise segmentation performances. The reasons why the proposed method has a better performance for segmenting all types of GGO and juxta-vascular nodules are as follows. In the proposed integrated ACM model, the local domain for local energy model is selected adaptively, and measurable distances between probability density functions of multi-dimension features with high class separability are used. The model using local



information commonly can obtain better performance than that of global statistic information in solving the segmentation problems of intensity inhomogeneity, such as part solid and nonsolid GGO nodules. Multi-dimension features are also important and helpful for the segmentation of GGO nodules. In many research papers, juxta-vascular nodules observed in CT images are outlined applying a global refinement procedure (i.e., throughout the initial segmentation boundary) after an initial rough segmentation. Differently, as is mentioned before, our solution for efficiently segmenting the potential nodule objects involves two steps: (1) a segmentation method is proposed for a whole





segmentation, which is based on the proposed integrated ACM model. The method uses an adaptive local region energy model with PDF-based similarity distance, which is especially used for low-contrast nodules such as part solid and nonsolid GGO nodules, to overcome the problems of boundary leakage, intensity inhomogeneity; (2) a segmentation refinement method based on multi-features dynamic clustering method, which is referred to as a fine segmentation, is used to segment potential juxta-vascular nodules. So the correction method has the advantage that it locally refines the nodule segmentation along recognized vessel attachments only, without modifying the nodule boundary elsewhere.

**Table 5 Segmentation measure results (error rate)**

CT image	The edge-based active contour model	The region-based active contour model	The traditional integrated active contour model	The proposed active contour model
GGO pulmonary nodule	Mean 0.31 Std. 0.11	Mean 0.36 Std. 0.13	Mean 0.36 Std. 0.09	Mean 0.17 Std. 0.07
Juxta-vascular pulmonary nodule	Mean 0.32 Std. 0.13	Mean 0.39 Std. 0.12	Mean 0.29 Std. 0.12	Mean 0.20 Std. 0.09
GGO Juxta-vascular nodule	Mean 0.43 Std. 0.12	Mean 0.47 Std. 0.09	Mean 0.41 Std. 0.13	Mean 0.24 Std. 0.08

Standard deviation (Std.)

"GGO nodule" and "Juxta-vascular nodule" do not include GGO Juxta-vascular pulmonary nodule in the table

However, some nodules are missed by the proposed segmentation method. Typically, these nodules are too small (almost 3 mm), or small juxta-pleural nodules with very low contrast, which makes it difficult to segment. The small GGO juxta-pleural nodule with pleural tail is very near to the edge of lung wall.

To further improve the segmentation performance, some improvements need to be further investigated as follows: in order to recognize small and juxta-pleural pulmonary nodules in noisy image more effectively, an adaptive smoothing method needs to be further investigated, and the juxta-pleural nodules should be further researched; This requires further investigated in more detail.

## Conclusions

When the traditional segmentation method is used to segment the GGO and juxta-vascular nodules with weak edges and intensity inhomogeneity characteristic, the problems of boundary leakage and small volume over-segmentation often appear. To solve these problems, a novel segmentation method is proposed for pulmonary nodules in CT images, which is based on the proposed integrated ACM model and multi-features dynamic clustering method, especially for GGO nodules (part solid and nonsolid) and juxta-vascular nodules. This study demonstrates the superiority of the proposed method. The described segmentation method outperforms the traditional methods, and evaluating the algorithm on the provided test data leads to an average Tanimoto/Jaccard error of 0.17, 0.20 and 0.24 for GGO, juxta-vascular and GGO juxta-vascular nodules, respectively.

### Authors' contributions

BL designed the study and carried out the whole proposed segmentation method. QC studies the local region energy model. GP, YG, SO and LW participated in data collection and constituted the members of a qualified panel. KC and LF participated in the statistical analysis. All authors read and approved the final manuscript.

### Author details

<sup>1</sup> School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, Guangdong, China. <sup>2</sup> Department of Radiology, Guangzhou General Hospital of Guangzhou Command, Guangzhou 510010, Guangdong, China. <sup>3</sup> Department of Radiology, Shenzhen Third People's Hospital, Shenzhen 518112, Guangdong, China.

### Acknowledgements

The authors would like to thank Dr. W.B. Zhu, Dr. P. Chen, Dr. R. Bai, Dr. L. Zhang, Dr. F. Long, Radiologist G.Q. Qiao, and Engineer L. Tang for their helpful comments and advice which contributed much to this paper. This work is supported by National Natural Science Foundation of China (61305038, 61273249), the Public Science and Technology Research Funds Projects of Ocean (201505002), the Fundamental Research Funds for the Central Universities, SCUT (No. 2015ZZ028), Key Laboratory of Autonomous Systems and Network Control of Ministry of Education (SCUT of China), the National Engineering Research Center for Tissue Restoration and Reconstruction and the Guangdong Key Laboratory for Biomedical Engineering (SCUT of China).

### Sources of support

This work is supported by National Natural Science Foundation of China (61305038, 61273249), the Public Science and Technology Research Funds Projects of Ocean (201505002), the Fundamental Research Funds for the Central Universities, SCUT (No.2015ZZ028), Key Laboratory of Autonomous Systems and Network Control of Ministry of Education (SCUT of China), the National Engineering Research Center for Tissue Restoration and Reconstruction and the Guangdong Key Laboratory for Biomedical Engineering (SCUT of China).

### Competing interests

The authors declare that they have no competing interests.

Received: 10 January 2016 Accepted: 25 April 2016

Published online: 05 May 2016

### References

1. Kubota T, Jerebko AK, Dewan M, Salganicoff M, Krishnan A. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Med Image Anal.* 2011;15:133–54.
2. Diciotti S, Lombardo S, Falchini M, Picozzi G, Mascalchi M. Automated segmentation refinement of small lung nodules in CT scans by local shape analysis. *IEEE Trans Biomed Eng.* 2011;58(12):3418–28.
3. Sluimer I, Schilham A, Prokop M, van Ginneken B. Computer analysis of computed tomography scans of the lung: a survey. *IEEE Trans Med Imaging.* 2006;25(4):385–405.
4. Valente IRS, Cortez PC, Neto EC, Soares JM, de Albuquerque VHC, Tavares JMRS. Automatic 3D pulmonary nodule detection in CT images: a survey. *Comput Methods Programs Biomed.* 2016;124:91–107.
5. Ye XJ, Lin XY, Dehmeshki J, Slabaugh G, Beddoe G. Shape-based computer-aided detection of lung nodules in thoracic CT images. *IEEE Trans Biomed Eng.* 2009;56(7):1810–20.
6. Chan T, Vese L. Active contours without edges. *IEEE Trans Image Process.* 2001;10(2):266–77.
7. Li CM, Kao CY, Gore JC, Ding Z. Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans Image Process.* 2008;17(10):1940–9.
8. Ma Z, Tavares JMRS, Jorge RMN. "A review on the current segmentation algorithms for medical images," In: 1st international conference on imaging theory and applications (IMAGAPP). ISBN: 978-989-8111-68-5, Portugal; 2009. p. 135–140.
9. Ma Z, Tavares JMRS, Jorge RMN, Mascarenhas T. A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. *Comp Methods Biomech Biomed Eng.* 2010;13(2):235–46.
10. Tavares JMRS. "Image processing and analysis: applications and trends," In: AES-ATEMA'2010 fifth international conference on advances and trends in engineering materials and their applications. ISBN:978-0-9780479-7-9(CD-ROM)-978-0-9780479-9-3 (Hard Copy), Canada; 2010. p. 27–41
11. Way T, Hadjiliski L, Sahiner B, Chan HP, Cascade PN, Kazerooni EA, Bogot N, Zhou C. Computer-aided diagnosis of pulmonary nodules on CT scans: segmentation and classification using 3D active contours. *Med Phys.* 2006;33(7):2323–37.
12. Tao WB. Iterative narrowband-based graph cuts optimization for geodesic active contours with region forces (GACWRF). *IEEE Trans Image Process.* 2012;21(1):284–96.
13. Chen S, Sochen NA, Zeevi YY. Integrated active contours for texture segmentation. *IEEE Trans Image Process.* 2004;1(1):1–18.
14. Li B, Chen K, Tian L, Yeboah Y, Ou S. Detection of pulmonary nodules in CT images based on fuzzy integrated active contour model and hybrid parametric mixture model. *Comput Math Methods Med.* 2013;2013:515386.
15. Ni K, Bresson X, Chan T, Eshedoglu S. Local histogram based segmentation using the wasserstein distance. *Int J Comput Vision.* 2009;84(1):97–111.
16. Lankton S, Tannenbaum A. Localizing region-based active contours. *IEEE Trans Image Process.* 2008;17(11):2029–39.
17. Wu H, Appica V, Anthony Y. Numerical conditioning problems and solutions for nonparametric statistical active contours. *IEEE Trans Pattern Anal Mach Intell.* 2013;35(6):1298–311.
18. Krinidis S, Chatzis V. Fuzzy energy-based active contours. *IEEE Trans Image Process.* 2009;18(12):2747–55.
19. Assen HC, Danilouchkine MG, Dirksen MS, Reiber J, Lelieveldt BPF. A 3-D active shape model driven by fuzzy inference: application to cardiac CT and MR. *IEEE Trans Inf Technol Biomed.* 2008;12(5):595–605.
20. Zhang Y, Li GY, Xiehua S, Xinmin Z. Geometric active contours without re-initialization for image segmentation. *Pattern Recogn.* 2009;42(9):1970–6.
21. Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI. Three-D segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. *IEEE Trans Med Imaging.* 2003;22(10):1259–74.
22. Kuhnigk JM, Dicken V, Bornemann L, Bakai A, Wormanns D, Krass S, Peitgen HO. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. *IEEE Trans Med Imaging.* 2006;25(4):417–34.
23. Lesage D, Angelini ED, Bloch I, Funka-Lea G. A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes. *Med Image Anal.* 2009;13:819–45.
24. Bauer C, Pock T, Sorantin E, Bischof H, Beichel R. Segmentation of interwoven 3D tubular tree structures utilizing shape priors and graph cuts. *Med Image Anal.* 2010;14:172–84.
25. Chen K, Li B, Tian L, Zhu W, Bao Y. Vessel attachment nodule segmentation using integrated active contour model based on fuzzy speed function and shape-intensity joint Bhattacharya distance. *Sig Process.* 2014;103:273–84.
26. Zhu W, Li B, Tian L, Li X, Chen Q. Topology adaptive vessel network skeleton extraction with novel medialness measuring function. *Comput Biol Med.* 2015;64:40–61.
27. Freedman D, Zhang T. Active contours for tracking distributions. *IEEE Trans Image Process.* 2004;13(4):518–26.

28. Michailovich O, Rathi Y, Tannenbaum A. Image segmentation using active contours driven by the Bhattacharyya gradient flow. *IEEE Trans Image Process.* 2007;16:2787–801.
29. Agam G, Armato SG, Changhua W. Vessel tree reconstruction in thoracic CT scans with application to nodule detection. *IEEE Trans Med Imaging.* 2005;24(4):486–99.
30. Frangi AF, Niessen WJ, Vincken KL, Viergever MA. Multiscale vessel enhancement filtering. *Medical image computing and computer-assisted intervention—MICCAI'98.* vol. 1496 of the series lecture notes in Computer science, p. 130–137, June 2006.
31. Zana F, Klein JC. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Trans Image Process.* 2001;10(7):1010–9.
32. Murphy K, van Ginneken B, Schilham AMR, de Hoop BJ, Gietema HA, Prokop M. "A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification". *Med Image Anal.* 2009;13(5):757–70.
33. Reeves AP, Biancardi AM, Apanasovich TV, Meyer CR, MacMahon H, Beek EJR, Kazerooni EA, Yankelevitz D, Gray MFM, McLennan G, Armato SG, Henschke CI, Aberle DR, Croft BY. The Lung Image Database Consortium (LIDC): a comparison of different size metrics for pulmonary nodule measurements. *Acad Radiol.* 2007;14(12):1475–85.
34. Dietenbeck T, Alessandrini M, Friboulet D, Bernard O. "CREASEG: a free software for the evaluation of image segmentation algorithms based on level-set." In: *IEEE international conference on image processing.* Hong Kong; 2010.
35. Okada K, Comaniciu D, Krishnan A. Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT. *IEEE Trans Med Imaging.* 2005;24(3):409–23.

Submit your next manuscript to BioMed Central  
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

