An unsupervised metaheuristic search approach for segmentation and volume measurement of pulmonary nodules in lung CT scans

Elham Shakibapour [a,*], António Cunha [a,b], Guilherme Aresta [a,c], Ana Maria Mendonça [a,c], Aurélio Campilho [a,c]

[a] *INESC TEC – INESC Technology and Science. FEUP Campus, Dr. Roberto Frias 4200 - 465, Porto, Portugal*
[b] *Universidade de Trás-os-Montes e Alto Douro. Quinta de Prados 5001-801, Vila Real, Portugal*
[c] *Faculdade de Engenharia, Universidade do Porto. FEUP Campus, Dr. Roberto Frias 4200 - 465, Porto, Portugal*

A R T I C L E   I N F O

A B S T R A C T

This paper proposes a new methodology to automatically segment and measure the volume of pulmonary nodules in lung computed tomography (CT) scans. Estimating the malignancy likelihood of a pulmonary nodule based on lesion characteristics motivated the development of an unsupervised pulmonary nodule segmentation and volume measurement as a preliminary stage for pulmonary nodule characterization. The idea is to optimally cluster a set of feature vectors composed by intensity and shape-related features in a given feature data space extracted from a pre-detected nodule. For that purpose, a metaheuristic search based on evolutionary computation is used for clustering the corresponding feature vectors. The proposed method is simple, unsupervised and is able to segment different types of nodules in terms of location and texture without the need for any manual annotation. We validate the proposed segmentation and volume measurement on the Lung Image Database Consortium and Image Database Resource Initiative – LIDC-IDRI dataset. The first dataset is a group of 705 solid and sub-solid (assessed as part-solid and non-solid) nodules located in different regions of the lungs, and the second, more challenging, is a group of 59 sub-solid nodules. The average Dice scores of 82.35% and 71.05% for the two datasets show the good performance of the segmentation proposal. Comparisons with previous state-of-the-art techniques also show acceptable and comparable segmentation results. The volumes of the segmented nodules are measured via ellipsoid approximation. The correlation and statistical significance between the measured volumes of the segmented nodules and the ground-truth are obtained by Pearson correlation coefficient value, obtaining an *R*-value $\geq$ 92.16% with a significance level of 5%.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Lung cancer is the deadliest type of cancer in both men and women (Stewart & Wild, 2014) and the early detection and diagnosis of lung masses, lung lesions or pulmonary nodules, is fundamental to improve the survival rate of patients and to improve the quality of the worldwide health.

Pulmonary nodules are detectable on chest X-ray and computed tomography (CT) scans (Tsao, 2007). CT produces a volume of lung structures captured from different angles. Modern scanners allow a lung volume to be reformatted in 3 major orthogonal anatomical planes: axial, sagittal, or coronal planes or even as a volumetric 3D representation. Lung computer-aided diagnosis (CAD) systems aim to assist expert radiologists in the analysis and interpretation of pulmonary nodules during lung screening processes.

Pulmonary nodule detection is the first important task in a lung CAD system. After nodule detection, follows nodule segmentation and characterization, to measure the nodule volume or characterize the nodule morphology. These are the major steps to estimate the nodule malignancy likelihood.

Pulmonary nodules with diverse shapes and lesion characteristics (size, lobulation, subtlety, spiculation, etc.) can be characterized in terms of location (isolated, juxta-vascular, pleural-tail and juxta-pleural nodules) and texture-wise (solid and sub-solid nodules). Sub-solid nodules are further divided into part-solid and non-solid - Ground Glass Opacity (GGO) nodules. Fig. 1 illustrates examples of some nodules with varied textures and appearances of the Lung Image Database Consortium and Image Database Resource Initiative – LIDC-IDRI dataset – (Armato et al., 2011).

Fig. 1 illustrates the large variability of cases that can occur, which makes more difficult the design of an accurate and robust

---

* Corresponding author.

E-mail addresses: elhamshakibapour@miau.ac.ir (E. Shakibapour), acunha@utad.pt (A. Cunha), guilherme.m.aresta@inesctec.pt (G. Aresta), amendon@fe.up.pt (A.M. Mendonça), campilho@fe.up.pt (A. Campilho).

(a) Isolated solid, part-solid and non-solid nodules

(b) Juxta-vascular solid, part-solid and non-solid nodules

(c) Pleural-tail solid, part-solid and non-solid nodules

(d) Juxta-pleural solid, part-solid and non-solid nodules
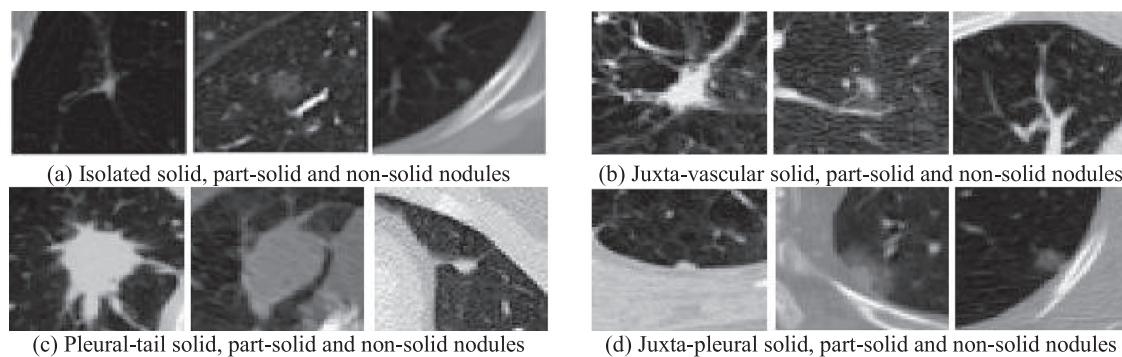
Fig. 1. Examples of pulmonary nodules of the LIDC-IDRI dataset (Armato et al., 2011).

segmentation task. An isolated nodule is the easiest to segment since it does not adhere to any lung structure such as vessel, fissure or chest wall (Fig. 1(a)). Segmentation of nodules attached to other structures (Fig. 1(b)–(d)) demands for an approach to segment them from the lung tissue. The size of nodules has a great range of variation. Their shape can be very irregular and their contrast also has a large variability, which may difficult the nodule discrimination from the lung tissue. The process of segmentation is also affected by the nodule texture diversity. Segmentation approaches often achieve a good performance on segmenting solid nodules. Part-solid nodules, which are more likely to be malignant (Tu et al., 2017), have characteristics of both solid and non-solid nodules. Such variability does no ease the segmentation task, resulting in many cases in less accurate results. Segmentation of non-solid nodules remains challenging because of shape variability with fuzzy contours and low contrast.

In a lung CAD system, an accurate segmentation is important to define correctly the nodule boundaries, and therefore allow a good estimate of the nodule volume, which is a relevant feature for lung cancer screening (Devaraj, van Ginneken, Nair, & Baldwin, 2017). It also assists radiologists predicting the degree of the nodule malignancy and defining the lesion response to therapy (McKee, Regis, McKee, Flacke, & Wald, 2015). We conceive the segmentation task as an unsupervised learning methodology and as a process of knowledge discovery through evolutionary computation. Unsupervised, to avoid the need of previous nodule annotation, which is a costly operation, and usually hard to obtain for a reasonable number of representative samples. Evolutionary computation, to optimize the search of the clusters inspired by biological evolution. Evolutionary algorithms approximate solutions by intelligently evolving the information with a metaheuristic or stochastic approaches. Evolutionary algorithms ideally have a good performance with the least knowledge. They start by initializing a set of parameters and attempt to optimize the series of the solutions and convert them to the new series such that it is confident that the new obtained outcomes are better than the previous ones.

The proposed unsupervised metaheuristic search via an evolutionary algorithm is schematized in Fig. 2. The goal is to optimally cluster a set of pixels characterized by feature vectors, composed by intensity and shape-related features in a given feature data space using a 2.5D representation. The feature vectors are extracted from a pre-detected pulmonary nodule, which is centered on the nodule's centroid. The pixels belonging to the same cluster of the nodule's centroid is the segmented region output. The nodule volume is measured via ellipsoid approximation, from a 2.5D nodule representation, using the 3 major orthogonal anatomical nodular planes: axial, sagittal and coronal planes of the middle slice.

In general, the advantages of the nodule segmentation and volume measurement schemes proposed in this work are four fold:

1) An unsupervised learning strategy helps a segmentation application to work with a large variety of inputs without any user-intervention for image annotation.
2) A metaheuristic search has the potential in revealing an outcome in a short fraction of time.
3) A metaheuristic search thru an evolutionary algorithm has the ability to search for knowledge by evolving their existing information intelligently.
4) In the proposed approach, the nodule size is obtained by measuring the volume of a segmented nodule via ellipsoid approximation using the equivalent diameters of the segmented regions in a 2.5D representation, not demanding for a full use of the CT slices in the nodule 3D volume.

The paper is organized as follows. Section 2 briefly surveys methods attempted to segment pulmonary nodules. Details of the proposed pulmonary nodule segmentation and volume measurement are given in Section 3. Section 4 provides the nodule segmentation and volume measurement results with two datasets from the LIDC-IDRI dataset. The performance of the segmentation method is discussed in Section 5. A comparative analysis of the proposed methodology with the state-of-the-art techniques and details on the performance of the evolutionary algorithms exploited in this paper are also described in Section 5. Section 6 concludes the paper.

2. Related work

A number of approaches attempted to first localize pulmonary nodules in lung CT scans and then segment the detected nodules (Gu et al., 2013; Keshani, Azimifar, Tajeripour, & Boostani, 2013; Song et al., 2016; Sun et al., 2014). In the work of Sun et al. (2014), nodules are first detected, followed by the segmentation of juxta-vascular nodules by extracting flow entropy and geodesic distance to distinguish the nodular area from vessel segments, using a k-means clustering approach. Keshani et al. (2013) employed active contours to extract the boundaries of the detected nodules followed by SVM classification to recognize the segmented nodules. Segmentation of juxta-pleural nodules using a single click and a region growing algorithm was provided by Gu et al. (2013). Song et al. (2016) proposed an automatic detection of initial seed points by an improved toboggan method. An iterative region growing approach with multi-constraints was applied for nodule segmentation.

3D segmentation of pulmonary nodules can be divided into two main groups, semi-automatic and automatic. In the first group, interactive segmentation is performed (Awad et al., 2012; Farag, Abd, Munim, Graham, & Farag, 2013; Kubota, Jerebko, Dewan, Salganicoff, & Krishnan, 2011; Lassen, Jacobs, Kuhnigk, van Ginneken, & van Rikxoort, 2015; Messay et al., 2015), while in the second one there is no user-intervention
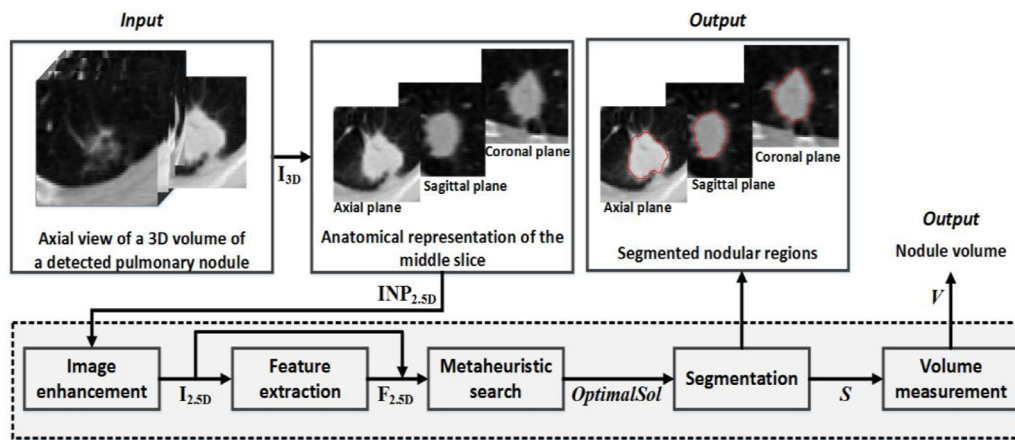
Fig. 2. An overview of the proposed unsupervised pulmonary nodule segmentation and volume measurement schemes.

(Gonçalves, Novo, & Campilho, 2016; Jung, Hong, & Goo, 2018; Liu, Geng, Guo, Zhang, & Zhang, 2017; Wang et al., 2017; Ye, Beddoe, & Slabaugh, 2010; Zhao, Ji, Xia, & Zhang, 2015).

Semi-automatic and automatic methodologies have been designed by thresholding and morphological methods (Messay, Hardie, & Tuinstra, 2015; Mukhopadhyay, 2016), region-growing approaches (Kalpathy-Cramer et al., 2016; Kubota et al., 2011; Lassen et al., 2015; Moltz et al., 2009; Zhao et al., 2015), deformable and level-set methods (Farag et al., 2013; Jung et al., 2018), graph-cut algorithms (Ye et al., 2010), shape-based segmentation strategies (Awad et al., 2012; Gonçalves et al., 2016), and unsupervised and supervised approaches (Liu et al., 2017; Wang et al., 2017).

A semi-automated system for 3D segmentation of solid and sub-solid nodules based on parameters estimation for thresholding and morphological operations using a regression neural network was developed in Messay et al. (2015). Kubota et al. (2011) proposed a solid and sub-solid nodule localization method using a manually selected starting point as a seed for 3D region growing on an Euclidean distance map. A convex hull was used for separating the attached structures. Lassen et al. (2015) implemented a 3D method starting by defining a region of interest around a nodule using an user-drawn stroke. Based on intensity analysis of the nodule region and the surrounding parenchyma, a threshold-based region growing approach was performed for sub-solid nodules. A 2D level-set method was used in Farag et al. (2013) to model the nodule image information. In a shape-based segmentation framework, the nodule center was marked and an ellipse was adapted to define the nodular region.

Zhao et al. (2015) used a region-growing method to roughly segment an image. A self-generating neural network and particular swarm optimization algorithms were used to cluster the nodular regions, particularly for 2D segmentation of cavitary nodules, without user-interaction. Ye et al. (2010) applied a 3D graph-cut segmentation algorithm using a new energy formulation. The method is automatic, where foreground and background seeds were obtained by considering spherical nodules. An important drawback is the failing of segmentation for some sub-solid nodules. Jung et al. (2018) proposed a 3D automatic segmentation of sub-solid nodules by extracting the non-solid component from the solid one using histogram modeling. Asymmetric multi-phase deformable model and pulmonary vessel removal were used in the segmentation process.

Conventional methods such as region-growing approaches, energy optimization algorithms like level-set based techniques or deformable models along with graph-cut algorithms involve the definition of convergence conditions. The convergence criteria can be very sensitive especially for segmentation of irregular-shaped nodules, as the sub-solid nodules (Wang et al., 2017). These methods often involve the selection of a threshold or a morphological template, in many cases using ad-hoc parameters.

A number of researchers worked on shape-based segmentation strategies. In Awad et al. (2012), solid nodules were segmented using a 3D based user-interaction approach through shape constrained Otsu multi-thresholding and sparse field active surface techniques. Gonçalves et al. (2016) proposed a 3D automatic segmentation of nodules using Hessian-based approaches, however it failed for some sub-solid nodules.

Unsupervised nodule segmentation was proposed by Liu et al. (2017), using a modified self-adaptive fuzzy C-means clustering algorithm for 2D automatic segmentation. This approach was not able to correctly segment non-solid and juxta-pleural nodules as well as small nodules of diameter less than 10.0 mm. Recently, Wang et al. (2017) proposed a data-driven model, the so-called central focused convolutional neural networks (CF-CNN) for 3D segmentation of nodules. The CF-CNN model demonstrated an ability to learn nodule-sensitive features automatically when trained on 0.41 million generated voxel patches. The performance of the CF-CNN model was assessed in terms of lesion characteristics (as size, subtlety, spiculation, among others) but not in terms of texture.

After segmentation, a lung CAD system should provide the nodule size, as the nodule volume, a major requirement in screening environments. Han, Heuvelmans, and Oudkerk (2017) compared semi-automated volume and diameter measurements of nodules. In Gavrielides, Kinnard, Myers, and Petrick (2009), volumetric analysis of nodules was reviewed. Simple voxel counting for nodule volume measurement was used by Awad et al. (2012), Diciotti, Lombardo, Falchini, Picozzi, and Mascalchi (2011), Heckel et al. (2014), and Kalpathy-Cramer et al. (2016). Volume measurement performance depends on image acquisition, nodule characteristics and importantly the performance on the nodule segmentation phase (Gavrielides et al., 2009).

In the study addressed in this article, we propose an unsupervised scheme able to segment nodules located in different regions of the lungs, robust to different lesion characteristics, including solid and sub-solid nodules. Our proposal has the ability to blindly segment nodules without any user-interaction. The nodule volume measurement method is approached with an ellipsoid approximation.
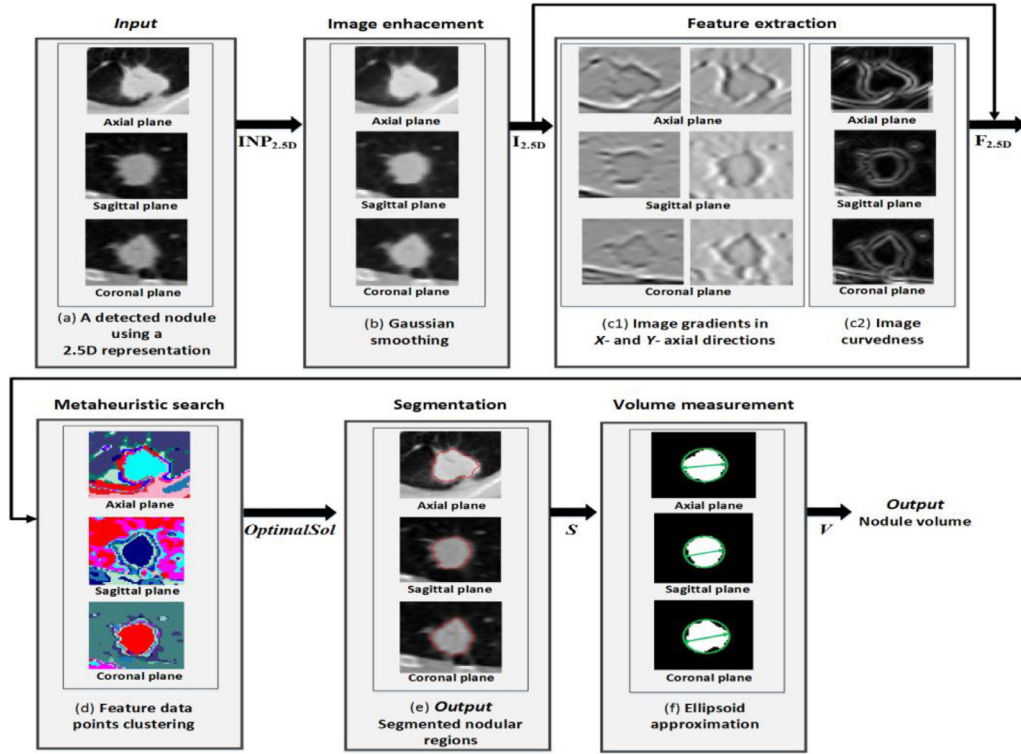
Fig. 3. Unsupervised pulmonary nodule segmentation and volume measurement schemes. (a) Detected nodule using a 2.5D representation; (b) Nodule image enhancement; (c) Nodule feature extraction; (d) Metaheuristic search for clustering; (e) Nodule segmentation; (f) Nodule volume measurement.

3. Unsupervised segmentation and volume measurement

The unsupervised segmentation task is developed by clustering a set of feature vectors characterizing the pixels belonging to the nodular region. Generally, in the clustering optimization process, each data point in a given feature data space is grouped around the closest cluster center, with a goal that maximizes the intra-cluster and minimizes the inter-cluster similarity, by optimally updating the position of the cluster centers on the feature space. In this paper, clustering is treated as a multi-objective optimization problem formulated by a metaheuristic search via an evolutionary algorithm. Fig. 3 gives an overview of the proposed unsupervised pulmonary nodule segmentation and volume measurement schemes. The input of the system is a 2.5D representation of a pre-detected nodule, which is centered on the nodule's centroid, with the 2D images corresponding to the axial, sagittal and coronal nodular planes (Fig. 3(a)). The output is the nodule segmentation of these 3 nodular planes and an estimation of the segmented nodule volume. The main phases of the approach are the following:

Step 1 **(Enhancement):** Given a 2.5D representation of a nodule, the 3 images of the nodular planes are smoothed using a Gaussian filter (Fig. 3(b)).

Step 2 **(Feature extraction):** The enhanced images are the input to the feature extraction phase, involving the computation of edge (Fig. 3(c1)) and curvature features (Fig. 3(c2)). The output is a matrix of feature vectors characterizing every pixel of each one of the 3 nodular planes. The corresponding details are given in Section 3.1.

Step 3 **(Segmentation by clustering using a metaheuristic search approach):** For each one of the nodular planes, different groups of feature vectors are randomly selected as the initial sets of cluster centers candidates. Each set of the selected feature vectors represents a candidate solution of the looked-for

optimum cluster centers. The feature vectors per group are iteratively evolved and updated by a metaheuristic search using an evolutionary algorithm. The best looked-for optimum candidate solution and the corresponding clusters of the feature vectors are the results of this phase (Fig. 3(d)). Thereafter, the nodule segmentation is performed by obtaining the positions of the pixels' feature vectors having the same cluster of the pixel's feature vector located in the centroid of each plane (Fig. 3(e)). Section 3.2 explains the details of this step.

Step 4 **(Volume measurement):** After segmentation, the nodule volume is estimated from the segmentations obtained for the 3 nodular planes (Fig. 3(f)), as described in Section 3.3.

3.1. Feature extraction

Firstly, the 3 nodular planes of a pre-detected nodule, are organized in $\mathbf{INP}_{2.5D}$ of size $N \times N \times 3$. This 3-valued image is pre-processed for reducing noise and smoothing the nodular surface. Each plane of $\mathbf{INP}_{2.5D}$ is convolved with a 2D Gaussian smoothing function (G) with standard deviation (σ). The output, $\mathbf{I}_{2.5D}$, given by Eq. (1), is the enhanced smoothed image.

$$\mathbf{I}_{2.5D} = \mathbf{INP}_{2.5D} * G(\sigma) \tag{1}$$

We use an Hessian matrix to measure the shape-related features, describing the nodule edgeness and local curvedness. The Hessian is the second order partial derivatives of an image. The Hessian matrix for $\mathbf{I}_{2.5D}$ at a pixel (x, y), of each plane, is defined by,

$$\mathbf{H} = \begin{bmatrix} \dfrac{\partial^2 \mathbf{I}_{2.5D}(x, y)}{\partial x^2} & \dfrac{\partial^2 \mathbf{I}_{2.5D}(x, y)}{\partial x \partial y} \\ \dfrac{\partial^2 \mathbf{I}_{2.5D}(x, y)}{\partial y \partial x} & \dfrac{\partial^2 \mathbf{I}_{2.5D}(x, y)}{\partial y^2} \end{bmatrix} \tag{2}$$

The positive curvedness value *CV* at a given pixel, of each plane, is given by,

$$CV = \sqrt{\lambda_1^2 + \lambda_2^2} \qquad (3)$$

where, λ_1 and λ_2 are the eigenvalues of **H** at a given pixel, of each one of the 3 nodular planes.

At a pixel (x, y) from each one of the 3 nodular planes, we organize the m-dimensional feature vector, in here $m = 4$, with the following characteristics:

- the first element of the vector is the image intensity at (x, y), $f_1 = \mathbf{I}_{2.5D}(x, y)$.
- the second element of the vector is $f_2 = \frac{\partial^2 \mathbf{I}_{2.5D}(x, y)}{\partial x^2}$, the first diagonal element of **H.**
- the third element is $f_3 = \frac{\partial^2 \mathbf{I}_{2.5D}(x, y)}{\partial y^2}$, the second diagonal element of **H.**
- the forth element is $f_4 = CV$.

Finally, a feature matrix for the pixels of the 3 nodular planes, $\mathbf{F}_{2.5D}$, is organized where each row of $\mathbf{F}_{2.5D}$ is the m-valued vector defined above. The number of rows of $\mathbf{F}_{2.5D}$ is the number of pixels in each image plane, equal to $N \times N$.

3.2. Segmentation by clustering using a metaheuristic search approach

The input of the segmentation algorithm is the 3 nodular planes characterized by the m-valued feature vectors per each one of the 3 nodular planes. For segmentation, we cluster the nodule feature space $\mathbf{F}_{2.5D}$ into k number of optimum clusters via a metaheuristic search using an evolutionary algorithm. In this study, we consider three evolutionary algorithms: a firefly algorithm (FA), a genetic algorithm (GA) and a teaching learning based optimization (TLBO) algorithm (see Table 1 for a brief description) to analyze their performance. The GA is one of the first most common evolutionary algorithms applied in numerous different applications. The FA and TLBO algorithms are recent effective evolutionary algorithms based on knowledge discovery among the candidate solutions. These algorithms are significantly less dependent on initialization settings and allow visiting multiple local optima by an iterative evolving process to achieve one global optimum.

Generally, in an evolutionary algorithm, a population of an initial set of candidate solutions in the feature data space is generated and iteratively evolved and updated. An evolutionary algorithm has two main parts: EXPLOITATION and EXPLORATION. EXPLOITATION tries to improve the current solutions in order to reach new local optimum, by evolving gradually the quality rather than only producing new results. EXPLORATION produces new solutions in the feature space search environment. One possible algorithm for exploration is random search, which can produce new results in every part of the feature space.

Specifically, in our segmentation proposal as schematized in Algorithm 1, an initial population [step 1] containing *nPop* random groups of candidate solutions from $\mathbf{F}_{2.5D}$ are generated where each candidate solution contains k cluster centers. Each candidate solution is evaluated by a cost function (*Cost*), [step 2]. Here, we use *Cost* function defined in Eq. (4). A candidate solution with minimum *Cost* a long with its corresponding clusters is saved as the present best looked-for optimum solution of k clusters, *OptimalSol*, [step 3].

$$Cost = \sum_{z=1}^{3} \sum_{r=1}^{N \times N} \min \left(D(i) = \sqrt{\sum_{j=1}^{m} \left(f_{rjz} - \text{cluster center}_{ijz} \right)^2} \quad for\, i = 1 : k \right)$$
$$(4)$$

where, $D(i)$ is an Euclidean distance computed between the m-valued feature vector in the rth row of $\mathbf{F}_{2.5D}$ and the ith cluster center's feature vector, per each one of the 3 nodular planes.

Afterwards, among *nPop* an exclusive elite sub-group is selected as *nElitePop* [step 4], following an algorithm-specific rule, e.g. for GA, *nElitePop* could be the set of candidate solutions with the best-fit (lower *Cost*) values of *nPop*. During the EXPLOITATION stage [step 6], each candidate solution of *nElitePop* [step 7] is evolved and updated once [step 8], e.g. for GA, this could be crossover and then mutation. Followed by *Cost* computation [step 9], if the updated candidate solution has a lower *Cost* value, then it replaces its original version [steps 10–11], and thus converge to the local optimum. Consequently, it guarantees a better overall clustering solution.

After updating, the so-called non-elite sub-group of *nNonElitePop* is picked [step 12], following an algorithm-specific rule, ensuring that the metaheuristic search dynamically explores other locations of the feature space. For instance, for FA, *nNonElitePop* could be the candidate solutions with lower brightness (higher *Cost*). During EXPLORATION [step 13], each one of the members of *nNonElitePop* is swapped by a new random candidate solution if the new one has a lower *Cost* [steps 14–18]. This process is repeated for number of iterations (*MaxIt*) and *OptimalSol* with a minimum *Cost* is computed per iteration [steps 19–20]. The final *OptimalSol* is returned as the best output looked-for optimum solution of k clusters [step 21]. The final nodule segmentation (*S*) is the set of positions of the pixels' feature vectors from *OptimalSol* having the same cluster of the pixel's feature vector located in the centroid of the 3 nodular planes. A morphological operation of dilation is applied to refine the border of the segmented areas**.**

Fig. 4 illustrates an example of the segmentation results obtained by the metaheuristic search using TLBO. This figure shows a juxta-vascular solid nodule using a 2.5D representation after image enhancement (Fig. 4(a$_1$)–(a$_3$)). The corresponding clustering results for each plane are shown in Fig. 4(b$_1$)–(b$_3$) (for better representation the clusters are displayed in different colors). Fig. 4(c$_1$)–(c$_3$) present the final segmentation result. The boundaries of the segmented nodular areas are represented in red, Fig. 4(d$_1$)–(d$_3$). The plot in Fig. 4(e$_1$) shows the evolution of the optimum *Cost* value obtained after each iteration of the metaheuristic search approach.

3.3. Volume measurement

The estimation of the segmented nodule volume (*V*) can be obtained by ellipsoid approximation, using the expression,

$$V = \pi d_1 d_2 d_3 / 6 \qquad (5)$$

where, d_1, d_2 and d_3 are the equivalent diameters of the segmented regions in the 3 nodular planes, corresponding to the approximate circular intrinsic shape of a pulmonary nodule in the planes.

4. Experimental set-up and results

The following settings are used through the experiments. For k-clustering, the number of clusters k is an important hyper-parameter to produce clusters with a low intra-cluster and a high inter-cluster distances. To find a good solution for k, we experimentally tested the performance of the clustering with $k = 5, 7, 9, 11, \ldots, 31$. The selected value for k is $k = 15$, as with this value the best performance was achieved among all the tested k values. Besides, *MaxIT* and *nPop* are set to 100 and 50. The evolutionary algorithms used, namely FA, GA and TLBO, by the metaheuristic search are initialized with the settings proposed by Holland (1992), Yang (2008), and Rao, Savsani, and Vakharia (2011).

Table 1
Brief descriptions of the evolutionary algorithms namely FA, GA and TLBO used in the metaheuristic search[a].

Evolutionary algorithm	Description
Firefly Algorithm (FA), Yang (2008)	- Attractiveness is proportional to the brightness (*Cost*) of the fireflies (candidate solutions containing cluster centers). Fireflies with better brightness are attractive to the ones with the less brightness, and so for any two fireflies, the less bright one will be attracted by (and thus move towards) the brighter one. However, distance between the fireflies is an important property. The intensity (apparent brightness) decreases as their mutual distance increases. - If there is no fireflies brighter than a given firefly, it would be moved randomly.
Genetic Algorithm (GA), Holland (1992)	- Selection is proportional to the individuals (candidate solutions containing cluster centers) with the best-fit (lower *Cost*). The best-fit individuals as parents are selected for reproduction. - Generate new individuals (offsprings) using crossover and mutation operations.
Teaching Learning Based Optimization (TLBO), Rao et al. (2011)	- Selection is proportional to a teacher (a candidate solution containing cluster centers) who has the highest knowledge (lower *Cost*). A teacher tries to improve the performance of his class (other candidate solutions as learners) by increasing the average knowledge of his class (mean (*Cost*)). - Learners increase their knowledge by interacting among themselves. A learner interacts randomly with other learners for enhancing his knowledge. For a learner, it moves toward a random selected learner if his knowledge is lesser than the selected one, and it takes some distance if his knowledge is higher. - If the knowledge of a class is not improved, the teacher would be changed randomly.

[a] *Note*: Further details of the metaheuristic search and the evolutionary algorithms are given in Holland (1992), Rao et al. (2011), Yang (2008) and Zäpfel, Bögl, and Braune (2010).
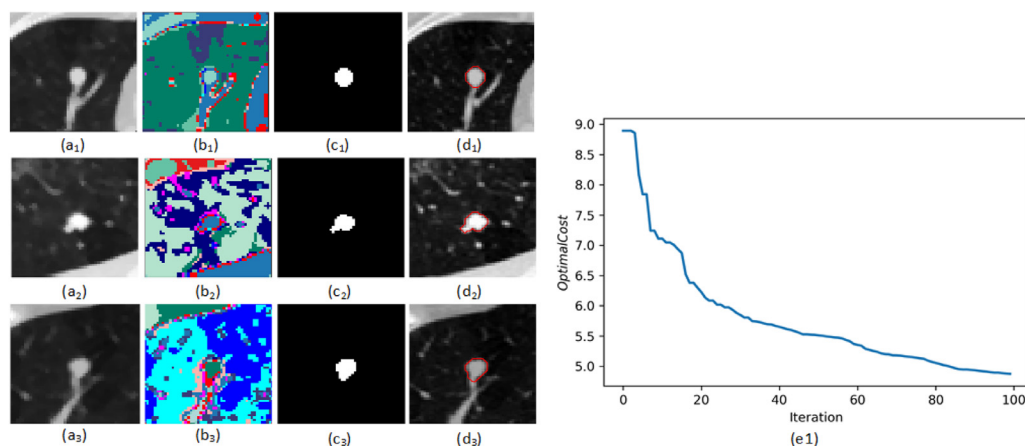
Fig. 4. Example of the segmentation results obtained by TLBO. (a_1)–(a_3) A juxta-vascular solid nodule using a 2.5D representation; (b_1)–(b_3) Results of the clustered pixels' feature data for each plane; (c_1)–(c_3) Binary segments corresponding to the region that hits the nodule centroid; (d_1)–(d_3) Contours of the segmented nodular areas are in red; (e_1) The optimum *Cost* value obtained after each iteration of the metaheuristic search approach. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1. LIDC-IDRI dataset

We assess the performance of the proposed segmentation and volume measurement on the LIDC-IDRI (Armato et al., 2011) dataset used by many researchers (Lassen et al., 2015; Messay et al., 2015; Wang et al., 2017). The LIDC-IDRI dataset contains lung CT scans of different patients. The dataset has 1018 scans with 2629 detected nodules. In total, 880 nodules have an agreement level equal to 4, indicating that all the 4 radiologists agree they are in presence of a nodule.

The nodule diameters range from 2.03 mm to 38.12 mm. The nodules show different lesion characteristics like subtlety, inter-

---

Algorithm 1 Metaheuristic search via an evolutionary algorithm.

Input: Feature matrix composed by feature vectors ($\mathbf{F}_{2.5D}$)
Output: Optimal Solution (*OptimalSol*)
1:　*Initialize* a population with *nPop* random candidate solutions from $\mathbf{F}_{2.5D}$
2:　*Evaluate* each candidate solution using *Cost* function
3:　*OptimalSol* = a candidate solution of min (*Cost*)
4:　Elite population (*nElitePop*) = *Evolutionary Computation* (*nPop*)
5:　**repeat**
6:　　**procedure** EXPLOITATION
7:　　　**for all** *nElitePop* **do**
8:　　　　an updated candidate solution = *Evolutionary Computation* (a present candidate solution)
9:　　　　*Evaluate* an updated candidate solution using *Cost* function
10:　　　　**if** an updated candidate solution.*Cost* < a present candidate solution.*Cost* **then**
11:　　　　　a present candidate solution = an updated candidate solution
12:　　Non Elite population (*nNonElitePop*) = *Evolutionary Computation* (*nPop*)
13:　　**procedure** EXPLORATION
14:　　　**for all** *nNonElitePop* **do**
15:　　　　*Generate* a new random candidate solution from $\mathbf{F}_{2.5D}$
16:　　　　*Evaluate* a new candidate solution using *Cost* function
17:　　　　**if** a new candidate solution.*Cost* < a present candidate solution.*Cost* **then**
18:　　　　　a present candidate solution = a new candidate solution
19:　　*OptimalSol* = a candidate solution of min (*Cost*)
20:　**until** number of iterations (*MaxIt*)
21:　**return** *OptimalSol*

---

nal structure, calcification, sphericity, margins definition, lobulation, spiculation, and have a degree of malignancy in the range [1, 6]. The texture of the nodules has the range [1, 5]. In our experiments, the nodules with an average radiologists' texture rating range of [1, 2.5) are considered as non-solid, those in a range [2.5, 4) are considered as part-solid, and a range of [4, 5] corresponds to the solid nodules.

The segmentation and volume measurement results are evaluated using two subsets of the LIDC-IDRI dataset. We consider as Dataset 1, 705 nodules located in different regions of the lungs, with an agreement level 4, 60 categorized as sub-solid (25 pure non-solid and 35 part-solid) and 645 as solid. We also consider the challenging dataset used in Lassen et al. (2015) (Dataset 2), consisting of 59 sub-solid nodules. Dataset 2 is organized in two subsets: 1) the first subset (Subset 1) has 19 sub-solid nodules of agreement level 4; 2) the second one (Subset 2) includes 40 sub-solid nodules, with radiologists' agreement greater equal to 3. The nodule volumes are rescaled to a size $51 \times 51 \times 51$ mm with $1 \times 1 \times 1$ mm voxel size.

4.2. Evaluation metrics

Dice similarity coefficient (*DSC*) is generally used to compute the degree of overlap of a segment *S* with a segment used as reference *G*, the ground-truth (Jung et al., 2018; Wang et al., 2017). It is defined in Eq. (6). The range of *DSC* is [0, 1], where low values correspond to a poor segmentation performance. The pixel classification performance and the correctness of the segmentation area are measured by the sensitivity (*SEN*) and the positive predictive value (*PPV*) (Wang et al., 2017), Eq. (7). *SEN* and *PPV* close to 1 indicate a high level of agreement. The average boundary distance (*AVD*) between surfaces *S* and *G* (Lassen et al., 2015; Wang et al., 2017) is measured by Eq. (8). The more similar the segmented area *S* with respect to *G* is, the closer *AVD* is to 0.

$$DSC = \frac{2|S \cap G|}{|S| + |G|} \tag{6}$$

$$SEN = \frac{|S \cap G|}{|G|}, PPV = \frac{|S \cap G|}{|S|} \tag{7}$$

$$AVD = \frac{1}{2}\left(\text{mean}_{i \in G}\ \min_{j \in S} d(i, j) + \text{mean}_{i \in S}\ \min_{j \in G} d(i, j)\right) \tag{8}$$

where, the Euclidean distance between pixels *i* and *j* measured in millimeters (mm) is denoted by $d(i, j)$.

The ground-truth boundaries delineated by expert radiologists, differ from observer to observer mainly at nodules with complex shape and irregular contours. Such lack of correspondence affects the validation of the segmentation process. As a reference, we measure the inter-observer segmentation performance by computing the corresponding inter-observer *DSC, SEN, PPV* and *AVD*. For each metric, the average value is considered as the inter-observer segmentation performance. In here, the values obtained for *SEN* and *PPV* are the same since each boundary *G* considered as *S* plays the opposite role in a next computation.

4.3. Segmentation performance on Dataset 1

Table 2 summarizes the segmentation performance of the metaheuristic search using FA, GA and TLBO on Dataset 1. Table 2 shows that we are able to achieve a segmentation performance comparable to the inter-observer performance. It can be also observed that TLBO provides a better segmentation than the other metaheuristics approaches. Comparison and discussion of the performance of FA, GA and TLBO will be given in Section 5, and hereinafter, in this Section 4, we will only present the segmentation performance of TLBO.

For all the nodules in Dataset 1, the performance of TLBO can be summarized by the following overall average values: $DSC = 82.35\% \pm 5.54$; $SEN = 87.11\% \pm 9.85$; $PPV = 85.59\% \pm 11.19$; and $AVD = 0.49$ mm ± 0.32. The distributions of *DSC, AVD, SEN* and *PPV* values for all the nodules are shown in Fig. 5. It can be observed that most of the nodules achieve a high segmentation similarity and a pixel classification performance with a low average boundary distance.

Fig. 6 shows the segmentation performance on 5 illustrative examples using TLBO, on each of the 3 nodular planes from Dataset 1. This figure shows the manual segmentations of 4 radiologists together with the segmentation results obtained by our proposed approach. There are several observations we can raise regarding Fig. 6. There exists a higher variability and inconsistency among observers' segmentations for the nodules with low contrast nodule boundaries, complex shapes and irregular contours. This inter-observer variability can be objectively measured by *DSC* and *AVD*, that ranges from 95.20% and 1.06 mm (an almost full agreement among observers) observed in the large well-defined

Table 2
Segmentation performance of the metaheuristic search using FA, GA and TLBO on Dataset 1.

Nodule texture [number of nodules]	DSC (%)	SEN (%)	PPV (%)	AVD (mm)
solid [645]				
method using FA	82.31 ± 5.25	84.84 ± 10.99	87.86 ± 10.91	**0.49 ± 0.33**
method using GA	82.23 ± 5.51	84.55 ± 10.69	**88.19 ± 10.50**	0.49 ± 0.36
method using TLBO	**82.73 ± 5.41**	**87.29 ± 9.52**	86.03 ± 10.84	0.50 ± 0.33
inter-observer performance	84.52 ± 5.97	85.56 ± 5.08	85.56 ± 5.08	0.45 ± 0.32
part-solid [35]				
method using FA	76.63 ± 5.60	81.63 ± 13.84	80.94 ± 14.51	0.51 ± 0.32
method using GA	73.63 ± 9.27	79.20 ± 15.85	78.84 ± 17.95	0.54 ± 0.34
method using TLBO	**78.04 ± 5.72**	**83.64 ± 13.24**	**81.69 ± 14.41**	**0.45 ± 0.19**
inter-observer performance	81.30 ± 5.95	82.53 ± 5.54	82.53 ± 5.54	0.43 ± 0.24
non-solid [25]				
method using FA	77.41 ± 6.12	87.50 ± 13.51	78.06 ± 12.56	0.48 ± 0.25
method using GA	77.56 ± 4.89	**87.85 ± 11.91**	77.80 ± 11.72	0.48 ± 0.24
method using TLBO	**78.50 ± 4.66**	87.17 ± 11.62	**79.86 ± 12.27**	**0.47 ± 0.23**
inter-observer performance	80.62 ± 4.36	81.93 ± 3.95	81.93 ± 3.95	0.43 ± 0.24
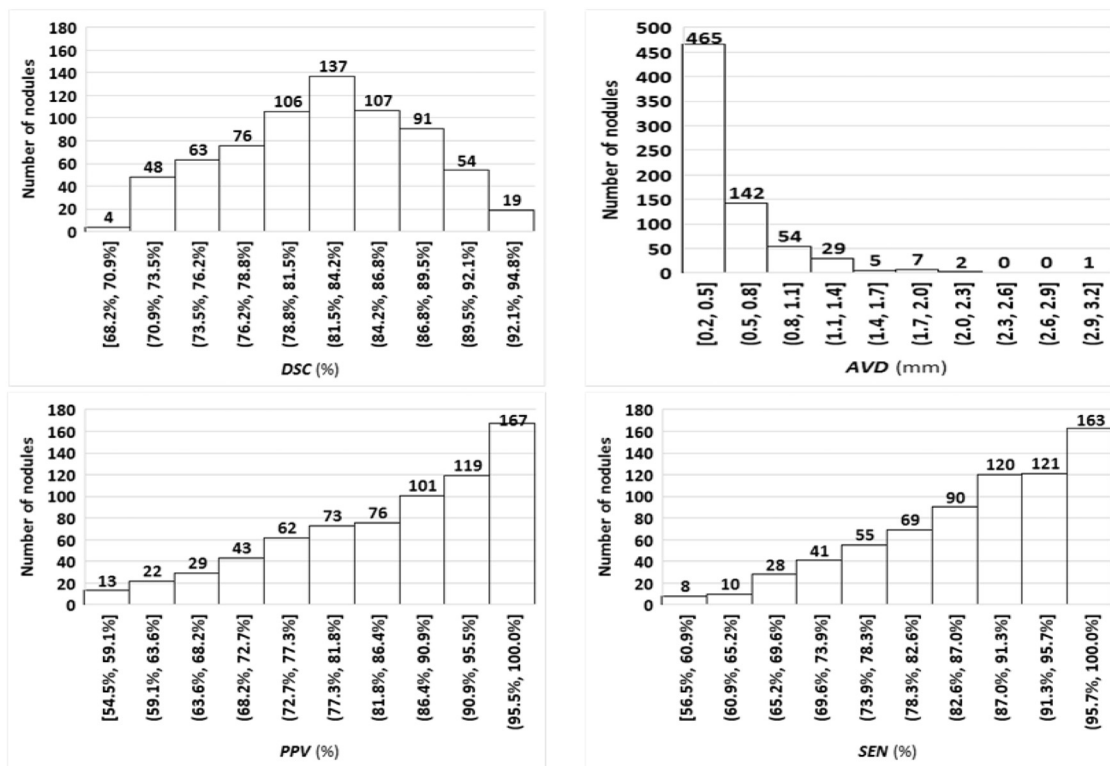
Fig. 5. Performance measure histogram for all the nodules in Dataset 1 obtained by TLBO.

juxta-vascular solid nodule (Fig. 6(a)), to a poor agreement for the pleural-tail solid nodule in Fig. 6(d) with a $DSC = 68.10\%$ and an $AVD = 2.32$ mm, respectively. An intermediate DSC and AVD is observed for the medium size pleural-tail solid nodule in Fig. 6(b) with a $DSC = 77.60\%$ and an $AVD = 1.10$ mm, and also for the small size isolated part-solid nodule in Fig. 6(c) with a $DSC = 86.50\%$ and an $AVD = 0.13$ mm.

It can be observed that the proposed segmentation method gives similar results comparable to the inter-observer DSC and AVD, showing also good performances for the nodules with different levels of complexity. However, as expected, the performance of the proposed segmentation method depends on the level of the nodule complexity, namely in terms of location, texture, shape and contour smoothness. As shown in Fig. 6(e), for a large size juxta-

pleural solid nodule with a $DSC = 91.90\%$ and an $AVD = 0.62$ mm (a nearly high agreement among observers), an over-segmented result is obtained where a connection between the nodule and the pleura makes the segmentation result containing both the nodule and the lung pleura in its axial plane.

4.4. Segmentation performance on Dataset 2

Dataset 2 is considered in this experiment to evaluate the segmentation performance of the metaheuristic search using TLBO. The obtained results are presented in Table 3. For comparison, the inter-observer average segmentation performance is also shown. This table shows that we achieve better segmentation results for the sub-solid nodules of Subset 1 than for Subset 2. This is because

Axial plane	Sagittal plane	Coronal plane	Nodule type and performance

(a) — **Large juxta-vascular solid nodule:**
DSC = 94.60%
AVD = 1.13 mm
DSC_inter-observer = 95.20%
AVD_inter-observer = 1.06 mm

(b) — **Medium pleural-tail solid nodule:**
DSC = 73.40%
AVD = 1.41 mm
DSC_inter-observer = 77.60%
AVD_inter-observer = 1.10 mm

(c) — **Small isolated part-solid nodule:**
DSC = 82.50%
AVD = 0.21 mm
DSC_inter-observer = 86.50%
AVD_inter-observer = 0.13 mm

(d) — **Large pleural-tail solid nodule:**
DSC = 61.80%
AVD = 3.05 mm
DSC_inter-observer = 68.10%
AVD_inter-observer = 2.32 mm

(e) — **Large juxta-pleural solid nodule:**
DSC = 60.10%
AVD = 1.83 mm
DSC_inter-observer = 91.90%
AVD_inter-observer = 0.62 mm

Color legend: Red − segmentation of the proposed approach. Yellow, green, blue and cyan − segmentation of 4 radiologists.

Fig. 6. Segmentation performance on 5 illustrative examples using TLBO, on each of the 3 nodular planes from Dataset 1.(a) A juxta-vascular solid nodule of large size; (b) A pleural-tail solid nodule of medium size; (c) An isolated part-solid nodule of small size; (d) A pleural-tail solid nodule of large size; (e) An over-segmented result of a juxta-pleural solid nodule of large size where the region of the lung pleura is segmented along with the nodule in its axial plane. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Segmentation performance of the metaheuristic search using TLBO on Dataset 2.

Sub-solid nodules [number of nodules]	DSC (%)	SEN (%)	PPV (%)	AVD (mm)
Subset 1 [19]				
method using TLBO	**76.92 ± 5.77**	**85.02 ± 12.96**	**78.95 ± 13.46**	**0.53 ± 0.24**
inter-observer performance	80.37 ± 4.54	81.80 ± 4.16	81.80 ± 4.16	0.47 ± 0.25
Subset 2 [40]				
method using TLBO	68.26 ± 12.37	80.01 ± 22.41	70.32 ± 17.33	0.69 ± 0.56
inter-observer performance	80.57 ± 6.43	82.12 ± 5.56	82.12 ± 5.56	0.49 ± 0.37

the nodular surfaces of Subset 2 have very irregular shapes and dissimilar features. For all the sub-solid nodules in Dataset 2, we obtain a DSC = 71.05% ± 11.43. As for the classification performance, we obtain SEN = 81.63% ± 20.00 and PPV = 73.10% ± 16.68. An AVD is equal to 0.63 mm ± 0.48.

In a Bland-Altman plot (Bland & Altman, 1999), the mean between metric values of two different methods are plotted against their differences to evaluate and compare the agreement between the two methods. The two methods are often considered interchangeable if the mean versus the difference values are within the standard deviation lines of the Bland-Altman plot (Kubota et al., 2011). For the sub-solid nodules in Dataset 2, Table 4 provides the global mean difference (*Mean*) and their corresponding standard deviation (*SD*) for the metrics DSC and AVD used in the Bland-Altman plots in Fig. 7.

In Table 4, the absolute high values of *Mean* and *SD* for the metrics DSC and AVD correspond to a low agreement between the

segmentation results and the inter-observer segmentation performance. The agreement between the segmentation similarity results and the inter-observer segmentation similarity performance for Subset 1 is higher than for Subset 2. It also shows that the proposed segmentation method achieves better agreement on average boundary distance for the sub-solid nodules of Subset 1 than for Subset 2. Furthermore, as seen in Fig. 7(a) and (b), our segmentation approach shows good agreement, with a mean DSC close or greater than 70.00%. Fig. 7(c) and (d) show a mean AVD for most of the nodules in the nodules in the range [0.20, 0.40] mm. The segmentation results show a performance comparable to the inter-observer segmentation performance especially for the sub-solid nodules in Subset 1.

Illustrative examples of the segmentation results from the two subsets of Dataset 2 are shown in Fig. 8.

Fig. 8(a)–(f) present the sub-solid nodules placed in a complex location as the pleura. For the sub-solid nodules in Fig. 8, our seg-

Table 4
Bland-Altman descriptive statistics for *DSC* and *AVD*.

Sub-solid nodules [number of nodules]	segmentation results obtained by our proposed method & inter-observer segmentation performance			
	Difference on *DSC* (%)		Difference on *AVD* (mm)	
	Mean	*SD*	*Mean*	*SD*
Subset 1 [19]	−3.46	7.94	0.06	0.13
Subset 2 [40]	−12.31	14.73	0.20	0.39
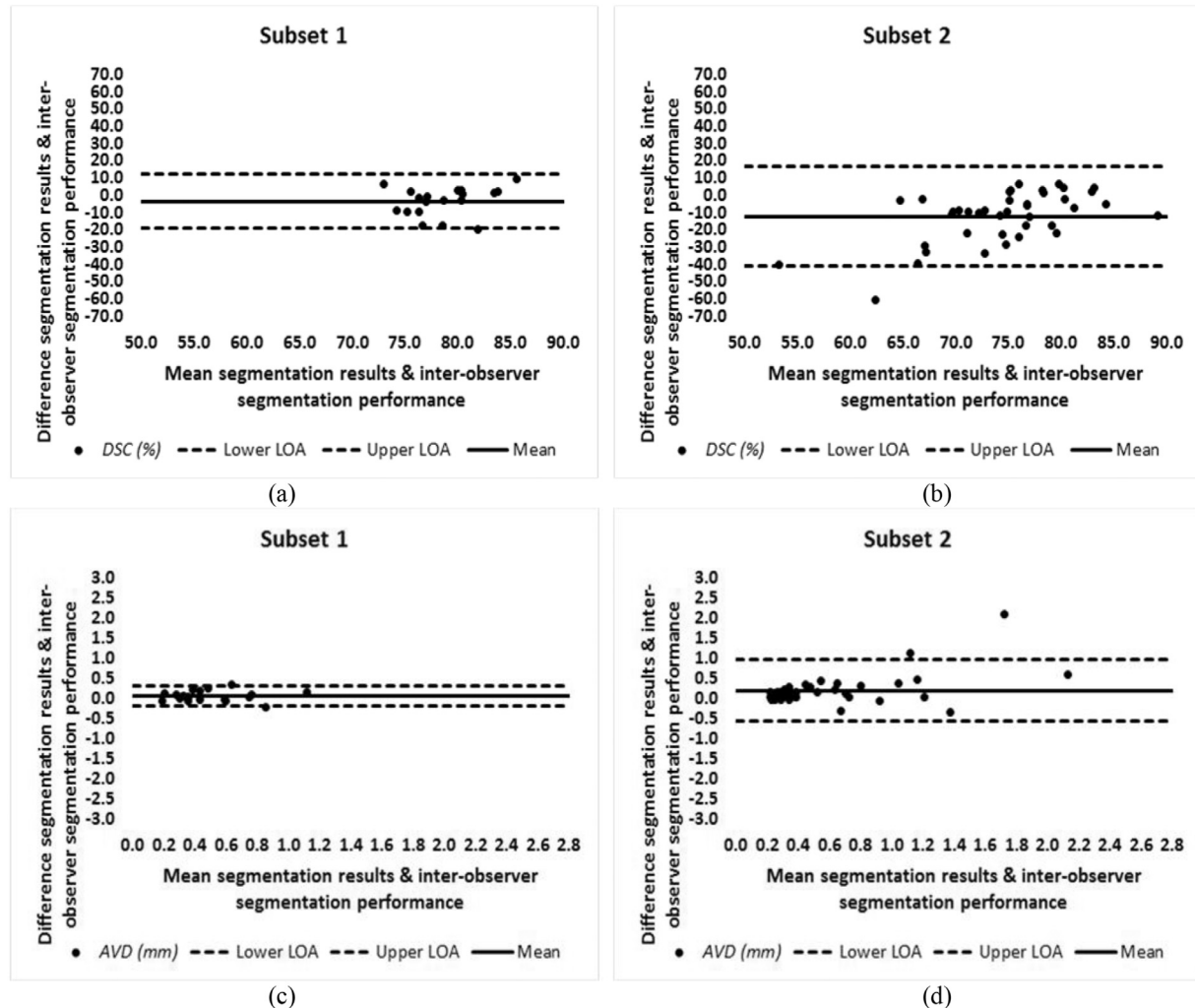
Fig. 7. Bland-Altman plots for segmentation of the sub-solid nodules in Dataset 2; the solid horizontal lines correspond the mean difference; the dashed horizontal lines correspond to ± 1.96 times the standard deviation of the differences; (LOA – Level of Agreement).
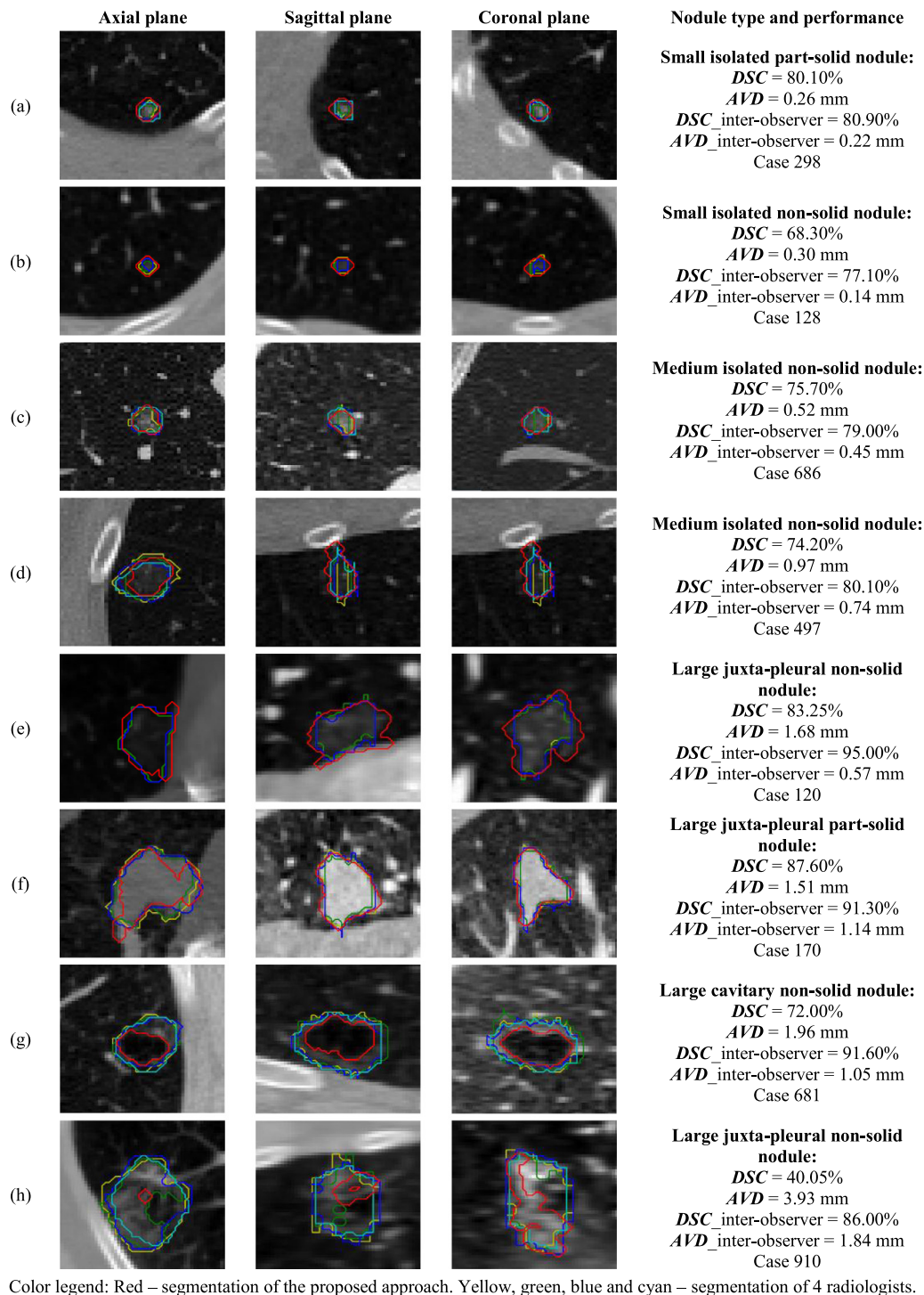
mentation proposal performs well, for different nodule shapes and irregular contours. In this figure, we also show cases with poor segmentation performance. Fig. 8(g) shows the effect of the segmentation for the cavitary nodule where the algorithm only segments the cavitary region. An under segmentation result is obtained for the non-solid nodule in Fig. 8(h), due to the inhomogeneous characteristics of its corresponding nodular region with fuzzy contour.

4.5. Volume measurement performance on Dataset 1 and Dataset 2

We measure the volume of S, V_S, using the expression given in Eq. (5). Similarly, a $V_{2.5G}$ is assessed by computing the average of the manual segmentations' volumes using a 2.5D representation obtained with Eq. (5). For evaluating the results of the volume measurement in 2.5D versus 3D, a V_{3G} volume is measured by computing the average of the total number of voxels of the man-

ual segmentations in 3D. For the nodules in Dataset 1 and Dataset 2, we compute the Pearson correlation coefficient, Pearson's r (R-value) between $V_{2.5G}$ and V_{3G}, and between V_S and V_{3G}. The results are shown in Table 5. We also analyze if the volume measurement has statistical significance. Table 5 includes the results of the t-test analysis (p-value) with a significance level of 5%.

In Table 5, the measured volumes for the 2.5D manual segmentations versus the 3D manual segmentations are highly correlated with the R-values ≥ 99.61%. The p-values also show that the volume measurement has significative correlation between the volumes measured in 2.5D versus 3D. For the nodules in Dataset 1 segmented using our segmentation proposal versus the 3D manual segmentations, the R-values are greater than 92.16%. The volume measurement shows a significative correlation for all the nodules, as well. The measured volumes for the sub-solid nodules of Subset 1 in Dataset 2 are also highly correlated with an R-value = 95.85%, with an p-value of 3.96%. For the sub-solid nodules of Subset 2

Axial plane	Sagittal plane	Coronal plane	Nodule type and performance

(a)
Small isolated part-solid nodule:
DSC = 80.10%
AVD = 0.26 mm
DSC_inter-observer = 80.90%
AVD_inter-observer = 0.22 mm
Case 298

(b)
Small isolated non-solid nodule:
DSC = 68.30%
AVD = 0.30 mm
DSC_inter-observer = 77.10%
AVD_inter-observer = 0.14 mm
Case 128

(c)
Medium isolated non-solid nodule:
DSC = 75.70%
AVD = 0.52 mm
DSC_inter-observer = 79.00%
AVD_inter-observer = 0.45 mm
Case 686

(d)
Medium isolated non-solid nodule:
DSC = 74.20%
AVD = 0.97 mm
DSC_inter-observer = 80.10%
AVD_inter-observer = 0.74 mm
Case 497

(e)
Large juxta-pleural non-solid nodule:
DSC = 83.25%
AVD = 1.68 mm
DSC_inter-observer = 95.00%
AVD_inter-observer = 0.57 mm
Case 120

(f)
Large juxta-pleural part-solid nodule:
DSC = 87.60%
AVD = 1.51 mm
DSC_inter-observer = 91.30%
AVD_inter-observer = 1.14 mm
Case 170

(g)
Large cavitary non-solid nodule:
DSC = 72.00%
AVD = 1.96 mm
DSC_inter-observer = 91.60%
AVD_inter-observer = 1.05 mm
Case 681

(h)
Large juxta-pleural non-solid nodule:
DSC = 40.05%
AVD = 3.93 mm
DSC_inter-observer = 86.00%
AVD_inter-observer = 1.84 mm
Case 910

Color legend: Red − segmentation of the proposed approach. Yellow, green, blue and cyan − segmentation of 4 radiologists.

Fig. 8. Segmentation performance on 8 illustrative examples using TLBO, on each of the 3 nodular planes from Dataset 2. (a)-(d) Isolated sub-solid nodules of different sizes; (e)–(f) Juxta-pleural sub-solid nodules of large size; (g) Under-segmented cavitary non-solid nodule of large size; (h) Under-segmented juxta-pleural non-solid nodule of large size. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in Dataset 2, a low correlation, with an R-value = 39.33%, is obtained. Table 5 shows an overall good performance for the nodule volume measurement in Dataset 1 and Dataset 2 using a 2.5D representation except for some of the sub-solid nodules of Subset 2 in Database 2.

Fig. 9 shows Pearson's r between V_S and V_{3G} for the nodules in Dataset 1 and Dataset 2. As seen in Fig. 9, the measured volume between the results obtained by our proposed segmentation method and the manual segmentations are highly correlated. This allows us to conclude, once again, that our segmentation proposal gives a reliable nodule segmentation in a 2.5D representation, which can be useful to measure the volume in a lung CAD system.

5. Discussion

In this paper, we use the LIDC–IDRI dataset with nodules of varied characteristics of lesions located in different regions of the lungs to evaluate the developed nodule segmentation and volume

Table 5
Statistical analysis of the nodule volume measurement on Dataset 1 and Dataset 2.

Nodule texture [number of nodules]	2.5D manual segmentations versus 3D manual segmentations			2.5D segmentation results versus 3D manual segmentations		
	R value (%)	p value (%)	significative	R value (%)	p value (%)	significative
Dataset 1:						
Solid [645]	99.61	$4.62e^{-23}$	Yes	94.10	$7.26\ e^{-36}$	Yes
part-solid [35]	99.72	0.00	Yes	92.16	0.00	Yes
non-solid [25]	99.94	0.01	Yes	95.64	1.77	Yes
Dataset 2:						
Subset 1-sub-solid [19]	99.94	0.04	Yes	95.85	3.96	Yes
Subset 2-sub-solid [40]	99.81	0.00	Yes	39.33	0.70	Yes
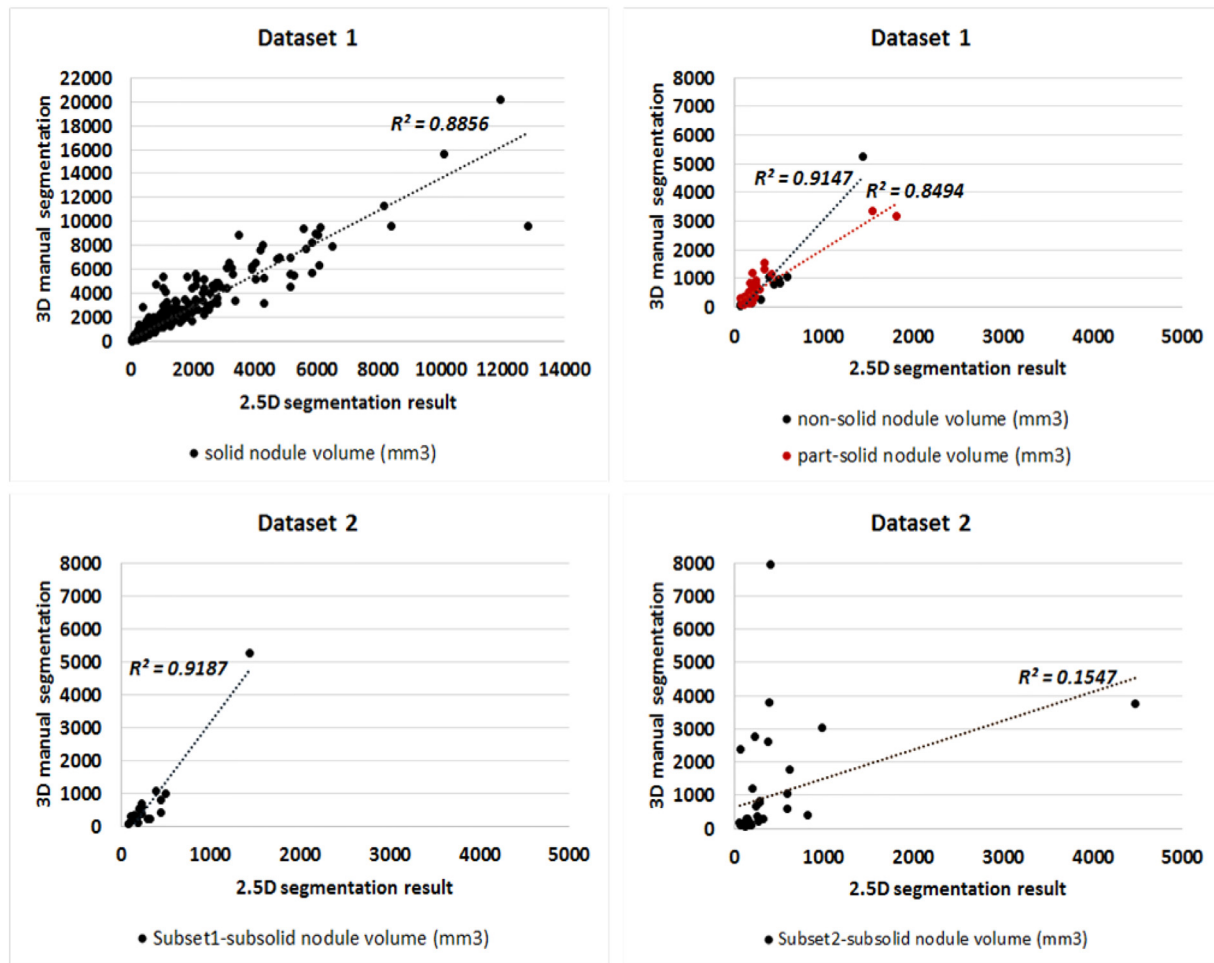
Fig. 9. Correlation coefficient for volume measurement of the nodules in Dataset 1 and Dataset 2.

measurement schemes. The experimental results presented in Section 4 show that the proposed segmentation method is able to segment nodules located in different regions of the lungs and with different texture, size and shape. To further evaluate the methodology on Dataset 1, we compute the average DSC for the nodules having the four lesion characteristics mentioned in Wang et al. (2017), namely the spiculation, the malignancy, the calcification and the sphericity. The corresponding average DSC is tabulated in Table 6. As represented in Table 6, the average DSC values do not show any significative difference for the different nodule characteristics, being the proposed segmentation method able to capture nodule shapes with respect to miscellaneous characteristics of lesions (Wang et al., 2017).

We also compare the segmentation performance with recently published state-of-the-art techniques evaluated for nodules of the LIDC–IDRI dataset by at least three radiologists

Table 6
Average DSC for different types of the nodules in Dataset 1 with ratings [1, 6].

Lesion characteristic ratings						
Lesion characteristics	1	2	3	4	5	6
Spiculation	82.25	82.47	83.36	82.98	78.58	–
[number of nodules]	[540]	[37]	[22]	[4]	–	
Malignancy	82.09	81.11	82.35	84.73	82.4	–
[number of nodules]	[94]	[219]	[266]	[123]	[3]	–
Calcification	–	82.15	82.67	81.29	83.00	82.26
[number of nodules]	–	[2]	[86]	[16]	[55]	[546]
Sphericity	–	79.09	81.61	83.23	83.68	–
[number of nodules]	–	[40]	[288]	[359]	[18]	–

(Kubota et al., 2011; Lassen et al., 2015; Messay et al., 2015; Messay, Hardie, & Rogers, 2010; Tan et al., 2013; Wang et al., 2017). Tables 7 and 8 show the performance achieved by different authors on Dataset 1 and Dataset 2, together with the performance

Table 7

Average *JI* for different nodule segmentation proposals on the LIDC-IDRI dataset evaluated by at least 3 radiologists.

Methods	Number of nodules		Performance *JI* (%)
	Training	Testing	
Messay et al. (2010)	–	68	63.00 ± 16.00%
Kubota et al. (2011)	–	23	69.00 ± 18.00%
		82	59.00 ± 19.00%
Tan et al. (2013)	–	23	65.00%
Messay et al. (2015)	300	66	71.70 ± 19.89%
		77	69.23 ± 13.82%
Wang et al. (2017)	350	493	71.16 ± 12.22%
Proposed method using TLBO on Dataset 1	–	705	70.37 ± 8.00%

Table 8

Average *JI* for the nodule segmentation proposals on the dataset given in Lassen et al. (2015).

Methods	Number of nodules		Performance *JI* (%)
	Training	Testing	
Lassen et al. (2015)	–	19	52.00 ± 7.00%
		40	50.00 ± 14.00%
Proposed method using TLBO on Dataset 2	–	19	62.53 ± 8.00%
		40	53.06 ± 13.32%

Table 9

Approximate processing speed for different nodule segmentation proposals on the LIDC-IDRI dataset.

Methods	Computer configuration	Computation time	User interaction
Kubota et al. (2011)	Not Provided	≤1 min	A click point initialization
Lassen et al. (2015)	Not Provided	≤1 s	A user-drawn stroke initialization
Messay et al. (2015)	MATLAB implementation on an HP workstation with processor speed of 3.22 GHz.	4 min	Click point(s) initialization
Wang et al. (2017)	Python 2.7 implementation on a machine with an Intel Core i7-4790K CPU, 8 GB RAM and an NVIDIA GTX-980Ti GPU with 6 GB on-board memory.	The CF-CNN model converged after 9 h of training on 0.41 million voxel patches.	Automatic
Proposed method using FA	Python 3. 6 implementation on an ASUS workstation with an Intel® Core™ i7-5960X CPU @ 3.0 GHz and 32 GB RAM.	4.75 min	Automatic
Proposed method using GA/TLBO	Python 3. 6 implementation on an ASUS workstation with an Intel® Core™ i7-5960X CPU @ 3.0 GHz and 32 GB RAM.	26 s	Automatic

of our segmentation approach using TLBO. To compare the methods, we use the Jaccard index (*JI*) (please note that *JI* and *DSC* are related by the expression $JI = DSC / (2 - DSC)$). Using the results in Table 7, we can argue that our method is in par with the best performant methods, with an advantage of getting a lower uncertainty for the *JI* similarity estimate, as we use a much larger dataset than the other approaches. Similarly, Table 8 shows the performance on Dataset 2, which are better than those reported by Lassen et al. (2015). More importantly, the works by Kubota et al. (2011), Lassen et al. (2015), Messay et al. (2015), and Tan et al. (2013) require a manual user-initialization whereas our segmentation approach is fully automatic, without any manual annotation. In addition, our method is unsupervised while the approach by Wang et al. (2017) relies on training samples. However, the nodule segmentation proposed in this paper is based on a 2.5D representation whereas the other approaches compared are based on a 3D representation.

The performance of the proposed segmentation method typically relies on feature extraction and the exploitation part of the evolutionary algorithm used by the metaheuristic search. As discussed in Section 3, this paper uses FA, GA and TLBO evolutionary algorithms, for image segmentation. TLBO proved to be the best performing algorithm. Using GA, the new candidate solutions containing cluster centers inherit the features of the existing candidate solutions. Thus, the intrinsic nature of the current candidate solutions continues to be spread in the iterations of the algorithm. As a result, segmentation of sub-solid nodules using GA shows low performance, due to the poor homogeneity and well-marked features of the sub-solid nodules, and since GA only attempts to use the information of the current candidates to produce the new ones, the optimum clusters are not obtained.

FA shows better segmentation performance in general and, in particular, for sub-solid nodules. The backbone of FA is based on attracting the candidate solutions with higher qualified features, termed as lights in the FA algorithm. The light absorption coefficient plays an important role in this algorithm. In addition to light, the distance of the candidate solutions with respect to each other is also considered. A candidate solution with higher qualified features but locating in a farther distance is not considered. This nature also aids to give a chance to the candidate solutions with lower qualified features to be exploited. In a population, any group of the candidate solutions will be attracted to all other groups, resulting in a high computational complexity.

Among all, TLBO has teaching and learning stages where the candidate solutions make an effort to compete with each other in order to gain knowledge by transmitting their own information. This allows the candidate solutions to improve their information. In the exploitation phase of this algorithm, the candidate solutions

are updated by information discovery of the current candidate solutions.

Regarding the processing speed for the proposed prototype Python implementation, we may say that the computation burden is mainly due to the time spent by the evolutionary algorithms, for evolving the candidate solutions. We use Python version 3.6 on an ASUS workstation with an Intel® Core™ i7-5960X CPU @ 3.0 GHz and 32 GB RAM. For each nodule, the FA approach, has a computation time in the range [4, 5.5) min with an average of 4.75 min. For GA and TBLO, the estimated computation time is in the range (23, 30) s with an average close to 26 s, being this low value one additional assert for the TLBO method. These estimates are average values, considering all the nodules. For comparison purpose, Table 9 shows the approximate processing speed for several nodule segmentation methods on the LIDC-IDRI dataset, where we can see the competitive performance of the proposed segmentation method using TLBO.

6. Conclusion

An unsupervised pulmonary nodule segmentation method on lung CT images was described in this paper. We proposed to segment a nodule using a 2.5D representation through the meta-heuristic search, in here based on evolutionary computation. Our focus was to segment different types of nodules in terms of texture and location without the need of any manual annotation or further post-processing steps. We also addressed the measurement of the volume of the segmented nodules via ellipsoid approximation. We validated the proposed approach along with volume measurement on two different datasets, subsets of the LIDC-IDRI dataset. The comparison between the segmented results and the ground-truth for the corresponding two datasets demonstrated the robustness of our system. However, it failed to segment a number of juxta-pleural nodules due to some of their features, like curvedness and intensity. Segmentation performance fell down for the sub-solid nodules such as cavitary ones because of their heterogeneous nodular surface characteristics. To improve the performance of the proposed segmentation method, particularly for this type of more complex nodules, we need to extract other type of features, trying to characterize the different levels of complexity that can appear in lung CT images.

Acknowledgment

References

Armato, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, Ch. R., Reeves, A. P., et al. (2011). The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics, 38*(2), 915–931.

Awad, J., Owrangi, A., Villemaire, L., O'Riordan, E., Parraga, G., & Fenster, A. (2012). Three-dimensional lung tumor segmentation from X-Ray computed tomography using sparse field active models. *Medical Physics, 39*(2), 851–865.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research, 8*(2), 135–160.

Devaraj, A., van Ginneken, B., Nair, A., & Baldwin, D. (2017). Use of volumetry for lung nodule management: Theory and practice. *Radiology, 284*(3), 630–644.

Diciotti, S., Lombardo, S., Falchini, M., Picozzi, G., & Mascalchi, M. (2011). Automated segmentation refinement of small lung nodules in CT scans by local shape analysis. *IEEE Transactions on Biomedical Engineering, 58*(12 PART 1), 3418–3428.

Farag, A. A., Abd, H. E., Munim, E., Graham, J. H., & Farag, A. A. (2013). A novel approach for lung nodules segmentation in chest CT using level sets. *IEEE Transactions on Image Processing, 22*(12), 5202–5213.

Gavrielides, M. A., Kinnard, L. M., Myers, K. J., & Petrick, N. (2009). Noncalcified lung nodules: Volumetric assessment with thoracic CT. *Radiology, 251*(1), 26–37.

Gonçalves, L., Novo, J., & Campilho, A. (2016). Hessian based approaches for 3D lung nodule segmentation. *Expert Systems with Applications, 61*, 1–15.

Gu, Y., Kumar, V., Hall, L. O., Goldgof, D. B., Li, Ch-Y., Korn, R., et al. (2013). Automated delineation of lung tumors from CT images using a single click ensemble segmentation approach. *Pattern Recognition, 46*(3), 692–702.

Han, D., Heuvelmans, M. A., & Oudkerk, M. (2017). Volume versus diameter assessment of small pulmonary nodules in CT lung cancer screening. *Translational Lung Cancer Research, 6*(1), 52–61.

Heckel, F., Meine, H., Moltz, J. H., Kuhnigk, J-M., Heverhagen, J. T., Kießling, A., et al. (2014). Segmentation-based partial volume correction for volume estimation of solid lesions in CT. *IEEE Transactions on Medical Imaging, 33*(2), 462–480.

Holland, J. H. (1992). Genetic algorithms. *Scientific American, 267*(1), 66–73.

Jung, J., Hong, H., & Goo, J. M. (2018). Ground-glass nodule segmentation in chest CT images using asymmetric multi-phase deformable model and pulmonary vessel removal. *Computers in Biology and Medicine, 92*, 128–138.

Kalpathy-Cramer, J., Zhao, B., Goldgof, D., Gu, Y., Wang, X., Yang, H., et al. (2016). A comparison of lung nodule segmentation algorithms: methods and results from a multi-institutional study. *Journal of Digital Imaging, 29*(4), 476–487.

Keshani, M., Azimifar, Z., Tajeripour, F., & Boostani, R. (2013). Lung Nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system. *Computers in Biology and Medicine, 43*(4), 287–300.

Kubota, T., Jerebko, A. K., Dewan, M., Salganicoff, M., & Krishnan, A. (2011). Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Medical Image Analysis, 15*(1), 133–154.

Lassen, B. C., Jacobs, C., Kuhnigk, J. M., van Ginneken, B., & van Rikxoort, E. M. (2015). Robust semi-automatic segmentation of pulmonary subsolid nodules in chest computed tomography scans. *Physics in Medicine and Biology, 60*(3), 1307–1323.

Liu, H., Geng, F., Guo, Q., Zhang, C., & Zhang, C. (2017). A fast weak-supervised pulmonary nodule segmentation method based on modified self-adaptive FCM algorithm. *Soft Computing.* https://doi.org/10.1007/s00500-017-2608-5.

McKee, B. J., Regis, S. M., McKee, A. B., Flacke, S., & Wald, C. (2015). Performance of ACR lung-RADS in a clinical CT lung screening program. *Journal of the American College of Radiology, 12*(3), 273–276. https://radiopaedia.org/articles/lung-rads.

Messay, T., Hardie, R. C., & Rogers, S. K. (2010). A new computationally efficient CAD System for pulmonary nodule detection in CT imagery. *Medical Image Analysis, 14*(3), 390–406.

Messay, T., Hardie, R. C., & Tuinstra, T. R. (2015). Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the lung image database consortium and image database resource initiative dataset. *Medical Image Analysis, 22*(1), 48–62.

Moltz, J. H., Bornemann, L., Kuhnigk, J-M., Dicken, V., Peitgen, E., Meier, S., et al. (2009). Excluído-advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans. *IEEE Journal on Selected Topics in Signal Processing, 3*(1), 122–134.

Mukhopadhyay, S. (2016). A segmentation framework of pulmonary nodules in lung CT images. *Journal of Digital Imaging, 29*(1), 86–103.

Rao, R. V., Savsani, V. J., & Vakharia, D. P. (2011). Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems. *Computer-Aided Design, 43*(3), 303–315.

Song, J., Yang, C., Fan, L., Wang, K., Yang, F., Liu, Sh., et al. (2016). Lung lesion extraction using a toboggan based growing automatic segmentation approach. *IEEE Transactions on Medical Imaging, 35*(1), 337–353.

Stewart, B. W., & Wild, C. P. (2014). *Cancer Worldwide* (pp. 16–69). IARC. World cancer report 2014.

Sun, S., Guo, Y., Guan, Y., Ren, H., Fan, L., & Kang, Y. (2014). Juxta-vascular nodule segmentation based on flow entropy and geodesic distance. *IEEE Journal of Biomedical and Health Informatics, 18*(4), 1355–1362.

Tan, Y., Schwartz, L. H., & Zhao, B. (2013). Segmentation of lung lesions on CT scans using watershed, active contours, and markov random field. *Medical Physics, 40*(4), 43502.

Tsao, A. S. (2007). Lung carcinoma: Tumors of the lungs. *Merck Manual Professional Edition, Online Edition.*

Tu, X., Xie, M., Gao, J., Ma, Zh., Chen, D., Wang, Q., et al. (2017). Automatic categorization and scoring of solid, part-solid and non-solid pulmonary nodules in CT images with convolutional neural network. *Scientific Reports, 7*(1), 8533.

Wang, S., Zhou, M., Liu, Z., Liu, Z., Gu, D., Zang, Y., et al. (2017). Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. *Medical Image Analysis, 40*, 172–183.

Yang, X. S. (2008). *Nature-inspired metaheuristic algorithms* (2nd Edition). Luniver Press (2010).

Ye, X., Beddoe, G., & Slabaugh, G. (2010). Automatic graph cut segmentation of lesions in CT using mean shift superpixels. *International Journal of Biomedical Imaging, 2010.*

Zäpfel, G., Bögl, M., & Braune, R. (2010). *Metaheuristic search concepts, a tutorial with applications to production and logistics* (pp. 67–155). Berlin, Heidelberg: Springer Berlin Heidelberg. Chp. 4–8.

Zhao, J.-J., Ji, G., Xia, Y., & Zhang, X.-L. (2015). Cavitary nodule segmentation in computed tomography images based on self–generating neural networks and particle swarm optimisation. *International Journal of Bio-Inspired Computation, 7*(1), 62–67.