Multimedia Systems (2022) 28:1151–1163
https://doi.org/10.1007/s00530-020-00698-x

SPECIAL ISSUE PAPER

Floor of log: a novel intelligent algorithm for 3D lung segmentation in computer tomography images

Solon Alves Peixoto[1] · Aldísio G. Medeiros[1] · Mohammad Mehedi Hassan[2] · M. Ali Akber Dewan[3] · Victor Hugo C. de Albuquerque[4] · Pedro P. Rebouças Filho[1]

Received: 8 July 2020 / Accepted: 20 September 2020 / Published online: 15 October 2020

Abstract
This work presents a high-performance approach for 3D lung segmentation tasks in computer tomography images using a new intelligent machine learning algorithm called Floor of Log(FoL). The Support Vector Machine was used to learn the better parameter of the FoL algorithm using the parenchyma and its border as labels. Sensitivity, Matthews Correlation Coefficient (MCC), Hausdorff Distance (HD), Dice, Accuracy (ACC), and Jaccard were used to evaluate the proposed algorithm. The FoL was compared with recent 3D region growing, 3D Adaptive Crisp Active Contour, 3D OsiriX toolbox, and Level-set algorithm based on the coherent propagation method algorithms. The FoL algorithm achieves good results with approximately 19 s in the most significant result in an exam with 430 slices and presents similarity indexes achieving HD 3.5, DICE 83.63, and Jaccard 99.73 and qualitative indexes achieving Sensitivity 83.87, MCC 83.08, and ACC 99.62. The proposed approach of this work showed a simple and powerful algorithm to segment lung in computer tomography images of the chest region by combining similar textures, highlighting the lung structure. The FoL was presented as a new supervised clustering algorithm which can be trained to achieve better results and attached to other approaches as Convolutional Deep Neural Networks applications.

Keywords Lung Segmentation · Image Processing · Clustering · Floor of Log · Deep learning

1 Introduction

In recent years, we have witnessed a significant growth of multimedia healthcare data in the form of text, computer tomography images, audio, video and so forth [16, 17, 44]. They are used in various healthcare applications such as 3D lung segmentation task in pulmonary pathologies [15, 33, 48]. The growing applications related to the detection of pulmonary pathologies have called attention worldwide. The patients with asthma overtake 300 million where some sudden symptoms are responsible for a big part of this value. Some factors involve some additional risk for the apparition of pulmonary diseases, including pollution, lifestyle related to sedentarism, smoking, which significantly affects the ratio of the number of lung pathologies per year [2].

✉ Mohammad Mehedi Hassan
mmhassan@ksu.edu.sa

Solon Alves Peixoto
solon.alves@lapisco.ifce.edu.br

Aldísio G. Medeiros
aldisio.medeiros@lapisco.ifce.edu.br

M. Ali Akber Dewan
adewan@athabascau.ca

Victor Hugo C. de Albuquerque
victor.albuquerque@ieee.org

Pedro P. Rebouças Filho
pedrosarf@ifce.edu.br

[1] Laboratório de Processamento Digital de Imagens, Sinais e Computação Aplicada, Instituto Federal de Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), Ceará, Brazil

[2] Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

[3] School of Computing and Information Systems, Faculty of Science and Technology, Athabasca University, Edmonton, Canada

[4] Programa de Pós-Graduação em Informática Aplicada, Laboratório de Bioinformática, Universidade de Fortaleza, Fortaleza-CE, Brazil

Some signals of that patience have the presence of the lung associated pathologies includes chronic sinusitis, degradation of the elastic movement of lungs, cough, fatigue and weakness. The severity of these pathologies requires fast methods that increase the prognosis performance and the treatment of the patient. In addition, a high demand of patients implies in necessary methods which help in aid of clinical diagnoses, avoinding the possibility erros from human physiological factors [6] and other technologies capable of interconnect such informations [9, 40, 41].

Recently deep learning-based algorithms have the capability to handle huge amounts of multimedia healthcare data and they are capable of extracting sufficient data for its analysis and have been created to meet this requirement in different medical scenarios [3, 10, 13, 30, 31, 38, 39, 42]. Some of these methods focus on the detection of nodules in pulmonary region [31], and the analysis of lung pathologies [31] as well as the study in other diseases such as Parkinson prognoses [43].

In bioimaging tasks related to lung segmentation, the detection of the pulmonary region is essential and computed tomography (CT) and magnetic resonance imaging (MRI) scans are important tools to get those [12]. These detection tasks normally are related to the a priori knowledge, in this case, the characteristics of an injury, as well as its shape and delimitation, are based on the knowledge of a specialist [5].

Maier et al. [19] presents a benchmark with comparison of algorithm related to ischemic stroke tasks in MRI exams. In his work, the development of approaches that combines methods to emphasize different characteristics presents the best results where the convolutional neural networks (CNN) are the most common approach to extract multiple patterns.

Rebouças Filho et al. [29] presented an powerful method using Active Contour Method called Adaptive Crisp Active Contour Method (ACACM). The results showed both the speed and effectiveness against commercial softwares. One of most important combination of this method is the inclusion of MultiLayer Perceptron in the algorithm making the shape adapt to the image with some intelligence [11].

Nithila and Kumar [21] also presents a method based on Active Contour model called Selective Binary and Gaussian Filtering-new Signed Pressure Force, where it can detect borders of parenchyma of lungs with much more efficient by highlighting the difference between the internal and external side of lungs.

Zhang et al. [50] presents an improved GrabCut algorithm for lung parenchyma segmentation. The main contribution of this work lies in optimizing the steps by automatically select the lung parenchyma which can be achieved by using a threshold OTSU algorithm.

Other important fields of related works are associated with lung nodules segmentation. In the work presented by Shakibapour et al. [35] a non-supervised approach based on metaheuristics is used for segmentation of pulmonary nodules in lung CT scans. Shakir at al. [36] extend the nodule segmentation to an 3D segmentation approach based on geodesic active contour.

Deep neural networks is also an powerful tool for segmentation process [49]. Skourt et al. [37] use the U-net architecture to compress and extract the data information. This process allows the model to learn the most important information related to the lung and later use this information to segment other lungs.

Thus, one can see that MRI or CT scans have some patterns with high discriminatory characteristics , preserving the structural information that helps the diagnoses of a lung associated pathology, mostly when the application highlight the specific desired pattern related to the disease.

This paper proposes a high-performance approach for 3D lung segmentation called Floor of Log(FoL). This clustering algorithm was previously tested in facial recognition study [24], where the features extracted from deep neural network method called Facenet [32] were compressed and evaluated for matching tasks, achieving the same results, but with 80–90% of data size reduction. In this work, the FoL was separated from the deep neural network approach and evaluated solely, focusing make clusters the main structures of the pulmonary region.

The most important impact of the proposed model lies in the increase of the optimization process to segment the pulmonary region of the CT scan using just the logarithm function and the 'floor' operator, making the lung almost segmented. Another important point lies in the using of machine learning in highlighting the clustering pattern of pulmonary region, different from other methods that focus on specific structures, loosing important characteristics that implies in a decrease results in the segmentation task.

This work evaluates the segmentation results by FoL using images of the CT scan of the chest region. The Similarity indexes, Qualitative indexes and Performance time was used as metrics for evaluating the proposed approach.

Thus, this work focuses on:

1. A new approach to lung segmentation in CT.
2. The analysis of lung structure and how to preprocess those structures.
3. The learning process behind the FoL algorithm.
4. Comparison with other segmentation methods.

The paper is organized as follows: Sect. 2 explains the image acquisition in Sect. 2.1 following by the Sect. 2.2 where the proposed method is presented. The Sect. 2.3 presents the process used to train the algorithm to better cluster the pulmonary region and how its output help in the Sect. 2.4. Finally, the Sect. 3 presents the results of this approach for

lung segmentation tasks and the Sect. 3.1 includes a discussion of these results.

2 Methodology

This section aims to help the understanding of the process executed to produce the results obtained. The Sect. 2.1 demonstrates the types of images used for the proposed work. The Sect. 2.2 presents a new method for clustering lung data in computed tomography images. A process for training the proposed method parameter has also been added and is presented by Sect. 2.3. Finally, the Sect. 2.4 section associates the clustering method with the segmentation step itself. The Sect. 2.5 section presents the key metrics for qualitative and comparative results. All the experiments and generated results were run in a Python 3.5 with using scikit-learn library in a hardware containing a Intel Core i7-8700 3.20 Ghz, 16 Gb RAM, 64 Bits Windows Operating System.

2.1 Images acquisition

An experimental process was made in LIDC/IDRI database using the CT scans of LUNA16 challenge [34]. The dataset of LUNA16 offers a large quantity of high-quality images with standard acquisition configurations, resulting in a better control of features like brightness and contrast of all CT scans.

Though the segmentation process indicates good results in LUNA16 dataset, at the presence moment of this work, this database does not have the ground truth of segmented lungs to apply the evaluation metrics.

For a comparison results , this work use the LAPISCO-LUNG database which can be accessed in https://lapisco.ifce.edu.br/producao-academica/private-datasets/database-lungs-from-chest-ct-scans/. This dataset was composed by 40 CT scan images of thorax region with all ground truth available.

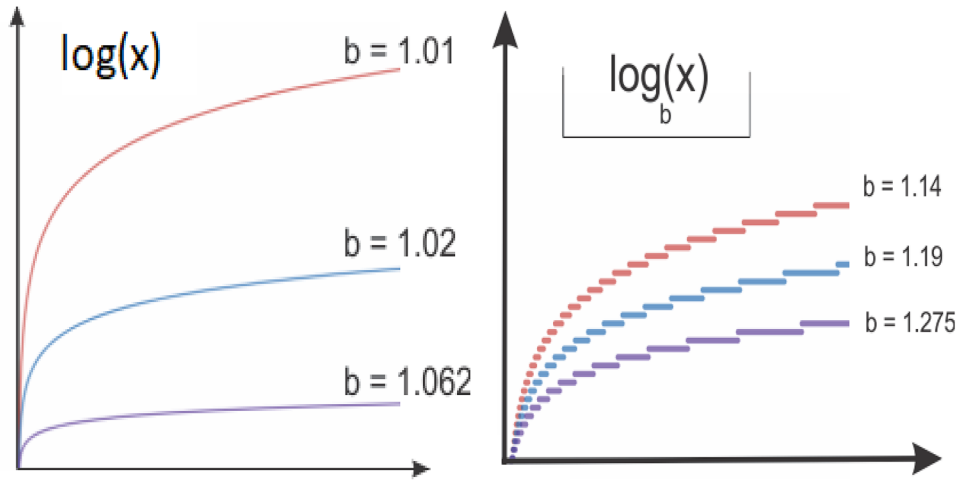
Approach proposed is evaluated from image set obtained in partnership with the Walter Cantídio Hospital of the Federal University of Ceará. Studies in this dataset were autorized by UFC Research Ethics Committee—COMEPE, protocol 35/06.

This image set has been previously evaluated for disease classifications [1, 23, 26, 28], automated recognition of lung [7, 8, 20], lung disease detection [27]. All samples were collected from the equipment Toshiba Aquilion GE Medical System LightSpeed16 and Philips Brilliance 10. The set is formed by 72 lung images, 512×512 pixels with 16 bits each. Segmentation ground truth was produced by a manual analyze from an experienced expert.

2.2 Floor of Log

This work proposes an important innovation in lung segmentation in computed tomography. Clustering of the data from its compression, this technique allows the object of interest to be highlighted in the other regions, thus allowing a much faster and more efficient segmentation.

From a data set, a logarithmic function is applied to apply a linear transformation to the data. Because it is a logarithmic function, this transformation is able to approximate values close to each other. Then, a floor function is applied to this logarithmic function, so only the lower limit of this function will be preserved, i.e its entire part. Figure 1 presents in the item 1a the logarithmic function different values of base and the item 1b presents the floor of the logarithmic function. Note that depending on the logarithmic base used, different types of grouping of the data occur.



(a) Logarithm function with different bases (b) Floor of Log with different bases

Fig. 1 Effect of floor applied in the logarithm function with different bases values. In **a** some logarithms functions with different bases. In **b** the floor function applied to logarithm with different bases values

Since in this work the data are pixel values, they will obtain similar values after the application of the step function that will remove the rational part of the data. The Eq. 1 below represents the main function of the clustering method, where b is the logarithmic basis of the log function. This process can be observed in Fig. 2, in which it presents all steps of transforming the data to obtain clustering.

$$f(x) = \lfloor \log_b(x+1) \rfloor \tag{1}$$

Note that when the Eq. 1 is applied to the data, a linear transformation occurs and, like other i.e. sigmoid transforms, it depends on an adjustment to its slope, in this case, the logarithmic basis. The logarithmic basis should be the most adequate to better separate the data and avoid that when applying the floor function, the data of different "classes" are clustered at the same point.

Therefore, since an object of interest varies according to the type of problem, a training step is required for this logarithmic base, this step will be presented in Sect. 2.3.

Even with the application of the logarithmic transform, some artifacts or noises in the image can disrupt the clustering process. Thus, a filter or pre-processing may assist in grouping the pixels in an image in a local manner so as to attenuate homogenizing the region of interest as the outer region and only then will the algorithm be applied. In this work, we added the Local Binary Patterns (LBP) to the original image, obtaining the result shown in Fig. 3.

Ojala et al. [22] presents the LBP as feature extractor. The image in this method can be analyzed in texture terms, with 8-bit code for each pixel after the texture analyses of neighboring pixels over the central value. The combination of the result of the pixels that which are bigger or smaller of the central pixel were used to represent a texture value.

Using the LBP rotation-invariant setting, local pixels added to the original image tend to approach smoothly without causing image damage. From the histogram below the images, it was observed that there is a significant increase in the peaks of each sample of pixels, that is, the image has undergone a contrast localization adjustment. In the Result column, we can see the output of this image when applied to the clustering process, described earlier.

2.3 Training algorithm

Although the clustering process is quite simple, a good clustering result will depend directly on the logarithmic base used. Some drawbacks related to FoL lies in the correct log base selection since the clustered Hounsfield Units may not present a fully connection of lung parts, like in advanced DPOC disease.

Those base selection can also include the trachea in the clusterization process, decreasing the algorithm performance. Thus, this topic presents the proposed methodology to find this parameter.

Using a set of images segmented by a medical specialist, we can use supervised techniques of artificial intelligence and analyze different values for the logarithmic base and thus choose a more appropriate value. Note that for this type of approach to be used, a cost function is normally required and thus simulate the downward gradient.

The creation of the cost function mainly involves the definition of the labels used for the pixel intensities of the object of interest. Therefore, since computed tomography has values of human tissues in Hounsfield units, these values will be separated to highlight the difference between regions of the lung and non-lung region, and therefore, can be interpreted as a separation problem binary.

A classic approach would be to apply the masking cleavage made by the specialist directly to the original image and use the values associated with the mascara as a lung class and the background as a non-lung class as described on the right side of Fig. 4.

This approach, however, may cause some separation problems in the classification algorithms, since the inner region of the lung is aerated but the patient's surroundings as well. Thus, we will have two regions with the same values representing different classes.

To solve this problem, this work uses only the mask contour as the background, and so the non-lung class is now related to the lung parenchyma as presented in Fig. 5, thus providing values that are sufficiently different to be analyzed in a training step.

Finally, the training step tends to iteratively minimize the cost function, adding the base a fixed value of *step*. Note that if the base is being updated positively or increasing and the value of the cost function has no change, it is interesting to use the largest value of the log base as it will further compress the data, thus avoiding unnecessary clusters. The algorithm that defines the training process can be demonstrated according to Algorithm 1.

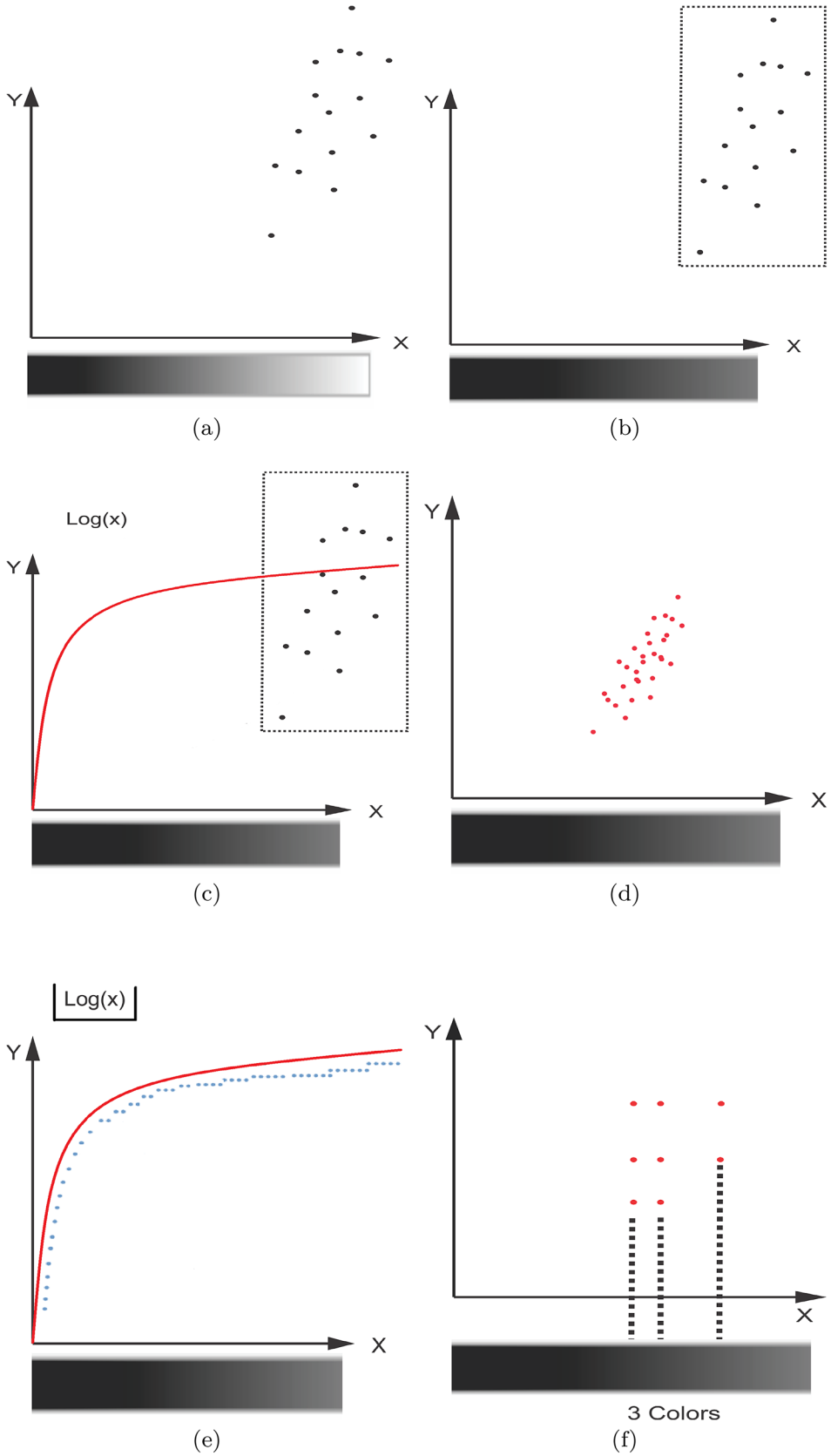


Fig. 2 Process for Data Clustering. **a**, **b** presents the Data used in the clustering process. The logarithm function is applied in the data (**c**), transforming it to closer values (**d**). After this transform, in **e** the blue dots represent the floor function applied in log function compressing the data, resulting in a few clusters of the data **f** (colour figure online)

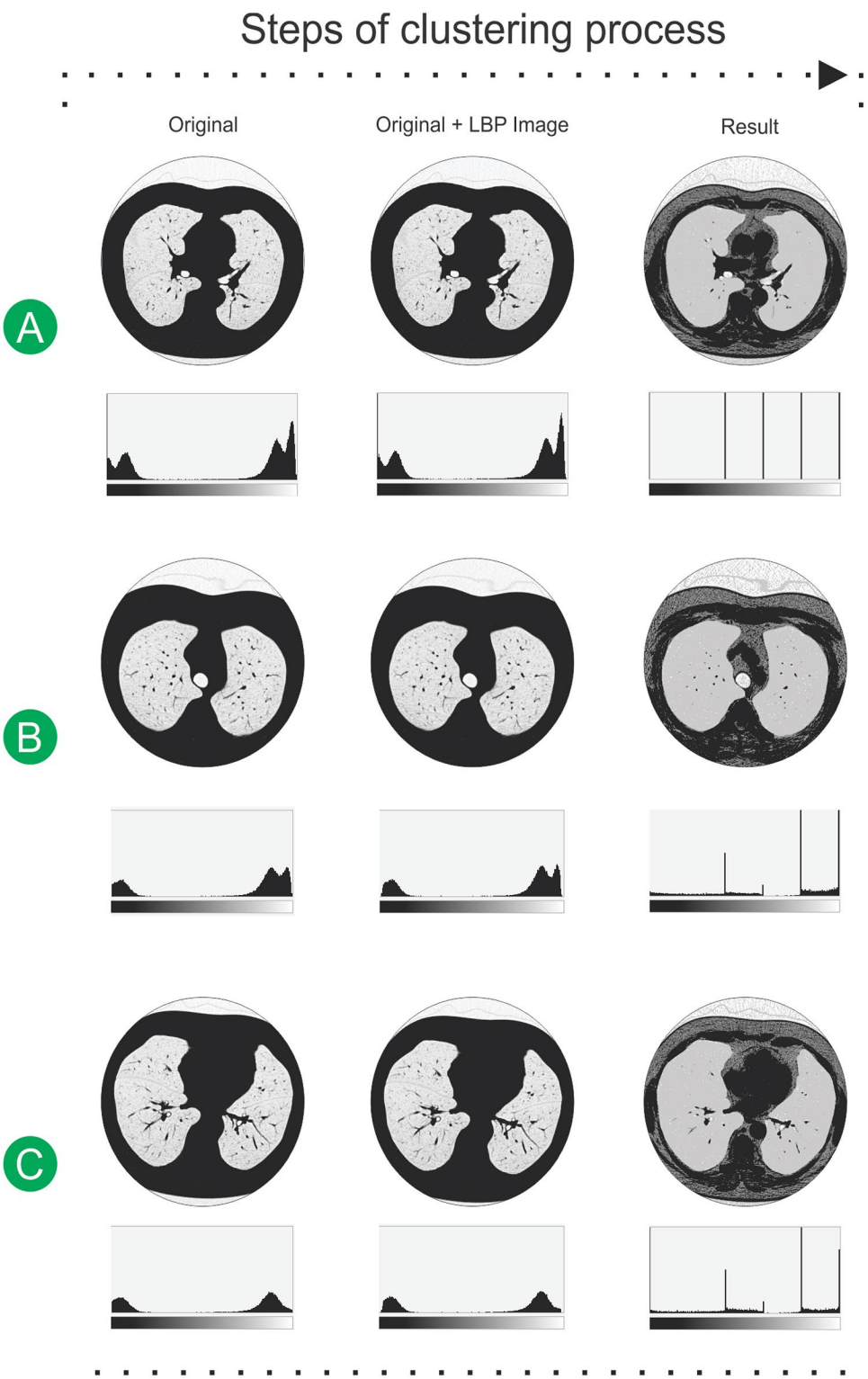


Fig. 3 Steps of clustering process. The **a**--**c** present some examples of the proposed method. After the input Image, this was combined with LBP image and later clustered by using the Floor of Log algorithm. The column Result presents the normalized image after the Floor of Log operation

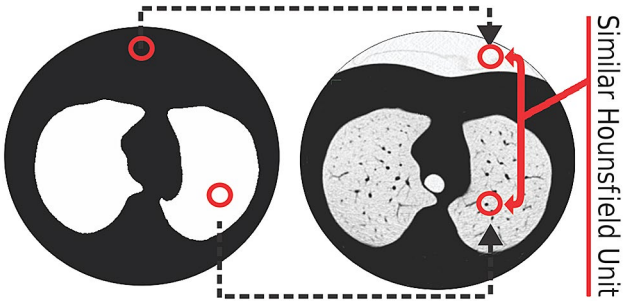


Fig. 4 Problems of in the creation of dataset used in the training process. The background and the ROI may have same values

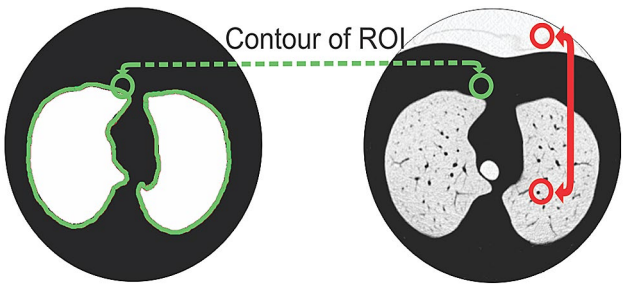


Fig. 5 The green contour in the middle image presents the background used in this work

In this work, the Support Vector Machine (SVM) as classifier was applied in the presented algorithm,specifically in performing cross-validation

This classifier creates a hyperplane that tries to separate the most important features of each class. This hyperplane can be interpreted as an approximation function that discriminates the data. Even with the presence of artifacts or noises, this classifier method is one of the most powerful tools in the machine learning area [45].

2.4 Segmentation

After the training analysis, the logarithmic basis will be used in the Eq. 1 and thus the entire dataset will be clustered. Note that each cluster offers a unique value that represents a portion of the data and, in this sense, we will extract from the tomography images the value corresponds to the clustered region of the lung. Even so, some regions remain in focus, as in other regions gas accumulation can also occur, such as the intestine.

Another fact known by other methods of lung segmentation is the presence of the ilium. This structure is usually present in the segmented images because it is directly connected to the lung as well as aerated. To cope with this

Algorithm 1 Proposed Algorithm to find a good logarithm base.

Input(data, step, $nInterations$, classifier , kFold)
$lastScore \leftarrow 0$
$bestBase \leftarrow base \leftarrow 1$
for $i \leftarrow 0$ to $nInterations$ **do**
  $cluterizedValues \leftarrow$floor(log($data, base$))
  x , $y \leftarrow$removeDuplicates($cluterizedValues$)
  stats $\leftarrow$crossValidation($classifier$,x,y, $kFold$)
  $meanScore \leftarrow$mean($stats.accuracy$)
  if $meanScore \geq lastScore$ **then**
    $lastScore \leftarrow meanScore$
    $bestBase \leftarrow base$
  end if
  $base \leftarrow base + step$
end for

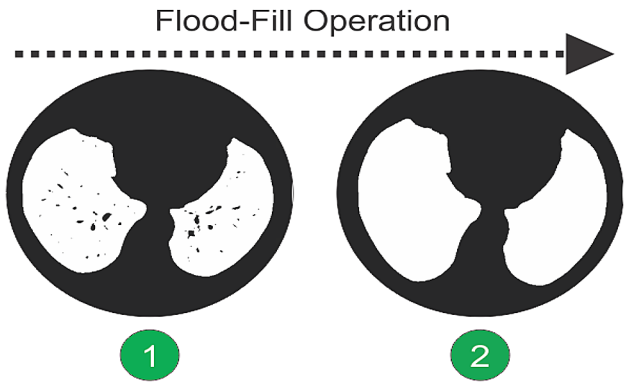


Fig. 6 Flood fill operation applied in image 1. The Result was presented by 2

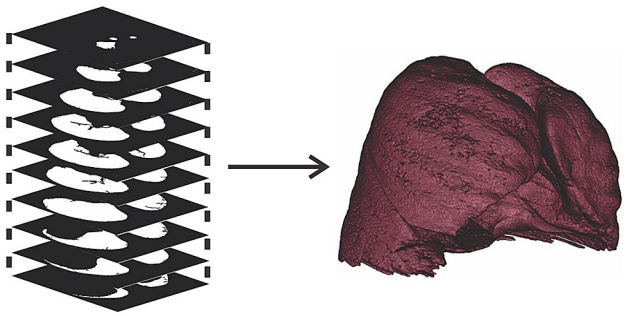


Fig. 7 The three-dimensional representation of the pulmonary region. After the process of each slice, the algorithm extract the entire lung in the third dimension by using the 2 bigger volumes

problem, this work used morphological techniques, specifically erosion, in which it can be used to detach objects and, therefore, to separate the ilium from the pulmonary structure. The regions of the blood vessels and airways can be complemented with a simple opening and closing functions, as illustrated by Fig. 6.

Finally, since the objects are still separated by clusters. We can segment the lung in a single operation, since it, in a three-dimensional space, has the largest volume. Therefore, it is enough that the two objects with the largest volume (left and right lung) in the three-dimensional space are extracted to have the lung segmented, finishing the process. Thus, it is necessary that each of the processed images be concatenated in sequence, to analyze the entire image in a three-dimensional space, as indicated by Fig. 7.

2.5 Evaluation metrics

To perform the comparison, false negative (FN), true positive (TP), false positive (FP) and true negative (TN) was extracted from confusion matrix of the image from the proposed approach and the specialist segmentation. Each one of these computed values is used in the evaluation metrics, presented below:

The ratio of the total errors and the correct segmentation possibility was evaluated by Accuracy (Acc) giving a qualitative index of how, in geral terms, the segmentation result was presented. Its formula is given by:

$$\mathrm{Acc} = \frac{\mathrm{TN} + \mathrm{TP}}{\mathrm{TN} + \mathrm{FP} + \mathrm{TP} + \mathrm{FN}} \tag{2}$$

Another qualitative index is presented by The sensitivity (Se) where it metric lies in the capacity to correctly predict a region in comparison with a set of regions that really belongs to pulmonary region presented by:

$$\mathrm{Se} = \frac{\mathrm{VP}}{\mathrm{VP} + \mathrm{FN}} \tag{3}$$

Another statistic metric is the Matthews Correlation Coefficient (MCC). In this case, the qualitative result between the proposed approach and the specialist segmentation are also analyzed even with different scale of those images. MCC correlation is calculated as presented below:

$$\mathrm{MCC} = \frac{(\mathrm{TN} * \mathrm{TP} - \mathrm{FN} * \mathrm{FP})}{\sqrt{(\mathrm{FP} + \mathrm{TP}) * (\mathrm{FN} + \mathrm{TP}) * (\mathrm{FP} + \mathrm{TN}) * (\mathrm{FN} + \mathrm{TN})}}, \tag{4}$$

For the similarity comparison, the Hausdorff distance [14] was included for analysis. Using the Hausdorff space, the maximum Euclidean distance subject to the segmented region by the proposed approach to the GT and the maximum Euclidean distance subject to GT in relation to segmented region is analysed. One can see that if the segmented region of both images are similar, this distance tends to be equal. The HD formula is presented by:

$$\mathrm{HD} = \max\left\{A, B\right\}, \tag{5a}$$

Where

$$A = \sup_{x \epsilon R_a} \inf_{y \epsilon R_{gt}} \mathrm{d}(x, y) \tag{5b}$$

$$B = \sup_{y \epsilon R_{gt}} \inf_{x \epsilon R_a} \mathrm{d}(x, y) \tag{5c}$$

The Dice coefficient (DC) [4], is also a metric used to verify the similarity between segmented regions. This method uses a overlap of images , presented as:

$$\mathrm{Dice} = 2 \cdot \frac{|S_a \cap S_{gt}|}{|S_a + S_{gt}|}, \tag{6}$$

where S_a stands for algorithmic segmentation, i.e, the segmented region by the proposed approach, and S_{gt} is the

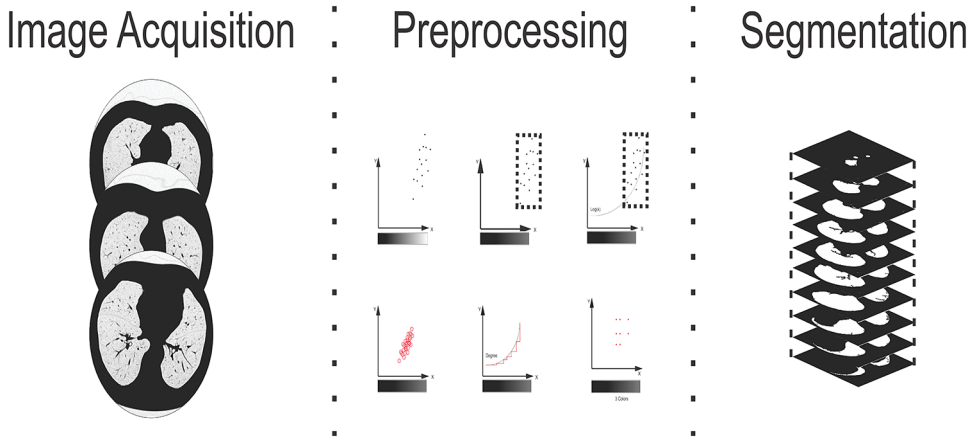


Fig. 8 Proposed methodology. In Image Acquisition steps, Lung images from CT Scan dataset are initially stored and separated by its classes. The Preprocessing column presents the newly proposed clustering method and the Segmentation step in the third column presenting the image segmentation and the construction of the 3D view

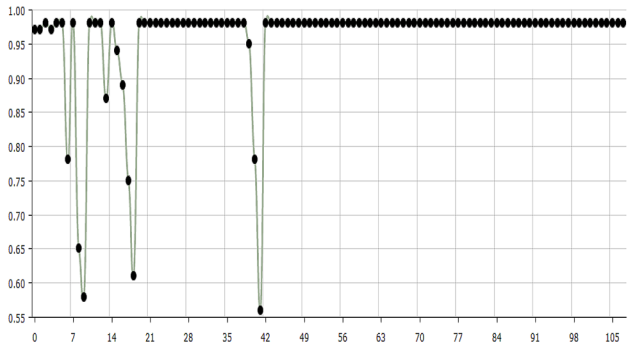


Fig. 9 Evaluation of learning process by proposed training algorithm with 10 Cross-validation mean results using 105 iterations. At each iteration, the algorithm changes the base 'b' used in FoL

region selected by the specialist. These overlapped regions allow the extraction of binary values after the comparison of each pixel in both regions. One can see that if the value 1 corresponding to the perfect overlap at each pixel, the sum of total matching corresponds to the higher similarity to specialist segmentation [25].

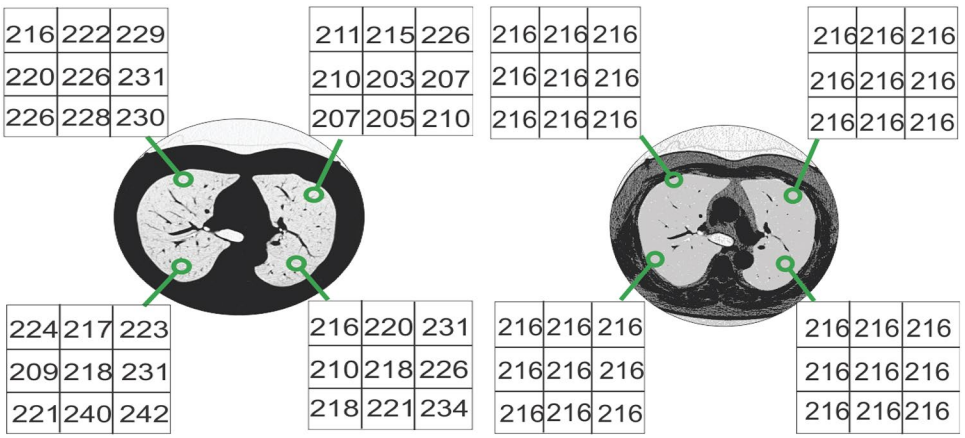
3 Results

Considering the Steps presented in Fig. 8 and the images from CT scan dataset, this section is divided into 2 main blocks: first, a presentation of the evaluation of the new proposed approach for 3D lung segmentation following by a second step showing a comparison with state-of-art methods related to 3D lung segmentation.

Showing each iteration of the learning process, Fig. 9 presents the mean accuracy after 10 Cross-validation in the learning process of the FoL algorithm, which needs to train its logarithm base to best fit the clustering regions.

For experimental approach, the Fig. 10b indicates a single-value of graylevel and some presence of holes that constitutes the vessels inn LUNA16 dataset . Those vessels are easily and fast removed by using flood-fill morphological operator as indicating in Fig. 11b.

The pre-segmented process is presented in Fig. 11 where the pixel value of the clustered regions are used to extract the aerated components. The Fig. 11a presents the results of the clusterized region using FoL algorithm.



(a) Normalized Original Image (b) Normalized Clustered Image

Fig. 10 Effects of the FoL in LUNA16 dataset. The original values presented in the image **a** were converted in the unique value of aerated region **b** after FoL method

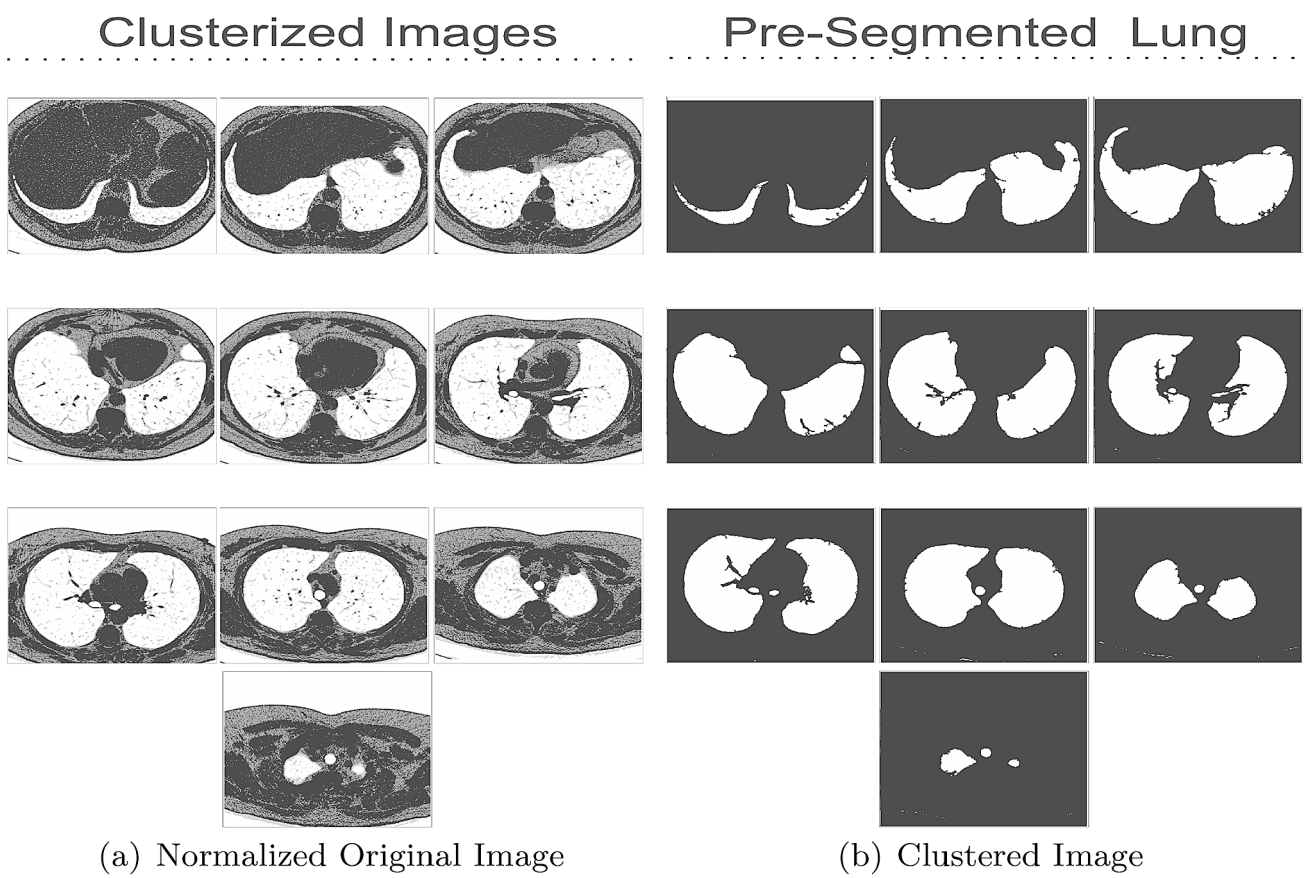


Fig. 11 Clustering process in LAPISCO dataset. After the FoL algorithm application, the pulmonary region was extracted by thresholding for a unique pixel value

Using this approach, the aerated region was extracted selecting the pixel value correponding this area, as showed by Fig. 11b.

Finally, using all the exam from slice, the lung was extracted in three-dimensional form, by selecting the 2 highest volumes, this approach is necessary to remove eventual artifacts as some unwanted gases and the trachea. The result of this operation is presented in Fig. 12.

From Tables 1, 2, 3, this work present the Average-Time of the segmentation task, the Similarity analysis and the Qualitative results, respectively. In all results, the FoL method was compared with methods normally used in this segmentation task: 3D Adaptive Crisp ACM [29], the automatic 3D Region Growing (3D RG) algorithm, the level-set algorithm based on the coherent propagation method (LSCPM) [46, 47], and the semi-automatic segmentation performed by an expert using the 3D OsiriX toolbox (EUOT) [18].

3.1 Overall discussion of the results

The Results generated related to 3D lung segmentation follows the sequence presented in Fig. 3.

By using the proposed method in an experimental approach, one can see that even with the presence of artifacts (in this case, the LUNA16 have some lungs with nodules indicating cancer), those items does not interferes negatively since its structure is also clustered with the neighbourhood.

An important difference of this database is in the requirement of a normalization of the CT scan images for most of the automatic methods since the Hounsfield units have a high variation between CT images.

Using the dataset offered by LAPISCO, where we can compare with the segmentation of a specialist, and the metrics discussed in Sect. 2.5.

The base of logarithm function used in FoL method was learned using the SVM as classifier and the proposed algorithm as presented in Algorithm 1. The Fig. 9 shows the mean result of 10 cross-validation after 105 iterations, using 0.01 as step. One can see that exist a high range of values of

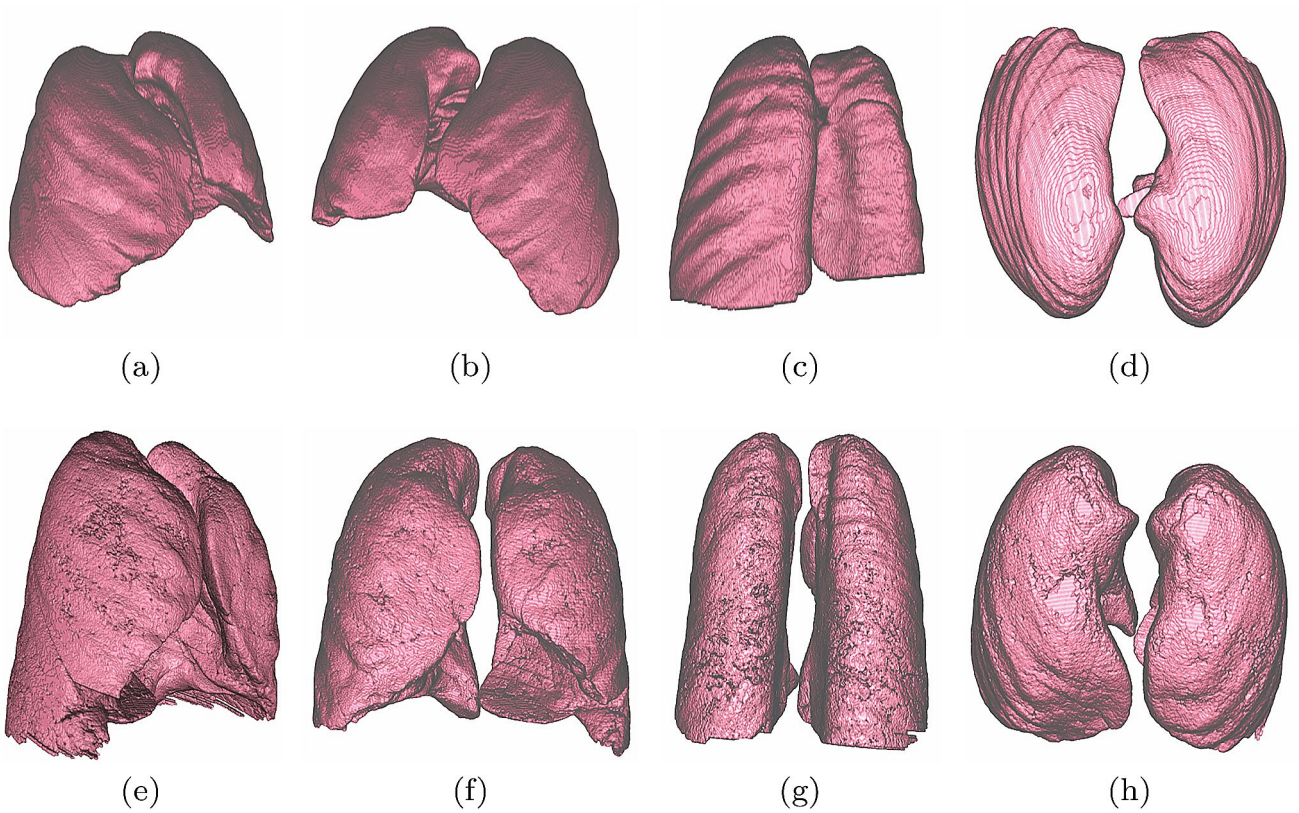


Fig. 12 2 examples of lung segmentation. After the extraction of the 2 highest volumes in the combination of the entire exam, the segmentation process is ended

Table 1 Segmentation average time of FoL against other methods

Average time	
Method	Time (s)
FoL	19 ±1.12
ACM 3D	441 ±0.11
3D RG	334 ±0.16
EUOT	643 ±0.11
LSCPM	235 ±0.11

Table 2 Similarity indexes to FoL against other segmentation methods

Similarity indexes			
Method	HD	DICE	Jaccard
FoL	**3.512±0.31**	**83.63 ±01.93**	**99.73 ±01.33**
ACM 3D	3.31±0.23	83.22 ±02.82	99.78 ±01.52
3D RG	3.71 ±0.63	82.33 ±01.84	99.66 ±01.41
EUOT	3.74±0.74	82.95 ±02.31	99.61 ±01.24
LSCPM	3.62±0.25	80.91 ±01.72	99.43 ±02.12

Best results are in bold

Table 3 Quantitatives indexes to FoL against other segmentation methods

Quantitative indexes			
Method	Se	Acc	MCC
FoL	**83.87±00.66**	**99.62±00.46**	**83.08±01.92**
ACM 3D	82.71±01.31	99.78±01.29	83.40±03.20
3D RG	83.01±05.10	99.66±03.61	82.79±06.45
EUOT	83.62±05.10	99.73±03.61	82.37±06.45
LSCPM	84.81±05.10	99.43±03.61	80.34±06.45

Best results are in bold

the logarithm base that establishes the classification results. This work used 1.72 as logarithm base.

After the training process, the FoL was applied to transform the original image into clusterized images as presented in Fig. 11a. Since the pulmonary region is a unique pixel value region, the lung can be extracted entirely by threshold this pixel value. The result of this operation was presented in Fig. 11b.

The performance was presented in Table 1. The LSCPM method achieves 235 s in average time over the segmentation task followed by 334 s with the RG 3D. The ACM 3D model approach spent 441 s, which is a little bit more since it uses machine learning approach inside to analyze the neighborhood. The EUOT expended the most time, which is expected since is a user-dependent model. The FoL approach drops its value with a huge gain in speed by achieving a just 19 s in the average time.

Table 2 presents the results concerning similarity. Note that there was a substantial growth of FoL compared with the others approaches in the metrics indicating good segmentation values when compared to a result of a specialist.

In Table 3, the values of Sensitivity, Accuracy and MCC are presented as qualitative measures. In this sense, it is evaluated the ability of the method to correct or not the regions related to the lung and not the similarity of the method over that of a specialist. One can see that FoL achieves almost the same values against the other methods with slightly better results.

4 Conclusion and future work

The proposes of this work showed a innovative and powerful approach to segment lung in CT images of the chest region. The FoL was presented as a new clustering algorithm which can be trained to achieve better results.

The results presented a high score rate with Se 83.62%, Acc 99.62%, MCC 83.08%, HD 3.51%, DICE 83.63% , Jaccard 99.73% in lung segmentation. On the other hand, the FoL reinforce the currents Hounsfield-based techniques used by the specialist and can be used as an important tool in CAD system to help in the diagnosis of diseases.

As a future work, this method will be combined with CNN to evaluate if this approach can effectively cluster the data or even its features to speed up the segmentation methods that uses CNN.

Acknowledgements The authors are grateful to King Saud University, Riyadh, Saudi Arabia for funding this work through Researchers Supporting Project number RSP-2020/18. This study was also supported in part by the CAPES - Finance Code 001, and by the CNPq via Grants Nos. 311973/2018-3 and 430274/2018-1.

References

1. Alves, S.A., Rebouças, E.S., Oliveira, S.F., Braga, A.M., Rebouças Filho, P.P.: Lung diseases classification by analysis of lung tissue densities. IEEE Latin Am. Trans. (2018)
2. Chatterjee, R., Maitra, T., Islam, S.H., Hassan, M.M., Alamri, A., Fortino, G.: A novel machine learning based feature selection for motor imagery eeg signal classification in internet of medical things environment. Fut. Gen. Comput. Syst. **98**, 419–434 (2019)
3. Chouhan, V., Singh, S.K., Khamparia, A., Gupta, D., Tiwari, P., Moreira, C., Damaševičius, R., De Albuquerque, V.H.C.: A novel transfer learning based approach for pneumonia detection in chest X-ray images. Appl. Sci. **10**(2), 559 (2020)
4. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945)
5. Fang, R., Gupta, A., Huang, J., Sanelli, P.: Tender: tensor non-local deconvolution enabled radiation reduction in ct perfusion. Neurocomputing **229**, 13–22 (2017)
6. Félix, J.H.d.S.: Metodos de contornos ativos hilbert 2d na segmentação de imagens dos pulmões em tomografia computadorizada do tórax. Ph.D. thesis, Universidade Federal do Ceará, Fortaleza, CE (2011)
7. Filho, P.P.R., Barros, A.C.d.S., Ramalho, G.L.B., Pereira, C.R., Papa, J.P., de Albuquerque, V.H.C., Tavares, J.M.R.S.: Automated recognition of lung diseases in ct images based on the optimum-path forest classifier. Neural Comput. Appl. (2017). 10.1007/s00521-017-3048-y
8. Filho, P.P.R., de S. Rebouças, E., Marinho, L.B., Sarmento, R.M., Tavares, J.M.R., de Albuquerque, V.H.C.: Analysis of human tissue densities: A new approach to extract features from medical images. Pattern Recogn. Lett. (2017). http://dx.doi.org/10.1016/j.patrec.2017.02.005
9. Gumaei, A., Hassan, M.M., Hassan, M.R., Alelaiwi, A., Fortino, G.: A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. IEEE Access **7**, 36266–36273 (2019)
10. Evolutionary algorithms for automatic lung disease detection: Gupta, N., Gupta, D., Khanna, A., Rebouças Filho, P.P., de Albuquerque, V.H.C. Measurement **140**, 590–608 (2019)
11. Hajimani, E., Ruano, M., Ruano, A.: An intelligent support system for automatic detection of cerebral vascular accidents from brain ct images. Comput. Methods Prog. Biomed. **146**, 109–123 (2017)
12. Han, G., Liu, X., Zhang, H., Zheng, G.W., Soomro, N.Q., Wang, M., Liu, W.: Hybrid resampling and multi-feature fusion for automatic recognition of cavity imaging sign in lung ct. Fut. Gen. Comput. Syst. **99**, 558–570 (2019)
13. Hassan, M.M., Alam, M.G.R., Uddin, M.Z., Huda, S., Almogren, A., Fortino, G.: Human emotion recognition using deep belief network architecture. Inform. Fus. **51**, 10–18 (2019)
14. Hausdorff, F.: Bemerkung über den inhalt von punktmengen. Math. Ann. **75**(3), 428–433 (1914)
15. Hossain, M.S.: Cloud-supported cyber-physical localization framework for patients monitoring. IEEE Syst. J. **11**(1), 118–127 (2015)
16. Hossain, M.S., Muhammad, G.: Emotion recognition using deep learning approach from audio-visual emotional big data. Inform. Fus. **49**, 69–78 (2019)
17. Hossain, M.S., Muhammad, G., Alamri, A.: Smart healthcare monitoring: a voice pathology detection paradigm for smart cities. Multimed. Syst. **25**(5), 565–575 (2019)
18. Hunter-Smith, D., Alexandra, R., Spychal, R., Michael, P.: 3D volumetric analysis and haptic modeling for preoperative planning in breast reconstruction. Anaplastology **4**(138), 1173–2161 (2015)
19. Maier, O., Menze, B.H., von der Gablentz, J., Häni, L., Heinrich, M.P., Liebrand, M., Winzeck, S., Basit, A., Bentley, P., Chen, L., et al.: Isles 2015—a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. Med. Image Anal. **35**, 250–269 (2017)
20. Medeiros, A., Peixoto, S., Barros, C., Albuquerque, V., Rebouças, P.P.: Uma nova abordagem para a segmentação de pulmões utilizando o método de contorno ativo não paramétrico optimum path

snakes em imagens de tomografia computadorizada. In: Workshop de Informática Médica, vol. 17 (2017)
21. Nithila, E.E., Kumar, S.: Segmentation of lung from ct using various active contour models. Biomed. Signal Process. Control **47**, 57–62 (2019)
22. Ojala, T., Pietikainen, M., Harwood, D.: Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In: Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision and Image Processing., Proceedings of the 12th IAPR International Conference on, vol. 1, pp. 582–585. IEEE (1994)
23. Peixoto, S., Pedrosa, P., Kumar, A., Albuquerque, V.: Automatic classification of pulmonary diseases using a structural co-occurrence matrix. Neural Computing and Applications (2018). https://doi.org/10.1007/s00521-018-3736-2
24. Peixoto, S.A., Vasconcelos, F.F., Guimarães, M.T., Medeiros, A.G., Rego, P.A., Neto, A.V.L., de Albuquerque, V.H.C., Rebouças Filho, P.P.: A high-efficiency energy and storage approach for iot applications of facial recognition. Image Vis. Comput, 103899 (2020)
25. Qian, X., Wang, J., Guo, S., Li, Q.: An active contour model for medical image segmentation with application to brain ct image. Med. Phys. **40**(2), (2013)
26. Ramalho, G.L.B., Ferreira, D.S., Filho, P.P.R., de Medeiros, F.N.S.: Rotation-invariant feature extraction using a structural co-occurrence matrix. Measurement **94**, 406–415 (2016). https://doi.org/10.1016/j.measurement.2016.08.012
27. Ramalho, G.L.B., Rebouças Filho, P.P., Medeiros, F.A.N.S.d., Cortez, P.C.: Lung disease detection using feature extraction and extreme learning machine. Revista Brasileira de Engenharia Biomedica **30**, 207–214 (2014)
28. Rebouças, P.P., de S. Rebouças, E., Marinho, L.B., Sarmento, R.M., Tavares, J.M.R., de Albuquerque, V.H.C.: Analysis of human tissue densities: A new approach to extract features from medical images. Pattern Recogn. Lett. **94**, 211–218 (2017). https://doi.org/10.1016/j.patrec.2017.02.005
29. Rebouças Filho, P.P., Cortez, P.C., da Silva Barros, A.C., Albuquerque, V.H.C., Tavares, J.M.R.: Novel and powerful 3d adaptive crisp active contour method applied in the segmentation of ct lung images. Med. Image Anal. **35**, 503–516 (2017)
30. Rebouças Filho, P.P., da Silva Barros, A.C., Almeida, J.S., Rodrigues, J., de Albuquerque, V.H.C.: A new effective and powerful medical image segmentation algorithm based on optimum path snakes. Appl. Soft Comput. **76**, 649–670 (2019)
31. Rodrigues, M.B., Da Nóbrega, R.V.M., Alves, S.S.A., Rebouças Filho, P.P., Duarte, J.B.F., Sangaiah, A.K., De Albuquerque, V.H.C.: Health of things algorithms for malignancy level classification of lung nodules. IEEE Access **6**18, 592–601 (2018)
32. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 815–823 (2015)
33. Selvanambi, R., Natarajan, J., Karuppiah, M., Islam, S.H., Hassan, M.M., Fortino, G.: Lung cancer prediction using higher-order recurrent neural network based on glowworm swarm optimization. Neural Comput. Appl. **32**(9), 4373–4386 (2020)
34. Setio, A.A.A., Traverso, A., De Bel, T., Berens, M.S., van den Bogaard, C., Cerello, P., Chen, H., Dou, Q., Fantacci, M.E., Geurts, B., et al.: Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. Med. Image Anal. **42**, 1–13 (2017)
35. Shakibapour, E., Cunha, A., Aresta, G., Mendonça, A.M., Campilho, A.: An unsupervised metaheuristic search approach for segmentation and volume measurement of pulmonary nodules in lung ct scans. Expert Syst. Appl. **119**, 415–428 (2019)
36. Shakir, H., Khan, T.M.R., Rasheed, H.: 3-D segmentation of lung nodules using hybrid level sets. Comput. Biol. Med. **96**, 214–226 (2018)
37. Skourt, B.A., El Hassani, A., Majda, A.: Lung ct image segmentation using deep neural networks. Proc. Comput. Sci. **127**, 109–113 (2018)
38. Sodhro, A.H., Fortino, G., Pirbhulal, S., Lodro, M.M., Shah, M.A.: Energy-efficiency in wireless body sensor networks. Netw. Fut. Arch. Technol. Implement. p. 339 (2017)
39. Sodhro, A.H., Li, Y., Shah, M.A.: Green and friendly media transmission algorithms for wireless body sensor networks. Multimed. Tools Appl. **76**, 20001–20025 (2016)
40. Sodhro, A.H., Luo, Z., Sodhro, G.H., Muzamal, M., Rodrigues, J.J., de Albuquerque, V.H.C.: Artificial intelligence based qos optimization for multimedia communication in iov systems. Fut. Gen. Computer Syst. **95**, 667–680 (2019)
41. Sodhro, A.H., Pirbhulal, S., de Albuquerque, V.H.C.: Artificial intelligence-driven mechanism for edge computing-based industrial applications. IEEE Trans. Ind. Inform. **15**(7), 4235–4243 (2019)
42. Sodhro, A.H., Pirbhulal, S., Qaraqe, M., Lohano, S., Sodhro, G.H., Junejo, N.U.R., Luo, Z.: Power control algorithms for media transmission in remote healthcare systems. IEEE Access **6**, 42384–42393 (2018)
43. Souza, J.W., Alves, S.S., Rebouças, E.d.S., Almeida, J.S., Rebouças Filho, P.P.: A new approach to diagnose parkinson's disease using a structural cooccurrence matrix for a similarity analysis. Computat. Intell. Neurosci. **2018** (2018)
44. Uddin, M.Z., Hassan, M.M., Almogren, A., Alamri, A., Alrubaian, M., Fortino, G.: Facial expression recognition utilizing local direction-based robust features and deep belief network. IEEE Access **5**, 4525–4536 (2017)
45. Vapnik, V.N., Vapnik, V.: Statistical learning theory, vol. 1. Wiley, New York (1998)
46. Wang, C., Frimmel, H., Smedby, Ö.: Level set based vessel segmentation accelerated with periodic monotonic speed function. In: Medical Imaging 2011: Image Processing, vol. 7962, p. 79621M. International Society for Optics and Photonics (2011)
47. Wang, C., Frimmel, H., Smedby, Ö.: Fast level-set based image segmentation using coherent propagation. Med. Phys. **41**(7), (2014)
48. Wang, E.K., Chen, C.M., Hassan, M.M., Almogren, A.: A deep learning based medical image segmentation technique in internet-of-medical-things domain. Fut. Gen. Comput. Syst. **108**, 135–144 (2020)
49. Yang, X., Zhang, T., Xu, C., Yan, S., Hossain, M.S., Ghoneim, A.: Deep relative attributes. IEEE Trans. Multimed. **18**(9), 1832–1842 (2016)
50. Zhang, S., Zhao, Y., Bai, P.: Object localization improved grabcut for lung parenchyma segmentation. Procedia Comput. Sci. **131**, 1311–1317 (2018)