



**HAL**  
open science

## 3D lung nodule segmentation from 2D annotations using morphological operations

Jean-Luc Dupire, Younes Belkouchi, Leo Stepien, Nicolas Billet, Lama Dawi, Felix Wirth, Nathalie Lassau, Hugues Talbot

► **To cite this version:**

Jean-Luc Dupire, Younes Belkouchi, Leo Stepien, Nicolas Billet, Lama Dawi, et al.. 3D lung nodule segmentation from 2D annotations using morphological operations. 2024. hal-04418151

**HAL Id: hal-04418151**

**<https://hal.science/hal-04418151>**

Preprint submitted on 25 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# 3D lung nodule segmentation from 2D annotations using morphological operations

Jean-Luc Dupire<sup>1</sup>, Younes Belkouchi<sup>1,2</sup>, Leo Stepien<sup>1</sup>, Nicolas Billet<sup>3</sup>, Lama Dawi<sup>3</sup>, Felix Wirth<sup>3</sup>, Nathalie Lassau<sup>2,3</sup>, and Hugues Talbot<sup>1</sup>

<sup>1</sup> CentraleSupélec, Université Paris-Saclay, Inria, 91190 Gif-Sur-Yvette

<sup>2</sup> BIOMAPS, Université Paris-Saclay, 94800 Villejuif

<sup>3</sup> Gustave Roussy, 94800 Villejuif

**Abstract.** Tumor volume and heterogeneity are important for patient diagnosis, and automatic lesion segmentation is needed to compute this information from routine CT-Scans. Training a supervised neural network to solve these tasks demands good quality annotations on a large quantity of fully annotated scans, which are difficult and time-consuming to obtain. We propose a fast automatic method using morphological operators to create 3D masks from hand drawn contours of the lesions on their largest axial slice. This type of annotation leads to more precise 3D masks than points or ellipses. Thus, the obtained mask may be used to train end-to-end neural networks for detection and semantic segmentation of lesions on CT-Scans in 3D. We tested this methodology on the LIDC-LUNA dataset to produce the 3D masks from automatically selected 2D annotations. We also produced 3D masks of 115 lung lesions from their 2D contours, and compared them to ground truth 3D masks on an in-house dataset. The results are promising, and the method could be adapted to other organs.

**Keywords:** Nodule Segmentation · Watershed · Medical Imaging.

## 1 Introduction

In oncology, radiologists often use the Response Evaluation Criteria In Solid Tumors RECIST 1.1 [6] to evaluate patient’s response to treatment. By measuring the largest diameter of 5 lesions on a 2D axial slice, the evolution of the sum of diameters of measured lesions is used to classify the response of patients to treatment. However, they provide less information than a volumetric assessment that can be used to extract useful heterogeneity biomarkers [9]. A potential drawback of the RECIST method is that it only considers up to 5 lesions, which reflects the challenge of collecting more data. However, this may not capture the full extent of tumor burden in patients with metastases [2]. Hence, it is important to develop a method for segmenting all lesions in whole body CT-Scans. Semi-supervised methods could be a promising solution, as they could reduce the annotation time by allowing radiologists to mark a lesion and then letting an algorithm complete the 3D segmentation. This could enable a better volumetric evaluation for therapeutic response, a more consistent analysis of tumor

heterogeneity change over time, and a more precise identification of biomarkers to predict treatment outcomes [4, 11]. Therefore, it is essential to develop 3D automatic segmentation models for lung tumors on CT-Scans, since this is the main modality for tumor assessment in oncology. However, training such models is challenging due to the scarcity of fully 3D annotated datasets. To create such datasets, trained radiologists have to annotate all tumors in a CT-Scan, which is time-consuming and laborious. Furthermore, the annotations may not be very precise because of various factors that affect the image quality and interpretation [13]. To facilitate the creation of large databases for training supervised neural networks that can detect and segment all the tumors in a scan, an algorithm for generating 3D masks from 2D annotations is desirable. The input markers for semi-supervised algorithms can be various shapes around the target, such as points, lines, boxes, circles or ellipses. In this study, we aim to explore the benefits of using a hand drawn contour around the target lesion as an input marker, and to obtain a 3D mask using morphological operators. We also aim to adapt and improve a previously developed algorithm [10] for 3D lung nodules segmentation from 2D annotations.

## 2 Related Works

Lung lesion segmentation methods often rely on deep learning, including variants of the U-Net architecture [8]. Jaeger, et al. [7] have developed the *Retina U-Net* architecture for both nodule detection and semantic segmentation. One drawback of deep-learning-based methods is that they need a lot of training data, which is not easy to acquire in practice. This is because creating such data requires a lot of work and time from expert radiologists, who have to annotate the images manually. Another challenge of neural networks is that they are black boxes whose behaviors are hard to interpret, which poses a problem for radiological applications where unexpected errors can have serious implications.

Radiologists may prefer using semi-supervised segmentation methods that rely on gray levels, graphs and morphological operations. These methods require less work and still involve the radiologist in the process. They are also fast enough to be used interactively. For example, Diciotti, et al. deduce 3D masks of small lung nodules from points marking them through thresholding and shape analysis [5]. Tan, et al. proposed an automatic segmentation of lung lesions [10] from a surrounding ellipse on a reference slice, based on marker-controlled watershed, active contours, and Markov random fields. Their method can be summed up as follow.

**Step 1. Markers.** On the reference slice containing the annotation, a threshold is determined via a Gaussian mixture model to distinguish the pixels of the lesion in the ellipse. After morphological operations to refine the estimation of the shape of the lesion, the center  $O$  of the region and the radius  $\hat{R}$  of the largest inscribed circle with center  $O$  are obtained. The object marker is the circle with

center  $O$  and radius  $\hat{R}/2$  on the reference slice. Voxels outside the ellipse or belonging to slices more than  $a + b$  away from the reference slice with  $a$  (resp.  $b$ ) the semi-major (resp. semi-minor) axis of the ellipse are marked as background.

**Step 2. Watershed.** The watershed followed by a morphological opening provides a first segmentation of the lesion. It is applied to a gradient image in which the strength of the edges inside the lesion is reduced. The gradient is multiplied by  $\min(1, (r/\hat{R})^2)$  where  $r$  describes the distance of a voxel to  $O$ .

**Step 3. Active contours.** The segmentation is refined by a geodesic active contour [3, 12].

**Step 4. Markov random fields.** Some lesions may have ground glass opacity (GGO) regions which have not been segmented. These cases are detected according to the mean density of the pixels in the lesion at **Step 1**. A Markov random field allows the segmentation of the image into three regions: the lung parenchyma, the GGO regions, and the high density regions. The densities of the voxels of each region are assumed to be drawn from a Gaussian distribution. The result of this step is combined with the one of **Step 3**. The final segmentation is obtained after morphological operations.

Our method aims to improve and adapt the method presented above to the initial data of a polygon delimiting a lesion on a reference slice (the slice on which the lesion has the largest area).

### 3 Method

Drawing a lesion’s 2D outline precisely provides more information than an ellipse around it. It shows the lesion’s shape better, which helps segment it more accurately in 3D. We reuse the watershed and Markov random fields principles given above with a few modifications. We propose a new way of defining markers for the watershed, which allows us to avoid the active contour step, since it does not improve results. The steps of our automatic lung lesion segmentation are as follows:

**Step 0. Preprocessing.** The image is resized to obtain an isotropic voxel spacing via linear interpolation. The new voxel spacing is chosen as the smallest voxel spacing in the three dimensions of the initial image.

**Step 1. Markers.** The polygonal 2D annotation delineates the lesion better. It delimits a region  $\mathcal{A}$  with centroid  $\mathcal{C}$  which represents the segmentation of the lesion on the reference slice  $S_0$ . Let  $R$  be the radius of the largest inscribed circle in  $\mathcal{A}$ , which gives information on the width of the lesion. The object marker is the erosion of  $\mathcal{A}$  on the slice  $S_0$  by a disk with radius  $R/3$ . The background marker on slice  $S_0$  is the complement of the dilation of  $\mathcal{A}$  on the slice  $S_0$  by a disk with radius  $R/3$ . Slices at a distance greater than  $2\rho$  are deemed to belong to the background where  $\rho$  is 1.5x of the maximum distance between a point of  $\mathcal{A}$  and  $\mathcal{C}$ . On the remaining slices, the voxels outside the dilation of  $\mathcal{A}$  on the slice  $S_0$  by a disk with radius  $R$  are considered to belong to the background. The erosions and dilations of  $\mathcal{A}$  are applied with a structuring element large enough to remedy any inaccuracies in the initial 2D manual annotation.

**Step 2. Watershed.** The marker-controlled watershed algorithm is applied to a modified gradient image. Here, the voxel intensities are smoothed inside an ellipsoid to avoid ragged edges. The gradient image is multiplied by the minimum between 1 and the following value:

$$\left(\frac{2\hat{Z}}{a+b}\right)^2 + \left(\frac{\cos(\theta)\hat{Y} - \sin(\theta)\hat{X}}{b}\right)^2 + \left(\frac{\sin(\theta)\hat{Y} + \cos(\theta)\hat{X}}{a}\right)^2 \quad (1)$$

where  $(\hat{X}, \hat{Y}, \hat{Z})$  are the coordinates of a voxel with origin  $\mathcal{C}$ ,  $a$  (resp.  $b$ ) is the semi-major (resp. semi-minor) axis of the ellipse that has the same second moments as  $\mathcal{A}$ , and  $\theta$  is the orientation of the ellipse. This modification of the gradient image provides prior information on the shape of the lesion to be used. The segmentation given by the watershed is refined by an opening followed by a closing to obtain a smoother contour. The radius of the structuring element of the closing is half that of the opening, which is 3/10 of the radius of the largest inscribed ball in the segmentation.

**Step 3. Markov random fields.** If GGO regions are detected, a new segmentation is computed using a classical Markov random fields (MRF) with pairwise and unary terms. In terms of energy minimization, the problem is to find a labeling  $f$  (lung parenchyma, GGO region, or high intensity region) that minimizes the energy

$$E(f) = \sum_{p \in \mathcal{P}} \sum_{q \in \mathcal{V}(p)} \beta \delta(f_p, f_q) - \sum_{p \in \mathcal{P}} \log \mathbf{P}(c_p | f_p) \quad (2)$$

where  $\mathcal{P}$  is the set of the voxels of the image,  $\mathcal{V}(p)$  is the set of the neighboring voxels of  $p \in \mathcal{P}$ ,  $f_p$  (resp.  $c_p$ ) is the label (resp. the intensity in HU) of  $p$ .  $\beta$  is set to 100 and  $\mathbf{P}(c_p | f_p)$  is given by the Gaussian distribution proposed by Tan, et al. [10].  $\delta(f_p, f_q)$  equals -1 if  $f_p = f_q$  and 1 otherwise. The new segmentation is the union of the previous segmentation with the result of solving the MRF

problem. However, if the volume of the MRF segmentation is too large (in our case if it is greater than  $3\times$  the volume of a ball with radius  $\rho$ ), the result is considered inconsistent and is not taken into account. Similarly, the MRF result is kept alone if the volume of the watershed segmentation is greater (by a factor of 1.2 or more), since the watershed is supposed to be more conservative. Finally, the contours are made smoother by morphological opening with a structuring element of minimum radius between 2mm and  $2\rho/5$ .

The active contours refinement was abandoned because it did not improve results after the watershed step.

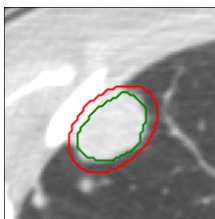
## 4 Experiments

### 4.1 Dataset

We evaluated the algorithms on 115 lung and pleural lesions from 6 contrast-injected Thoracic CT-Scans from Gustave Roussy hospital to compare the method of Tan, et al. [10] (denoted SEG0) to our method (denoted SEG). Scans 1 and 2 each have a single solid lung lesion. Scan 5 has 98 lesions in total, with a mix of solid, GGO, necrotic and subpleural pulmonary nodules.

An expert radiologist first drew a polygon around each lesion on the axial slice with the largest diameter. These were the 2D manual annotations. For SEG0, we converted them into ellipses that enclosed the lesions, as Fig. 1 shows. These annotations were the algorithms' starting point.

These annotations were also the starting point of manually drawn 3D reference annotations, obtained by completing the annotations of each lesion on all its axial slices. These were only used to compare the performance of both algorithms and never as input.



**Fig. 1.** Initial 2D annotations: a polygon for SEG (green) and an ellipse (red) for SEG0

In addition we evaluated the algorithms on 868 low-dose lung CT-Scans from the Lung Image Database Consortium image collection (LIDC-IDRI) [1] for a total of 6804 lesions. LIDC-IDRI contains both screening and diagnostic CT-Scans collected from 7 academic centers and 8 medical imaging companies. The 3D annotations performed by multiple radiologists are available. Our ground truth is

the mask consisting of the consensus of the radiologists if multiple annotations were present (a pixel is added to the mask if it belongs to more than half of the segmentations made by the radiologists). Using this ground truth, we extract a 2D hand contour and a 2D ellipse at the slice where the tumor was largest.

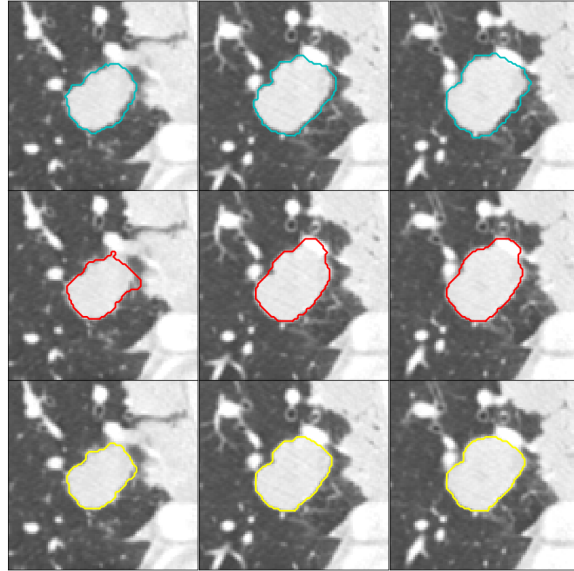
## 4.2 Results

**In-house dataset** The comparison was evaluated using Intersection over Union (IoU), Recall, Precision, and Dice scores to get an overall view of the behavior of the two methods. These scores are evaluated with respect to the 3D reference annotation from Gustave Roussy. The results of the two methods on all the scans are given in Table 1.

**Table 1.** Performance comparison on Gustave Roussy CT-Scans

Scan	Method	Metrics			
		IoU	Precision	Recall	Dice
1(1)	SEG0	0.5921	0.6484	0.8721	0.7438
	SEG	<b>0.6035</b>	<b>0.6533</b>	<b>0.8877</b>	<b>0.7527</b>
2(1)	SEG0	0.7106	0.7275	<b>0.9682</b>	0.8308
	SEG	<b>0.7650</b>	<b>0.7878</b>	0.9634	<b>0.8668</b>
3(5)	SEG0	0.6949	0.9627	0.7142	0.8200
	SEG	<b>0.7302</b>	<b>0.9698</b>	<b>0.7472</b>	<b>0.8441</b>
4(4)	SEG0	0.5774	0.7859	0.6852	0.7321
	SEG	<b>0.7347</b>	<b>0.7870</b>	<b>0.9171</b>	<b>0.8471</b>
5(98)	SEG0	0.6676	0.8653	0.7450	0.8007
	SEG	<b>0.7541</b>	<b>0.9163</b>	<b>0.8099</b>	<b>0.8598</b>
6(6)	SEG0	0.5921	0.6484	<b>0.8721</b>	0.7438
	SEG	<b>0.7519</b>	<b>0.8767</b>	0.8408	<b>0.8584</b>

Our method seems to perform better than SEG0 on all criteria and seem to better take into account thin details of the lesions (Fig. 2). The use of a precisely drawn polygon as input allows for instance to remedy the problem of strong edges for watershed in the lesions in case of cavities. Nodules with elongated or flattened shapes, a case notably encountered among subpleural lesions, are better handled as the algorithm has more input information on the global shape. Solid lesions surrounded by lung parenchyma are segmented nearly perfectly by both methods. However, subpleural lesions for which the borders are not well defined were the most difficult to correctly annotate by either methods.



**Fig. 2.** Comparison of SEG and SEG0. *Cyan: Reference; Red: SEG0; Yellow: SEG.* Each column displays mask contours on an axial slice of the same lesion. The right-hand column correspond to the slice of the initial 2D annotation.

SEG differentiates the lesion from vessels, which is not the case here with SEG0 that includes a vessel on Figure. 2.

**LIDC-IDRI** The results of the two methods evaluated on the LIDC dataset are displayed in Table 2.

SEG surpasses SEG0 on average for each metric, coupled with more consistent scores, as the standard deviation is lower for our method. The lesions on the LIDC dataset are usually small (a few mm wide), with a few on each scanner, meaning that these algorithms might easily overflow, especially when close to the pleura. This also implies that small misses might be more impactful on the computed metrics. Overall, having the exact form of the lesion seems to yield a more appropriate marker for morphological operations.



**Table 2.** Performance comparison on LIDC-IDRI CT-Scans

Metrics	SEG0		SEG	
	Mean	SD	Mean	SD
IOU	0.3996	0.2250	<b>0.6008</b>	0.1600
Precision	0.6720	0.3227	<b>0.7397</b>	0.2015
Recall	0.5455	0.2652	<b>0.7816</b>	0.1165
DICE	0.5317	0.2482	<b>0.7366</b>	0.1417

## 5 Conclusion

Our methodology was able to successfully segment in 3D lung lesions on scanners from Gustave Roussy hospital (with multiple lesions in a single scan), and from a large scale database (LIDC-IDRI). Being able to use precisely drawn 2D annotations seems to help segment lung lesions in 3D more accurately than an ellipse. Furthermore, our refinement of the markers seems reliable, and allowed us to discard the active contour computation from the segmentation pipeline. Our algorithm works well for a variety of lung lesions: solid, part-solid, necrotic, sub-pleural or pleural. It also has proven to be reliable on a large scale dataset such as LIDC-IDRI. However, the methodology requires a precisely hand drawn contour on the largest axial slice of each lesion, which remains challenging, especially in routine clinical practice. Moreover, we did not test the capacity of this method to generalize to other tissues in a CT-Scan. Lung nodule segmentation remains a relatively easy task due to the high contrast between nodules and pulmonary tissues. Lastly, evaluating if deep neural networks trained on datasets generated with both algorithms yield a significant difference in automatic detection remains to be explored. Nevertheless, these methodologies could serve in creating large scale databases faster, to constitute a bigger learning base for automatic lung segmentation algorithms.

## References

1. Armato, S.G. et al.: 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 (2011). <https://doi.org/10.1118/1.3528204>.
2. Belkouchi, Y. et al.: Predicting immunotherapy outcomes in patients with MSI tumors using NLR and CT global tumor volume. *Front. Oncol.* 12, 982790 (2022). <https://doi.org/10.3389/fonc.2022.982790>.
3. Caselles, V. et al.: Geodesic active contours. In: *Proceedings of IEEE International Conference on Computer Vision*. pp. 694–699 IEEE Comput. Soc. Press, Cambridge, MA, USA (1995). <https://doi.org/10.1109/ICCV.1995.466871>.

4. Dercle, L. et al.: Early Readout on Overall Survival of Patients With Melanoma Treated With Immunotherapy Using a Novel Imaging Analysis. *JAMA Oncol.* 8, 3, 385 (2022). <https://doi.org/10.1001/jamaoncol.2021.6818>.
5. Diciotti, S. et al.: Automated Segmentation Refinement of Small Lung Nodules in CT Scans by Local Shape Analysis. *IEEE Trans. Biomed. Eng.* 58, 12, 3418–3428 (2011). <https://doi.org/10.1109/TBME.2011.2167621>.
6. Eisenhauer, E.A. et al.: New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *European Journal of Cancer.* 45, 2, 228–247 (2009). <https://doi.org/10.1016/j.ejca.2008.10.026>.
7. Jaeger, P.F. et al.: Retina U-Net: Embarrassingly Simple Exploitation of Segmentation Supervision for Medical Object Detection. (2018). <https://doi.org/10.48550/ARXIV.1811.08661>.
8. Ronneberger, O. et al.: U-Net: Convolutional Networks for Biomedical Image Segmentation. (2015). <https://doi.org/10.48550/ARXIV.1505.04597>.
9. Sun, R. et al.: A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *The Lancet Oncology.* 19, 9, 1180–1191 (2018). [https://doi.org/10.1016/S1470-2045\(18\)30413-3](https://doi.org/10.1016/S1470-2045(18)30413-3).
10. Tan, Y. et al.: Segmentation of lung lesions on CT scans using watershed, active contours, and Markov random field. *Medical Physics.* 40, 4, 043502 (2013). <https://doi.org/10.1118/1.4793409>.
11. Tselikas, L. et al.: Role of image-guided biopsy and radiomics in the age of precision medicine. *Chin Clin Oncol.* 8, 6, 57–65 (2019). <https://doi.org/10.21037/cco.2019.12.02>.
12. Yezzi, A. et al.: A geometric snake model for segmentation of medical imagery. *IEEE Trans. Med. Imaging.* 16, 2, 199–209 (1997). <https://doi.org/10.1109/42.563665>.
13. QIBA CT Volumetry Biomarker Committee. CT Tumor Volume Change for Advanced Disease. Quantitative Imaging Biomarkers Alliance. Profile Stage: Clinically Feasible. July 21, 2022. Available from: <http://qibawiki.rsna.org/index.php/Profiles>