ISSN 2278-3091
Volume 8, No.4, July – August 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse24842019.pdf
https://doi.org/10.30534/ijatcse/2019/24842019

3D Lung Segmentation on CT Images Using Region-Based Method

S. Akila Agnes[1], J. Anitha[2*]
[1]Dept. of CSE, Karunya Institute of Technology and Sciences, Coimbatore, India, akilaagnes@karunya.edu
[2*]Dept. of CSE, Karunya Institute of Technology and Sciences, Coimbatore, India, anitha_j@karunya.edu

ABSTRACT

Computed Tomography (CT) images are becoming a valuable aid in advanced lung cancer investigation. Automated segmentation of lung structures from CT images is the elemental step in many lung cancer diagnosis systems, where the malignant nodules are identified at its initial stage. Lung segmentation is a challenging task because of the variant density distribution amongst the lung region, similar densities in the pulmonary structures, and scanning protocols. This work intended to recognize the efficient region-based segmentation framework to segment lung volumes. The standard region-based approaches, such as watershed and region growing, are adapted to segment the lung region. The specified segmentation methods segment lung in the axial space and achieve promising segmentation results. The performance of the system is assessed by computing the DSC, Hausdorff distance and AVD between the automatically segmented lung volumes and the volumes manually outlined by radiological experts. The marker-controlled watershed method followed by a sequence of morphological operations segments the lung parenchyma accurately with a dice score of 92.5 and the region growing method with a dice score of 93.8 concerning the expert-traced contours. The obtained results confirm the effectiveness of the system. It is suggested that the performance of the described framework could be further enhanced by fusing shape-based features.

Key words : lung cancer diagnosis, region growing, volumetric lung segmentation, watershed algorithm.

1. INTRODUCTION

Lung cancer endures the preeminent cause of cancer death among men and women in the world [1]. The National Lung Screening Trial (NLST) has manifested that the low dose helical CT based cancer-screening process demotes the mortality rate by detecting lung cancer at the preclinical stage. Volume-based image investigation possesses a significant role in intensifying the detection and diagnosis of lung cancer on CT images. Thoracic computed tomography imaging system permits the physician to observe the inner organs, bones, and tissues of the individual. Chest CT scan image has four organs-at-risk such as the esophagus, heart, left and right lungs and spinal cord. The right and left lungs can be bounded separately, but they are considered as a single structure for lung dosimetry. CT screening is an important procedure routinely conducted to identify cancer in the anatomical structures (organs-at-risk) at the initial stage [2]. Hand-operated investigation demands a notable amount of time to segment the ROIs, and it is essential for the ROIs (lung parenchyma) to be segmented precisely to discover the changes in the anatomical structure. Automatic analysis of lung pathologies and 3D lung parenchyma segmentation are the current interest in the field of medical image analysis.

Lung segmentation is defined as the delineation of the region enveloped by the pleura, which comprises lung parenchyma commonly with pulmonary bronchi, vessels, and nerves. Lung segmentation is a pre-processing step in computer-aided diagnosis systems such as lung cancer and pulmonary diseases. This paper intended to examine the performance of two traditional region-based methods in segmenting the lung volume from the pile of CT slices in axial orientation. In literature, various methods are reported to separate the lung parenchyma from the CT scan images. Conventional segmentation methods are categorized into four categories: 1) Intensity-based segmentation methods that segment the image based on the intensity value of a pixel or a voxel. Thresholding [3], [4] technique departs the image based on pixel intensity. This method uses a threshold value to convert a grayscale input image into a binary output that designates the segmented regions. The technique compares each image pixels with the defined threshold value, and if the pixel value is higher than the threshold, the binary output pixel is set to the one; unless, it is set to zero. The threshold is determined by previous domain knowledge or histogram based algorithms. Hence, the CT images are obtained from distinct scanner machines, identifying the optimal threshold value is still a challenging assignment. Global thresholding fails to operate when the foreground and background pixels have coinciding grayscales but are separated in space. The CT image could not be segmented properly with global thresholding.

2) Region-based segmentation methods set apart the identical region attentively by developing the region of interest from a seed point or a preliminary target volume. The region-growing method starts with a pixel or voxel as a prime region and explores the neighborhood pixels or voxels that meet a similarity condition with the region of interest. This process develops the regions continuously until no other similar pixel or voxel is observed. Hao et al. [5] have proposed

a region-growing strategy to segment ultrasound images. Though this method is simple to implement, it takes more time because it requires examining every possible pixel or voxel. 3) The edge-based segmentation method segments the images based on the predominant edges present in the image. This method is not reliable when the surfaces of the objects are, and the edges are not connected. In CT image, the lung area includes air valves and veins, and this distracts the algorithm by producing wrong edges. 4) Traditional clustering algorithms such as K-means and FCM are commonly used to segment the images based on the spatial relationship between the pixels. These techniques are fit to 2D images, and for 3D images, it consumes longer time and memory. These methods are limited by computer power and used no or little prior-knowledge.

Advanced techniques have developed in an attempt to overcome the difficulties in heuristic approaches [6-8]. Numerous ambiguity models and optimization methods have proposed to segment the lung from chest CT scan. Cootes et al. [9] have proposed a shape-training models based on the collection of shapes, where a flexible point distribution model is produced during the object search into the image. A fully automated lung segmentation algorithm based on adaptive border marching has been proposed to segment the lungs, including juxta-pleural nodules that are adjacent to the chest wall [10]. This method minimize the over segmentation of adjacent regions. A robust active shape model approach [11] has been proposed to segment the lungs with pathological conditions, and an optimal surface finding approach is adapted to improve the segmented results. Besides, not all these approaches are standard for distinct applications.

In this work, two common region-based segmentation methods have been adapted for automatic 3D-lung segmentation. The first method is based on a region-growing algorithm [12] and the second method utilizes a marker controlled watershed approach [13]. The performances of these approaches are analyzed in terms of dice coefficient (DSC), hausdorff distance and average hausdorff diatance (AVD) and recommendations are given based on the obtained results. The structure of this paper is as follows: This section studies the related research carried on lung segmentation. Section 2 provides the detailed methods for 3-D lung segmentation. The experimental results are discussed in section 3. Finally, the conclusion and informations for future work are provided in Section 4.

2. METHODOLOGY

2.1 Dataset

The Cancer Imaging Archive (TCIA) is a service that affords a free archive of cancer images for public access [14]. In TCIA, the medical image data (both CT and RTSTRUCT) are stored in Digital Imaging and Communications in Medicine (DICOM) file format. RTSTRUCT file carries the hand-operated contours that have been used in the clinic for treatment planning, and this is used as ground-truth for validation. Figure 1 presents the axial and sagittal orientation of the chest CT scan image. From Figure 1, it is apparent that the bone structures (white region) have a higher density than body muscles (gray region) and muscles have a higher density than lung parenchyma (dark region).

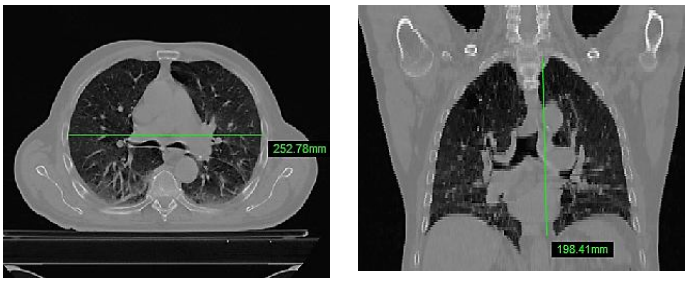


Figure 1: CT sample images in axial and sagittal plane views

Hounsfield Unit (HU) is a conventional system for measuring the density of a voxel in CT scan. In a CT scan, HU is equivalent to the degree of x-ray attenuation, and it is allocated to each voxel to show the image that represents the density of the anatomical tissue [15]. The HU value for the anatomical structure of chest CT scan is presented in Table 1.

Table 1: Hounsfield unit value for the anatomical structure of chest CT scan

Anatomical structure or substance	Hounsfield Unit (HU) value
Air	-1000
Lung	-700 to -900
Water	0
Muscle	50
Bone	>1000

2.1 2.2 Marker controlled watershed algorithm

A marker-based watershed segmentation method is developed to segment lung parenchyma from chest CT images. The original watershed segmentation has a great response to thin edges; this may stop the algorithm by returning over-segmented results. Hence, lung segmentation with regular watershed transform to the gradient CT image can be settled in over segmentation due to other dominant edges. The idea of markers is included to control the over segmentation problem of the traditional watershed method. A marker is a coupled element associated with an image that modifies the gradient image. Two types of markers, such as internal and external, are used in marker-controlled watershed segmentation. The internal marker expresses the region of interest, and an external marker expresses the boundary region. The marker-controlled watershed segmentation is an effective method for segmenting the regions with shuttered contours, where the edges are meant as ridges. After the initial segmentation, the boundaries of the watershed regions

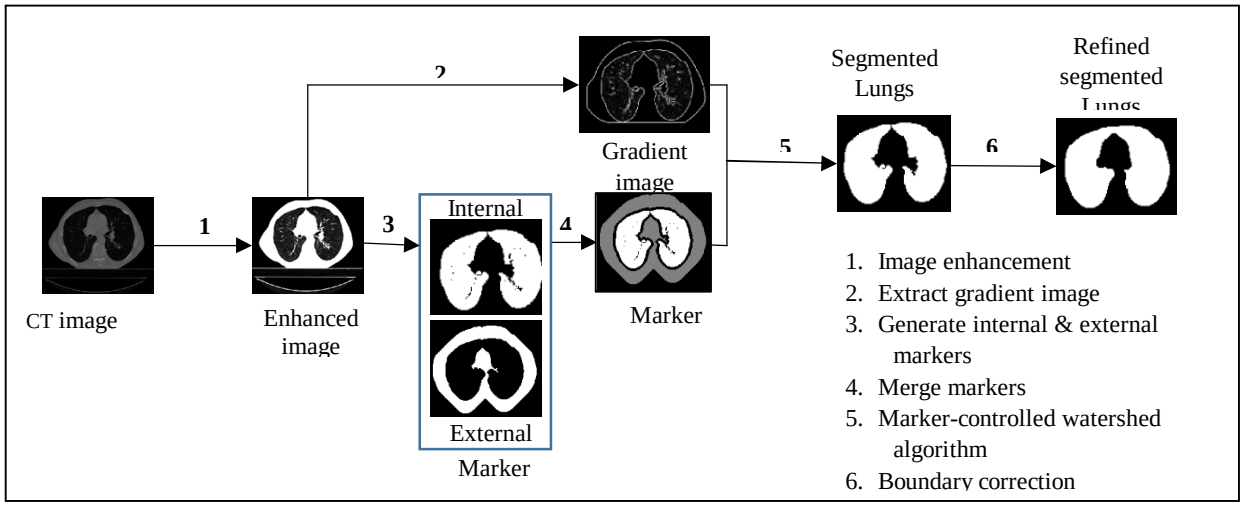


Figure 2: Segmentation of lung volumes using marker-controlled watershed

are smoothened by adopting proper morphological operation. Figure 2 shows the framework for segmenting lung volumes using marker-controlled watershed approach.

A. Algorithm

1. The input CT image is intensified using the windowing operation (1), where the lung region voxels are

$$v_{gray} = Window\left(v_{HF}, w_{\min}, w_{\max}\right) = \begin{cases} 0 & \forall\ v_{HF} < w_{\min} \\ 255 \times \dfrac{v_{HF} - w_{\min}}{w_{\max} - w_{\min}} & \forall\ w_{\min} \le v_{HF} \le w_{\max} \\ 255 & \forall\ w_{\max} < v_{HF} \end{cases} \quad (1)$$

where $w_{\min}$ is minimum range in HU and $w_{\max}$ is maximum range in HU.

2. For each axial slice in the input 3D CT scan image, do steps 3 - 7.

3. Obtain the gradient image G from the CT image using the Sobel masking operator (2). Sobel convolution kernels applied separately to the image in horizontal and vertical direction and computed the gradient images G_x and G_y respectively.

$$G = \sqrt{G_x^{\,2} + G_y^{\,2}} \quad (2)$$

4. Generate internal and external markers from the CT image. Apply the global thresholding to generate the internal marker, which represents the region of interest. Two larger regions are selected, and the corresponding voxels are set as internal markers. External marker is a wide strip around the internal marker with the thickness of 10mm. The binary dilation is applied to the internal marker to obtain the external marker.

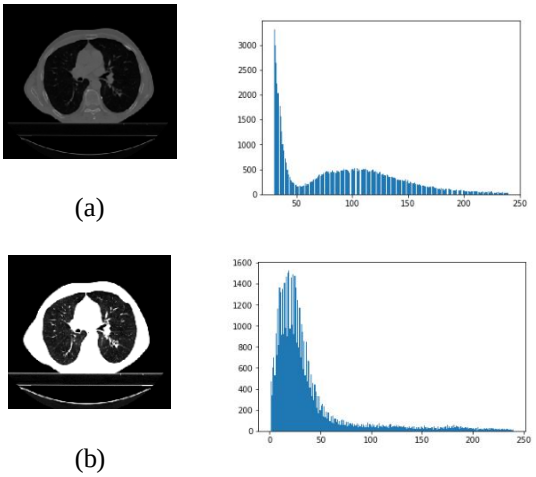


Figure 3: Histogram plot of the chest CT image before and after windowing operation; (a) Input CT image & Histogram; (b) Enhanced Image & Histogram

5. Watershed marker is obtained by merging the internal and external markers.

6. Watershed transform algorithm takes the gradient image and watershed marker as input, and then segments the image by retaining the significant edges between the markers.

7. Top-hat transform is applied to the obtained segmented image that detects and includes the small details on the wall of the lung region.

2.3 Region growing algorithm

Region-growing algorithm consists of four main stages: automatic seed selection, algorithm-stopping condition, region growing, and process to correct the segmented lungs. This algorithm is applied to all the slices with the thickness of minimum 2.5 mm in axial direction.

A. Algorithm

1. The input CT image is intensified using the windowing operation (1), where the lung region voxels are highlighted. Intensity windowing function converts the v_{HF} (voxel in HU) in the range of [-1000, 3054] into v_{gray} (voxel in grayscale) in the range of [0, 255].

2. Initialize the volume mask $V \quad (i = 0, \ldots, n;\ j = 0, \ldots, m)$ and set the voxel's values as zero. $V_{ij} = 0$

3. For each axial slice in the input 3D CT scan image, do steps 4 - 6.

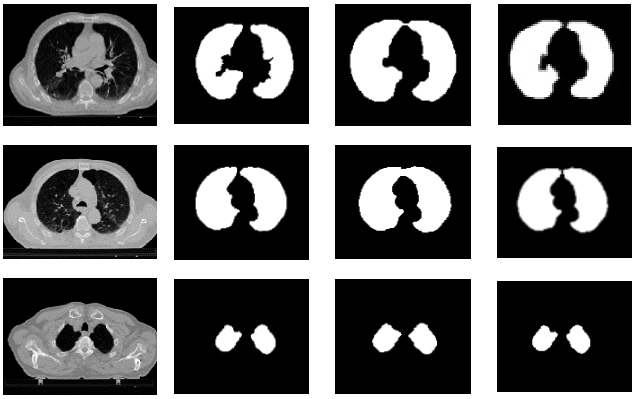
4. Set the initial seed by histogram peak finding method [16] for the current slice in the volume mask and enqueue all seeds into queue.

5. For each voxel in the queue, check its neighbor and enqueue the voxel into queue, if the neighboring voxel is similar to the current voxel. Repeat this step until the queue is empty.

6. Rolling ball filter is applied to fill the gaps on the boundary and 20mm diameter of ball is used that includes juxta- pleural nodules.

3. RESULTS AND DISCUSSION

In this experiment, Lung CT Segmentation Challenge 2017 (LCTSC) database of 60 thoracic CT studies has been employed. Evaluation has performed against the manual contours available in DICOM RTSTRUCT file that is provided in the LCTSC dataset. These manual contours are used as ground truth label to assess the performance of the proposed approach. The lung data considered for this test is excluded from tumor, main bronchi, airways and vessels greater than $5 \pm 2\ mm$ diameter.



(a) Input image (b) Ground truth (c) Region growing (d) Marker-controlled watershed

Figure 4: Axial position CT images show the segmented results of region growing and marker-controlled watershed algorithm. Each row shows the segmented results for different slices of a case (LCTSC-Train-S1-001)

The visualization of the segmented results shown in Figure. 4 and Figure. 5 confirms that the above-mentioned algorithms segments 3D lung parenchyma from the CT scan images. The effectiveness of these approaches is analyzed further in quantitative measures with dice coefficient, hausdorff distance and AVD.

(a) (b) (c)

Figure 5: Volumetric lung segmentation. (a) Ground truth (left column, green color); (b) lungs delineated by region growing (middle column, red color); (c) lungs delineated by watershed method (last column, blue color)

3.1 Evaluation metrics

Dice Coefficient

Dice coefficient (DSC) is a measure of relative overlap between the segmented result and the ground truth result, where 1 represents perfect match and 0 represents no overlap.

$$D = \frac{2|X| \cup |Y|}{|X| + |Y|} \tag{3}$$

where X is ground truth and Y is the segmented result.

Table 2: Results on lung segmentation using marker controlled watershed

Study ID (Train)	Marker Controlled Watershed		
	DSC	Hausdorff	AVD
LCTSC-Train-S1-001	0.93	21.26	15.54
LCTSC-Train-S1-002	0.94	13.34	9.64
LCTSC-Train-S1-003	0.96	14.89	6.64
LCTSC-Train-S1-004	0.92	19.12	10.07
LCTSC-Train-S1-005	0.92	13.87	8.56
LCTSC-Train-S1-006	0.93	21.45	9.67
LCTSC-Train-S1-007	0.89	29.01	12.45
LCTSC-Train-S1-008	0.93	14.98	8.57
LCTSC-Train-S1-009	0.95	13.78	8.01
LCTSC-Train-S1-010	0.88	24.36	11.52

Hausdorff Distance

The Hausdorff Distance (HD) between two finite vectors X and Y is defined by (4)

$$HD(X,Y) = \max(h(X,Y), h(Y,X)) \quad (4)$$

where $h(X,Y)$ is called the directed Hausdorff distance is computed by (5)

$$h(X,Y) = \max\ x \in X\ \min\ y \in Y \|x - y\| \quad (5)$$

The HD is sensitive to noise and outliers. Since, noise and outliers are usual in medical image segmentation; it is advised that do not assess the performance of algorithms with HD alone.

Table 3: Results on lung segmentation using region growing

Study ID (Train)	Region Growing		
	DSC	Hausdorff	AVD
LCTSC-Train-S1-001	0.94	15.17	6.59
LCTSC-Train-S1-002	0.93	14.97	8.63
LCTSC-Train-S1-003	0.97	12.65	7.03
LCTSC-Train-S1-004	0.95	16.51	5.98
LCTSC-Train-S1-005	0.94	15.23	6.90
LCTSC-Train-S1-006	0.95	17.75	5.45
LCTSC-Train-S1-007	0.90	23.15	9.65
LCTSC-Train-S1-008	0.92	15.76	5.78
LCTSC-Train-S1-009	0.97	10.82	6.54
LCTSC-Train-S1-010	0.91	18.48	9.23

Average Hausdorff Distance

The Average Hausdorff Distance (AVD), is the averaged Hausdorff Distance across all points. The AVD is less sensitive to noise and outliers than the HD. It is defined by (6)

$$AVD(X,Y) = \max\{d(X,Y), d(Y,X)\} \quad (6)$$

where $d(X,Y)$ is the directed Average Hausdorff distance that is computed by (7)

$$d(X,Y) = \frac{1}{|X|}\sum_{x \in X} \min\ y \in Y \|x - y\| \quad (7)$$

The Hausdorff distance and the average Hausdorff distance are computed based on the distances among all pairs of voxels. As all possible pairs are considered, these metrics are reliable and robust, especially in volumetric images. The experiment results of 3D Lung segmentation using watershed and region growing algorithm is validated with 10 clinical CT images from LCTSC dataset. The performance of these two algorithms in terms of DSC, Hausdorff distance and AVD are presented in Table 2 and Table 3 for comparative analysis.

Table 4: Performance comparison of lung segmentation methods

Metrics	Watershed	Region Growing
DSC	0.92 ± 0.02	0.94 ± 0.02
Hausdorff	18.06 ± 5.34	16.04 ± 3.35
AVD	10.06 ± 2.55	7.17 ± 1.47

The statistical results of the algorithms are provided in Table 4. Compared to marker-controlled watershed method, region-growing method segments the lung volumes accurately. However, the result of region growing method is fully depends on the initialization of the seed points. Though it adapts automatic initialization of seed using the histogram peak finding method [13] , it requires the expertise input for the CT images with pathological conditions. As well, this could segment the lung, which has homogeneous region. From the experimental results, it is noted that the watershed method followed by morphological operations segment the lung volumes efficiently without demanding initializations.

4. CONCLUSION

The segmentation framework implemented with watershed and region-growing algorithms segment the lung parenchyma from the chest CT images precisely. The experiment result confirms that two traditional approaches such as watershed and region-growing automatically segment the lung volume from CT image with an average dice coefficient of 0.92 and 0.94 respectively. In order to include the juxta-pleural nodules, morphological operations are performed after segmenting the regions from the 3D volume CT scan images. Region-growing algorithm achieves the average Haousdorff distance of 7.17 and watershed algorithm is 10.06. Watershed based segmentation is robust against the lung with high pathological conditions, but the region-growing algorithm has limitations when the lung region has abnormalities. Though these methods have

achieved promising results; still, the overall system needs to be renewed. 3D and 2D visualization of obtained results reveal that the inclusion of other regions parts, such as the trachea and spinal cord, which has the same intensity. In future, a segmentation algorithm needs to be designed to segment all these organs too.

REFERENCES

1. American Lung Association. **Lung cancer fact sheet**. *American Lung Association* http://www. lung. org/lung-health-and-diseases/lung-disease-lookup/lung-cancer/resource-library/lung-cancer-fact-sheet. html (accessed 11 September 2018) (2015).
2. M. Puderbach and H.-U. Kauczor, **Can lung MR replace lung CT?**, *Pediatr. Radiol.*, vol. 38, no. 3, pp. 439–451, 2008. https://doi.org/10.1007/s00247-008-0844-7
3. A. Karthikeyan and M. Valliammai, **Lungs segmentation using multi-level thresholding in CT images**, *IJECSE*, 2012.
4. L.-Y. Tseng and L.-C. Huang, **An adaptive thresholding method for automatic lung segmentation in CT images**, in *AFRICON 2009*, 2009, pp. 1–5. https://doi.org/10.1109/AFRCON.2009.5308100
5. X. Hao, C. Bruce, C. Pislaru, and J. F. Greenleaf, **A novel region growing method for segmenting ultrasound images**, in 2000 *IEEE Ultrasonics Symposium. Proceedings. An International Symposium* (Cat. No. 00CH37121), 2000, vol. 2, pp. 1717–1720.
6. R. Biswas and S. Roy, **Content based CT image sign retrieval using Fast Discrete Curvelet Transform and deep learning,** *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 854–863, 2019. https://doi.org/10.30534/ijatcse/2019/80832019
7. R. Jothikumar, G. Siva Shanmugam, M. Nagarajan, S. Premkumar, and A. Asokan, **Analyzes of mouth cancer using max-min composition in soft computing**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 825–830, 2019. https://doi.org/10.30534/ijatcse/2019/76832019
8. S. Rajurkar, U. Kodwani, A. Singh, and S. Mundada, **Online prescription for skin diseases**, *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 2, pp. 176–181, 2019. https://doi.org/10.30534/ijatcse/2019/11822019
9. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, **Active shape models-their training and application**, *Comput. Vis. image Underst.*, vol. 61, no. 1, pp. 38–59, 1995. https://doi.org/10.1006/cviu.1995.1004
10. J. Pu, J. Roos, A. Y. Chin, S. Napel, G. D. Rubin, and D. S. Paik, **Adaptive border marching algorithm: automatic lung segmentation on chest CT images**, *Comput. Med. Imaging Graph.*, vol. 32, no. 6, pp. 452–462, 2008. https://doi.org/10.1016/j.compmedimag.2008.04.005
11. S. Sun, C. Bauer, and R. Beichel, **Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach**, *IEEE Trans. Med. Imaging*, vol. 31, no. 2, pp. 449–460, 2011. https://doi.org/10.1109/TMI.2011.2171357
12. R. Adams and L. Bischof, **Seeded region growing**, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 6, pp. 641–647, 1994. https://doi.org/10.1109/34.295913
13. X. Yang, H. Li, and X. Zhou, **Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy**, *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 53, no. 11, pp. 2405–2414, 2006. https://doi.org/10.1109/TCSI.2006.884469
14. K. Clark et al., **The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository**, *J. Digit. Imaging*, vol. 26, no. 6, pp. 1045–1057, 2013. https://doi.org/10.1007/s10278-013-9622-7
15. P. W. Goaz and S. C. White, **Oral radiology: principles and interpretation**. *Mosby Louis*, 1994.
16. M. Xess and S. A. Agnes, **Analysis of image segmentation methods based on performance evaluation parameters**, *Int. J. Comput. Eng. Res*, vol. 4, no. 3, pp. 68–75, 2014.