REVIEW ARTICLE

Evaluation of thresholding and region growing techniques in segmenting lung nodules in computed tomography images: a systematic review

Siti Nur Atiqah Mohamad Sabri, Noor Shafini Mohamad*

[1]*Centre of Medical Imaging, Faculty of Health Sciences, Universiti Teknologi MARA Cawangan Selangor Kampus Puncak Alam, 42300 Bandar Puncak Alam, Selangor, Malaysia*

***Corresponding Author**

Noor Shafini Mohamad
Email: shafini.mohamad@uitm.edu.my

Abstract:

Image segmentation is an essential step in computer-aided diagnosis and treatment planning of lung nodules. Therefore, the purpose of this study was to perform a systematic review and provide an overview of the literature available on image segmentation algorithm, which is thresholding and region growing method regarding the optimization (of the different methodologies developed) of lung nodules in the lung CT scan prior for the lung nodule segmentation. This systematic review was compiled according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. A total of 687 articles were retrieved from the databases, and six articles were selected for this review. The finding showed that a 3D Automatic Lung Parenchyma Extraction and Border Repair (ALPE&BR), which consists of an Automatic Single Seeded Region Growing (ASSRG) and a 3D Spherical region-growing method (SPRG), showed the highest sensitivity of 98.5% and 83.245%, respectively. Improvement of the existing methods or proposing a new one may be the best option. Standardization of the evaluation metrics is needed to allow a direct comparison between methods

Keywords: Computed Tomography, lung nodule, region growing, segmentation, thresholding

1. INTRODUCTION

Lung cancer is one of the most chronic forms of cancer in the world [15]. It accounts for 19.8% of all cancer-related mortality and with the lowest cancer survival (11.0%) in Malaysia [4]. According to the National Cancer Registry (NCR), lung cancer is the third most common cancer in the country, the second-most in males, and the 4th most in females. It has been concluded that cigarette smoking is a major etiological risk factor [4]. Lung cancer is usually being detected at an advanced stage when surgery is no longer an option [7]. According to Loh et al (2007) the reasons for refusable to be operated were patients' beliefs in traditional, complementary treatment rather than medical treatment.

Evaluation of lung nodules in the past is usually by CT scan with thin-slice thickness, contrast media, and image post-processing such as using multiple imaging techniques, including shaded surface display (SSD), volume-rendering (VR), and multiplanar reconstruction (MPR) [5]. Contrast media helps the radiologist to have extra information about the vasculature of the suspected pulmonary nodular disease. Nevertheless, a thinner slice produces a greater detail, but the system needs to increase the mAs to reduce images noise-producing more exposure to the patient. Contrast media is highly potential for an adverse reaction, especially for a patient who in advanced age, elevated serum creatinine level, dehydration, diabetes, and renal disease. However, the detection of faint abnormal regions, such as Ground-glass opacity nodules (GGNs), can be challenging and time-consuming since GGNs are usually small and show contrast with the surrounding lung parenchyma [7].

However, these days, researchers more likely to solve segmentation problems that involve machine learning rather than the simplest yet effective segmentation algorithm [6]. Therefore, the purpose of various thresholding and region growing algorithms is to provide an overview of the literature available on image segmentation algorithms, which are thresholding and region growing methods for optimization of lung nodules in the lung CT scan. A key to better diagnosis of the pathological condition in the image can be achieved through a correct segmentation process. Therefore, the rationale of doing the literature search is to help to find a suitable segmentation method that will help the radiologist to improve for a better delineation of lung nodules effectively.

2. MATERIALS AND METHODS

The eligibility criteria and analysis in this review were performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines 2009. Research question was formulated based on the patient, index test, comparator, outcome, and study design (PICOS) criteria as the following: What is the performance of the thresholding and region growing methods in segmenting the lung nodule in CT images, as compared with the radiologist's manual segmentation?

2.1 Literature Search

The review methods of the present studies were conducted using two primary databases which are PubMed and IEEE Xplore. A literature search was also done through two supporting databases, such as Taylor & Francis Online and Dimension. Search strings were developed in April 2020, using the following expressions:

Search string number 1: ("Threshold" OR "Thresholding" OR "Otsu" OR "Multi-Thresholding" OR "Adaptive Thresholding") AND ("Segmentation" OR "Delineation") AND ("Lung" OR "Lungs" OR "Pulmonary" OR "Chest") AND ("Nodule" OR "Nodules" OR "Cancer" OR "Tumor" OR "Tumors" NOT "Emphysema") AND ("Computed Tomography" OR "Computed Tomographic" OR "CT" OR "CT SCAN" NOT "PET/CT" NOT "Positron Emission Tomography" NOT "Radiotherapy").

Search string number 2: ("Region Growing" OR "Seeded Region Growing" OR "SRG") AND ("Segmentation" OR "Delineation") AND ("Lung" OR "Lungs" OR "Pulmonary" OR "Chest") AND ("Nodule" OR "Nodules" OR "Cancer" OR "Tumor" OR "Tumors" NOT "Emphysema") AND ("Computed Tomography" OR "Computed Tomographic" OR "CT" OR "CT SCAN" NOT "PET/CT" NOT "Positron Emission Tomography" NOT "Radiotherapy").

In total, 687 articles were retrieved in the first stage of the literature search process.

2.2 Screening Criteria

The purpose of the first stage of screening was to remove duplicate articles. The selection criteria for the publications included were the journal (research articles), articles written in English and published between 2012 and 2020. Then duplicated articles were removed. After the removal of duplicates, a total of 34 articles were prepared for the next stage, known as eligibility.

2.3 Eligibility Criteria

Data for the included studies were extracted using a standardized form according to the following attributes: (1) Data presented in an article must at least show the accuracy, sensitivity, specificity, Dice similarity coefficient, and Jaccard's index; (2) Segmented images from the segmentation method must be compared to the radiologist's manual segmentation, and (3) Evaluation must be done on CT images that are more than 10.

2.4 Data Extraction and Quality Assessment

Data regarding the accuracy, sensitivity, and specificity of the segmentation algorithm for the lung nodule segmentation are recorded as well as the Dice similarity coefficient and Jaccard's index. The total number of CT images that were stated in the articles was recorded as well as where the data sets were acquired. General information related to the interesting data were also included. Those **were** the name of authors, publication time, type of used scanner, acquisition and technical parameter, and image resolution.

Quality assessment was conducted based on the best evidence synthesis, were all retrieved articles eligible for inclusion have to undergo a quality assessment process during the synthesis of results. The studies that meet the predefined inclusion criteria for the review were appraised by the two people based upon predefined guidelines, and the studies are judged as being scientifically acceptable or not. This process relied on the skill and expertise of the reviewer, which might be a source of potential bias. Where a study is not regarded as scientifically admissible, it is rejected from the review process.

2.5 Data Synthesis and Analysis

A qualitative synthesis of quantitative studies was used to qualifying all selected data. Qualitative synthesis refers to the systematic review part, in which there are several aspects of the review process that requires searching, critical appraisal, extraction, and summarising with or without additional quantitative meta-analysis

3. RESULTS AND DISCUSSION

3.1 Literature search

The current research work successfully retrieved a total of 679 articles from the databases. Manual searching efforts were also conducted on other databases based on similar keywords, which resulted in an additional number of 8 articles. In total, 687 articles were then screen based on the inclusion and exclusion criteria. The criteria for the publication include (1) Journal (research article), (2) publication timelines between 2012 and 2020, and (3) published in English.

Overall, a total of 456 articles were excluded based on these criteria, leaving a total of 231 articles to be included. After the removal of 35 duplicates studies, the remaining 196 studies were examined thoroughly based on the title, abstract, and the main contents of all the articles to ensure that they fulfilled the inclusion criteria of the study selection and fit to be employed in the present study in order to achieve the objectives of the current research. A total of 20 full-text articles were assessed for eligibility, and 14 of these articles were excluded because lack of data requirements, thus, leaving only six articles that were included in this systematic review. Two studies focused the thresholding method [9] [10], one study combined the thresholding and region growing method [11], and three studies focused on the region growing method [1] [3] [7]. The study selection was summarised in Figure 1.

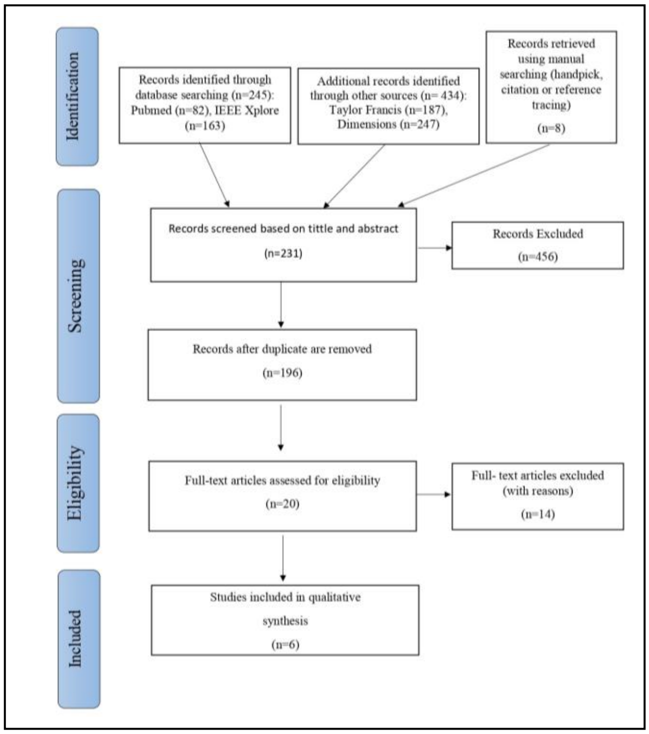

Figure 1. Flow diagram of the study selection process.

Table 1. The summarization of the lung nodules segmentation methods.

Author	Type	Method	Nodule types	Evaluation Metric
Samundesswari et al. [1]	Region growing	ALPE&BR	All types	ACC: 99.4 SN: 98.5 SP: 98.8 Dice:0.982
Kim et al. [3]	Region growing	SPRG	Solid	ACC: 99.9 SN:83.245 SP: 99.97 Dice: 0.87
Kim et al. [7]	Region growing	Seeded region growing	GGN	Dice: 0.6290
Lassen et al. [11]	Thresholding & region growing	Thresholding based region growing.	Sub-solid (n=19)	Jaccard: 0.52
			Sub-solid (n=40)	Jaccard: 0.50
Mukhopadhyay [10]	Thresholding	Thresholding (-500 HU) followed by connected component analysis.	Solid/part-solid nodule	ACC:99.00 SN: 58.00 SP: 99.00
		Thresholding (-800 HU) followed with anisotropic diffusion application.	Non-solid nodule	ACC:98.00 SN: 44.00 SP:100
Guo et al. [9]	Thresholding	OTSU Thresholding	All types	Jaccard: 0.586
		Fixed thresholding		Jaccard: 0.601

*ACC (%): Accuracy. SN (%): Sensitivity. SP (%): Specificity

3.2 Review of Segmentation Algorithm for Lung Nodule Segmentation

3.2.1 Thresholding Technique

Some authors take advantage of the easiness of the threshold and region growing methods (Table 1). Therefore, in this section, explanations regarding the work of authors and results would be explained, along with the advantages and disadvantages discovered by the authors with their proposed methods.

In a study conducted by Guo et al. [9], Otsu and fixed thresholding methods were compared on a total of 85 CT scans images, which consisted of 111 various types of pulmonary nodules. In fixed (or global) thresholding, the threshold value is held constant throughout the image. In contrast, Otsu's thresholding chooses the threshold to minimise the intraclass variance of the thresholded black and white pixels constant throughout the image. To obtain the ideal initial contour by choosing a suitable global threshold was difficult because of the differences in the grey levels of lung nodules [2]. However, in Otsu method had a Jaccard's index of 0.586, while the fixed thresholding method had a higher index, 0.601. If the global threshold value is correctly selected, it will yield the highest value of overlap. This method offers a high calculation speed and accuracy.

Mukhopadhyay et al. [10] employed a thresholding method that involved the application of connected component analysis and anisotropic diffusion application. The authors isolated the lung nodules into solid and part-solid grouped in one category and non-solid in another since these two groups have different density distribution. A solid lung nodule usually lies within the range of -200 HU to 200 HU in the density distribution, while the part-solid lung nodule ranged between −733 HU to 295 HU [13]. The authors selected -500 HU to HU to separate the nodule from lung parenchyma since these nodules are heavily overlapped in this region. They were considering that solid or part-solid nodules are different from the sub-solid nodules. Kauczor et al. [14] reported that the non-solid nodule could be ranged between −750 HU to −300 HU for its density distribution. The selection of threshold at −800 HU, which is well below minimum attenuation for a non-solid nodule, was employed by Mukhopadhyay et al. [10] prior to lung parenchyma removal.

Next, a connected component analysis was performed to obtain the object containing nodule after thresholding. Only sub-solid lung nodules underwent the Anisotropic diffusion filter application on the grey-scale version of the thresholded image after the connected component analysis. If any holes present in the nodule, they will be filled up using morphological closing operation. Lastly, an ellipsoid fitting was performed on all types of nodules irrespective of attachment to reduce computation time.

The procedure of removal of pleural attachment was the same for solid/part-solid and non-solid nodules. The pleural surface removal technique of Kuhnigh et al. [12] is applied in the ellipsoid volume of interest (VOI). Any unwanted pleural surface portion was removed using a refinement step through the morphological dilation application compared to removing the attached blood vessel, the vasculature removal procedure for solid nodule did not apply to non-solid nodules. The solid nodule used a Vasculature Pruning Technique to remove attached vessels from the nodule. In contrast, a Non-solid nodule used Selective Enhancement Filtering for Removal of Vessels in the ellipsoid VOI.

The accuracy and specificity values of thresholding value at -500 HU were measured as 99% and 58%, respectively. For thresholding value at -800 HU, it was measured as 98% for accuracy and 100% for specificity. Contrarily, the sensitivity showed as 44%. The authors faced a problem where about 6 out of 891 lung nodules were failed to be segmented. They identified that the pruning technique assumes there were no vascular attached to a well-circumscribed nodule.

3.2.2 Region Growing Technique

Lessen et al. [11] has proposed a 3D method starting by defining a region of interest around a pulmonary nodule using a user-drawn stroke. A threshold-based region growing approach in a 6-connected neighborhood was performed for sub-solid nodules based on intensity analysis of the nodule region and the surrounding parenchyma. The chest wall was then removed by a combination of connected component analysis and convex hull calculation. A convex hull helps to reduce the over- and under-repair when separating the attached structures. Finally, attached vessels were detached by morphological operations. The proposed method can handle subsolid nodules of different densities automatically. The authors conducted two sets of studies where different radiologists acquired the ground truths. Study A used a total of 19 subsolid nodules that yield a Jaccard's index of 0.52. Whereas in study B, Jaccard's index was measured at 0.50 on a total of 40 subsolid nodules.

The advantage of the threshold-based region growing technique is, it offers a solution for a large set of subsolid nodules (n=59) collected from many different institutes and subjects. It was found that the adjacent vasculature is still included in the segmentation of the nodule (n=?), although the vasculature removal is persistent in follow-ups.

Kim et al. [7] conducted several methods using 2D semi-automatic Segmentation Methods for Persistent Ground Glass Nodules on Thin-Section CT Scans. One of the methods included the region growing method, particularly the Seeded region growing. The ground truth is the result of the manually segmented area by two radiologists. The mean Dice similarity coefficient for this method was 0.629 and been performed on 40 persistent GGNs.

The region growing method is well-known as a tool that is efficient at extracting a homogenous region in an image. Additionally, this method is efficient and robust in dealing with attenuation variations, for instance, attenuation variations caused by the pathologic conditions and imaging artifacts by reinforcing spatial neighborhood information and a regional term. In any case, post-processing is needed after the application of the region growing method because of the noise magnitude and the precision neighborhood criteria, which can suffer from false positive [16]. Furthermore, manual interaction or another intelligent algorithm for the selection of seed may be preferred to be run before the segmentation process. The delineation process is manually select the seed to find a suitable voxel or region.

To improve the accuracy of the traditional region growing method, Kim et al. [7] performed a 3D Spherical region-growing method (SPRG) on CT lung images that contained the solid nodules. The authors explained, SPRG is a modification of an existing region-growing method that is based on a sphere instead of pixels. Usage of the sphere in this method prevents the leakage of growing into the neighboring tissues during the nodule detection since there was some requirement to add a neighboring voxel. They obtained an accuracy of 99.9%, sensitivity of 83.245%, and a specificity of 99.97%. The Dice similarity coefficient was about 0.87. A negative point about this work concerns the validation of the system in which it is merely tested with only solitary nodules. Thus, it is not guaranteed that this system presents the same performance in other circumstances because the system was not tested with a broad range of types of nodules.

Samundesswari et al. [1] conducted a segmentation study using a 3D Automatic Lung Parenchyma Extraction, and Border Repair (ALPE&BR) consists of Automatic Single Seeded Region Growing (ASSRG) on 30 CT scan lung images acquired from LIDC/IDRI public database. The authors claimed that this is a Hybrid Segmentation and Refinement Method for Automatic Lung Cancerous Nodules Extraction. ALPE&BR consists of five main steps. Avoiding the over or under segmentation effect and easing the process of automatic seed selection are the main points why the authors divide the lung into the right and left the region. Then, the isolated type of nodule was segment using the ASSRG Algorithm. The border concavity closing makes use of the morphological operation, connected component analysis, logical rules to extract the lung lobe region based on its area.

The inner and outer boundary of the thick boundary line was detected using the canny edge detection. Calculation of the minimum area of the entire lung lobe region eventually prevents the over the segmented region of the lung wall. Then, elimination of the small vessel structures, non-cancerous nodules that is less than 1cm, or other impurities within the lung parenchyma were done through a Connected Component Analysis (CCA) and Threshold Based Mathematical Nodule (TBMN) refinement.

The acquired images consist of all types of lung nodules. The accuracy, sensitivity, and specificity values in this study were measured as 99.4%, 98.5%, and 98.8, respectively. The Dice similarity coefficient was 0.982. The only dilemma was that the authors separated the left and right lung to achieve the highest accuracy value. They assumed that the inclusion of other findings within both sides of the lung parenchymal could thoroughly affect the performance of nodules detection accuracy. Elimination of other findings in the lung parenchyma included a small non-cancerous nodule (<1cm), vessel structures, bronchioles, and other impurities in the lung parenchyma to mimic clinical situations may consume much time and inefficient to be applied [1].

Choosing segmentation methods that give the most adequate and accurate evaluation metrics eventually assists radiologists in defining the pulmonary nodule and predict the degree of malignancy through the feature extraction and classification. Strengths and flaws were the indicators of choosing the desired segmentation methods. The complexity level of the nodule directly reflects the performance of the segmentation method.

There were some promising results of the segmentation methods [1]. The ALPE&BR method that consists of ASSRG showed a sensitivity of 98.5% (Accuracy;99.4 %, specificity;98.8%) and the Dice similarity coefficient that was near to 1 [1]. However, validation of these systems was limited to tests with up to 30 CT scans, although being tested on all types of lung nodules. The authors did state that within these 30 CT Scan images, there was variation regarding the lung nodules that consist of solid, partially solid or non-solid, dissimilar texture (shape, size, area, and perimeter) and different locations (isolated, multinodule, juxtavascular and juxtapleural) to evaluate the effectiveness of the proposed method. However, it can be among 30 CT scans; there were some nodules not selected equally in terms of quantity, which may directly affect these results. Also, the authors purposely separate the right and left lung to ease the lung nodule segmentation to achieve the highest sensitivity, although, in the clinical setting, the simultaneous segmentation method helps enhance the radiologist efficiency.

The SPRG method [3] stood the second highest to the ALPE&BR method [1] with a sensitivity of 83.245%. They only used one type of nodule, which was the solid nodules were one of the easiest to segment. Thereby, the same performance may not guarantee to present the same performance in other conditions, considering this method was not tested with a vast range of the lung nodules.

Lessen et al. [11] have achieved the satisfactory for the segmentation processing speed where they manage to segment the subsolid nodule in less than 1 second of computation time. However, not all the acquired articles have provided such information. Some authors have overlooked this information; the sufficient algorithm time is needed along with a high sensitivity segmentation method so that this information can be used and compare with other methods since robustness is needed once the methods are applied in the clinical environment.

Moreover, there was a lack of standardized evaluation metrics to assess the segmentation performance, which can curb a correct comparison between studies. A standardized evaluation metrics can ease a future researcher in developing a systematic review that is appropriately evaluated and compared among the selected publication. Besides, additional surveys on the lung segmentation methods are needed in the technical aspects. These aspects can be varied from the ability of the segmentation methods to segment the pulmonary nodule in different scanning parameters to the degree of severity of the inhomogeneous of the nodule.

Most methods achieved an accuracy that was more than 97%. However, sensitivity is more appealing to the radiologist in terms of evaluating the segmentation performance. Therefore, only works by Kim et al. [3] and Samundesswari et al. [1] showed the highest sensitivity. Nevertheless, it is necessary to improve the specificity of the purpose methods in future research.

4. CONCLUSION

The ALPE&BR method that consists of ASSRG was considered the best segmentation method in segmenting the lung nodule in CT images among the acquired articles although, this method needs to undergo several things such as using a larger amount of sample image with a variety of lung nodule, hence confirming its ability to perform lung nodule segmentation accurately.

ACKNOWLEDGEMENTS

Authors thanks the librarian for assistance in resource searching.

REFERENCES

[1] Samundeeswari, P., & Gunasundari, R. (2019). A novel hybrid segmentation and refinement method for automatic lung cancerous nodules extraction. *International Journal of Recent Technology and Engineering*, *7*(6), 28–35.

[2] Xiao, X., Zhao, J., Qiang, Y., Wang, H., Xiao, Y., & Zhang, X. (2018). applied sciences An Automated Segmentation Method for Lung Parenchyma Image Sequences Based on Fractal Geometry and Convex Hull Algorithm. *Applied Sciences*.

[3] Kim, Y. J., Lee, S. H., Lim, K. Y., & Kim, K. G. (2018). Development and Validation of Segmentation Method for Lung Cancer Volumetry on Chest CT. *Journal of Digital Imaging*, *31*(4), 505–512.

[4] National Cancer Institute. (2018). Malaysian study on cancer survival (MySCan). In *National Cancer Institute, Ministry of Health Malaysia* (Vol. 4).

[5] Nanusha Assistant professor ECE, Institue of Aeronautical Engineering, H. (2017). Lung Nodule Detection Using Image Segmentation Methods. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume*, *6*(7), 690–697.

[6] Uzelaltinbulat, S., & Ugur, B. (2017). Lung tumor segmentation algorithm. *Procedia Computer Science*, *120*, 140–147.

[7] Kim, Y. J., Lee, S. H., Park, C. M., & Kim, K. G. (2016). Evaluation of semi-automatic segmentation methods for persistent ground glass nodules on thin-section CT scans. *Healthcare Informatics Research*, *22*(4), 305–315.

[8] Soon Hin How, Teck Han Ng, Yeh Chunn Kuan, Abdul Rahman Jamalludin and Abdul Rani Fauzi. Survival of lung cancer patients in a resource-limited country. Asia Pac J Clin Oncol 2015; 11(3): 221-7

[9] Guo, W., & Li, Q. (2014). Effect of segmentation algorithms on the performance of computerized detection of lung nodules in CT. *Medical Physics*, *41*(9), 1–8.

[10] Mukhopadhyay, S. (2016). A Segmentation Framework of Pulmonary Nodules in Lung CT Images. *Journal of Digital Imaging*, *29*(1), 86–103.

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

[12] Kuhnigk JM, Dicken V, Bornemann L, Bakai A, Wormanns D, Krass S, Peitgen HO. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. IEEE Trans Med Imaging. 2006;25(4):417– 434.

[13] Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ. On measuring the change in size of pulmonary nodules. IEEE Trans Med Imaging. 2006;25(4):435–450.

[14] Kauczor HU, Heitmann K, Heussel CP, Marwede D, Uthmann T, Thelen M. Automatic detection and quantification of groundglass opacities on high-resolution CT using multiple neural networks: comparison with a density mask. Am J Roentgenol. 2000;175(5):1329–1334.

[15] Sharma, P., Mehta, M., Dhanjal, D. S., Kaur, S., Gupta, G., Singh, H., Thangavelu, L., Rajeshkumar, S., Tambuwala, M., Bakshi, H. A., Chellappan, D. K., Dua, K., & Satija, S. (2019). Emerging trends in the novel drug delivery approaches for the treatment of lung cancer. *Chemico-biological interactions*, *309*, 108720.

[16] Mansoor, A., Bagci, U., Foster, B., Xu, Z., Papadakis, G. Z., Folio, L. R., Udupa, J. K., & Mollura, D. J. (2015). Segmentation and Image Analysis of Abnormal Lungs at CT: Current Approaches, Challenges, and Future Trends. *Radiographics: a review publication of the Radiological Society of North America, Inc*, *35*(4), 1056–1076.