International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Involuntary Lung Nodules Recognition in CT Images using 3D Features Mining and Neural Systems

Jasbeer Narwal

---***---

Abstract - The use of image processing and visualization methods for volumetric CT data sets may enable radiologists to identify tiny lung nodules more easily. For instance, it has been claimed that improving tiny nodule detection involves reconstructing CT images with close inter scan spacing and interpreting images using cine rather than film-based viewing techniques. A challenge in image processing is nodule detection. The objective is to locate the locations (and shape) of particular diseased nodules in the lungs. A nodule is a tiny, rounded lung lesion or a worm-shaped lung lesion attached to the pleura (the lung boundary) that have a radio density larger than lung parenchyma.

***Key Words*:** Computed tomography, National Lung Screening Trial, Low-Dose Computed tomography, Computer-aided Diagnosis, Region of Interest

1. INTRODUCTION

Lung cancer is the prominent source of cancer-related mortality worldwide. The American Cancer Society estimates that lung cancer accounts for 28% of all cancer associated demises in 2015; and the 5-year survival rate is only 17% on average. One major fact related to this low survival rate is that only 15% of lung cancers are diagnosed in an early stage, when no obvious cancer symptoms are evident. In contrast, survival increases upwards of 54% if the lung cancer is detected early on. Initial discovery of lung cancer is a motivating issue, with research ongoing to optimize identification of nascent disease.

1.1 Computed Tomography for Lung Cancer Screening

Computed tomography (CT) is now the most widely used screening modality to detect early stage lung cancer. In 2011, the landmark National Lung Screening Trial (NLST) showed a 20% mortality reduction for individuals with lung cancer who underwent screening using low-dose CT (LDCT) relative to plain chest radiography. Subsequently, based on this evidence the United States Preventative Services Task Force (USPSTF) gave a Grade B recommendation that annual screening for lung cancer with low-dose CT (LDCT) be performed in adults ages 55-80 who have a 30 pack-year (number of packs of cigarettes smoked per day multiplied by the numeral of ages an individual has smoked) smoking history and either presently smoke or have leave in the previous 15 years. This policy has spurred development and implementation of new lung cancer screening programs using LDCT.

Understanding of CT scans is both labor-intensive and potentially challenging. Initial phase lung cancer manifests itself as pulmonary nodules, which appear as small round or oval shaped opacities on CT studies with diameters less than 30mm. With a rising amount of CT scans to read, as well as the increasing resolution (e.g., typical thoracic CT scans presently have 200-500 slices), interpreting such large sets of data may lead to visual fatigue and/or strain, contributing to a decrease in diagnostic accuracy. In addition, not as much of knowledgeable radiologists have marked variability in detecting subtle lung cancers, as understanding deeply depend on previous practice. Substantial variability in the performance between radiologists has been reported for the discovery of lung nodules. The rich airway and vessel structure further complicates the understanding of CT scans [1]. Even among experienced radiologists, CT screening yields a large number of false positives, leading to a severe over-diagnosis. For example, in the NLST, a total of 96.4% of the positive screening results in the low-dose CT group were found to be false positives. The benefits of CT screening for detection of early lung cancer is likely to be reduced by the high false positive rates due to benign nodules; reducing these false positives and identifying patients who need intervention could reduce costs and morbidities associated with unnecessary invasive interventions. Thus, it is increasingly imperative to distinguish benign from malignant nodules. However, there are substantial challenges related to distinguishing benign from malignant nodules, and indolent vs. aggressive cancers. As such, computer-aided diagnosis (CAD) systems have been actively studied to assist physicians in solving this problem [2].

2. RELATED WORK

Nisha et al. objective of our work is to take CT scan images and perform segmentation of their using OTSU's thresholding method and various shape parameters include optimal thresholding, area, energy, entropy etc. On the basis of their parameter back propagation network is trained to classify the tumor as per its extent. If the calculated parameter value above the threshold value then extent of cancer is high otherwise low. This work has been done on few CT images and results are analyzed graphically as well as numerically [5]. Preprocessing and segmentation make up the first two stages of the primary phase, according to B. C. Preethi et al. Via a wiener filter, contrast is enhanced using contrast-limited adaptive histogram equalization during the preprocessing stage (CLAHE). To extract the lung tissues, segmentation using Otsu thresholding is done in the second stage. The primary phase of the lung nodule identification system's output is the segmented lung tissues [6]. By utilizing a deep convolutional neural network (CNN) that has been trained for the categorization of medical pictures, Adnan Qayyum et al. present a frame of profound knowledge for the CBMIR system. The network is trained using an intermodal dataset made up of twenty-four classes and five modalities. To get medical images, learnt features and classification results are applied. The best retrieval outcomes come from using class-based predictions. For the retrieval task, a mean typical cataloging accurateness of 99.77 percent and a nasty mediocre exactness of 0.69 are achieved [7]. A single feature-based lung nodule identification approach was put forth by Xuechen Li et al. in 2018. To mine the quality features, we applied the fixed wavelet alter and union catalogue filter. AdaBoost was then used to create a white nodule-likeness map. To measure candidates' levels of isolation, a lonely feature was developed. Concluding assessment of lung nodule contenders was based on both the level of isolation and the resemblance to a white

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

nodule [8]. To reduce the amount of time doctors spend interpreting computer tomography (CT) scan images, K. Senthil Kumar et al. investigate a quick image segmentation technique for medical images. Utilizing adaptive histogram equalization improves the intensity of the image [9]. A deep learning architecture developed by RuoXi et al to combine the grained features from PET and CT images to enable the noninvasive analysis of lung cancer [10].

3. PROPOSED WORK

For a better understanding of the implemented methodology some considerations should be done on the dataset and on the data type of the employed images. CT images are DICOM®, which is a standard format for medicinal imageries and their related information, which allows the storage and the exchange of the data with the required quality for clinical purposes. Moreover, it is recognized as ISO 12052 (International Organization for Standardization) and it is implemented in the majority of the medical imaging devices. For this motive the formation of a pipeline able to manage this file structure is essential for the establishment of an analytical tool exploitable in the medical field [1] [3] [5].

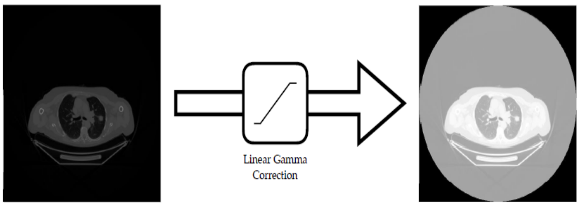

Figure 3.1: Contrast adjustment with linear gamma correction

The employed images are composed by the raw image, which is a matrix that, considering both the CT and the PET, is stored as grayscale images with a bit depth of 16 for each pixel. Additionally, the DICOM® provides the meta-data structure in which some information are stored such as the dimensions of the matrix containing the acquired image, the image type and the location of every slice based on the integral reference system to the table of the machinery. Considering all these data, each DICOM® image is around 520 KB for the CT and 40 KB for the PET, reflecting the different dimensions of the matrices. Meta-data are not reliable information since they can be customized by the different manufacturer. For this reason the created pipeline will rely, when possible, on the image itself instead of such meta-data. Being the aim of this work not only to help the physician in the identification of the cancer lesions but also the construction of a apparatus able to standardize the construction of datasets for machine learning approaches, it is clear the need to limit as much as possible the employment of variable characteristics [4, 6].

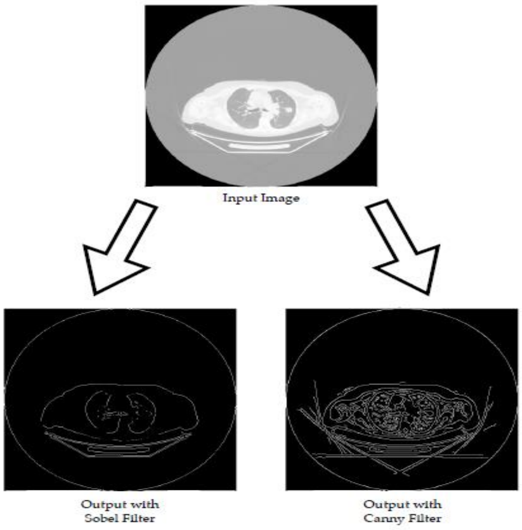

Figure 3.2: Sobel vs. Canny edge detection approaches

CT images present the gantry artifact which has been defined as the circular shadow around the patient that prevents the correct working of the processing pipeline. In order to identify it presence and the subsequent elimination, as first step an adjustment of the contrast has been performed using a gamma correction algorithm. As visible in Figure 3.1, starting from a gray scale image the output is still a gray scale image where the top and bottom 1% of all the pixel values are saturated. On the obtained adjusted image it is now possible to apply the edge detection using the Sobel kernel which, if compared to the Canny as shown in figure 3.2.

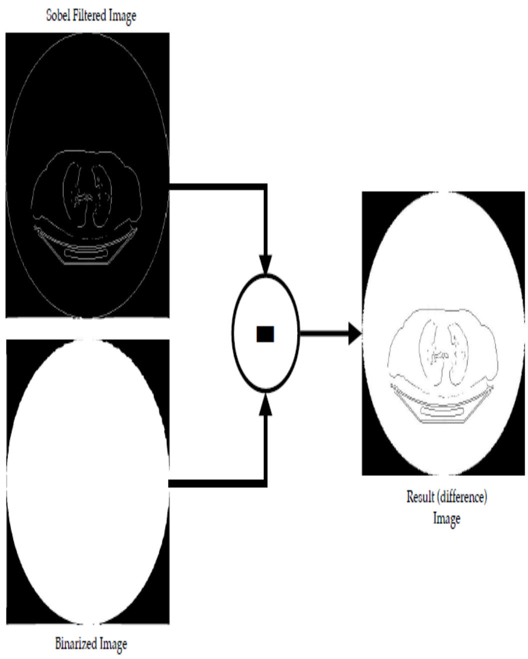

Figure 3.3: Creation of the mask that will be employed for the removal of the gantry artifact

For the creation of the mask that will be later used for segmenting the lungs it is essential to get a binary image where only the lungs should appear as white spots as shown in figure 3.3. All the small details inside of them, still visible in Figure 3.4, should be deleted.

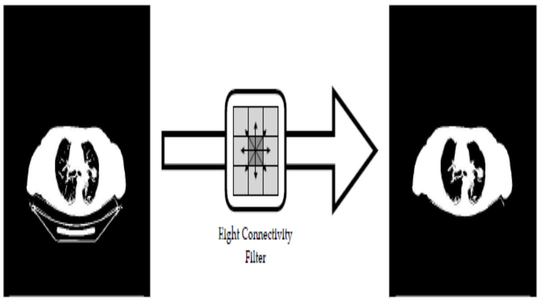

Figure 3.4: Identification of the major associated constituent centered on the 8-connectivity filter

This approach is based on the suppression of structures that are connected to the image border and that result lighter if compared to their eight adjacent pixels. So considering that all the pixels outside the thorax are white they certainly are lighter compared to it. Moreover, they are all connected to the image perimeter according to the definition of 8-connectivity. Satisfying both the conditions all the pixels outside the thorax can be set to black (background) [7].
Having correctly selected the contour of the lungs, there is now the need of creating a mask, where all the small linear structures inside the lungs are removed in order to create them distinguishable later on during the merging of the data coming from both the PET and the CT. In fact, all the details inside the lungs should be preserved on the original CT after the filtering with the created mask, since they can be possible tumors. The applied technique for the removal of the linear structures is a morphological closure of the image is a procedure composed of two sequential steps: dilation followed by erosion. These algorithms require the employment of a constructing element, which is used to probe the pixels surrounding the one selected and to decide if that specific element is part of a construction alike to the shape of the probe. The dilation process fills the possible holes present in the image in order to better fit the shape of the constructing element, while the erosion has the opposite effect. Applying the dilation followed by the erosion allows the preservation of those shapes that match the chosen probe and the filling of possible white isolated pixels inside those regions. Taking into account what have been said, in order to remove the linear ramification departing from the central structure, it has employed a round structuring element for preserving only the circular shaped region [5] [9]. At this point it is necessary to choose the radius of the probing circular element.

Figure 3.5: Morphological closure with different constructing element radius; the green one is the selected one

Figure 3.5 shows some attempts with different radii: experimentally it has been established that a good tradeoff is a radius equal to 12, which is the one not only able to remove all the linear structures but also to preserve the main circular shaped regions. Now that the mask has been obtained, the last step of the CT processing pipeline is the filtering of the preprocessed image with the morphological closed image [10]. In this way the segmented ROI corresponding to the lungs contour and their inside details is obtained as shown in figure 3.6.

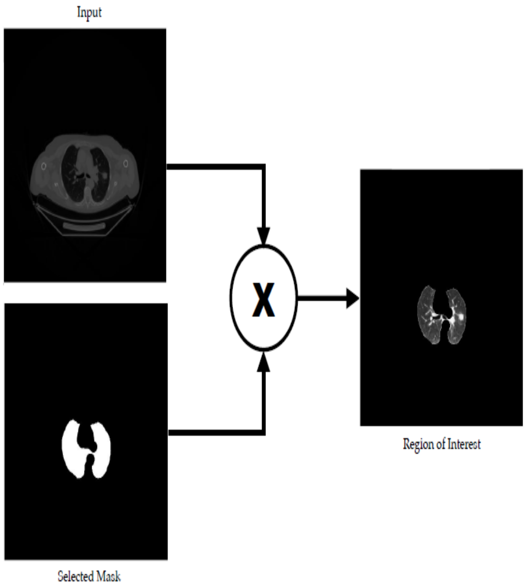

Figure 3.6: Identification of the Region of Interest within the CT processing pipeline

4. RESULTS AND DISCUSSION

LIDC-IDRI database contains CT scans of more than 1,000 patients, in which each nodule has been examined by four experienced radiologists. Some of the images from dataset are shown in figure 4.1. We developed a Feed Forward Neural Network model to automatically learn feature representation for retrieval. The model was trained to detect circle for each CT image. Each lung nodule is composed of a set of slices, and each slice has an associated feature vector.

To calculate the resemblance among two nodules A and B, the distances between each slice of A and all the slices of B are added together and divided by the total number of slices.

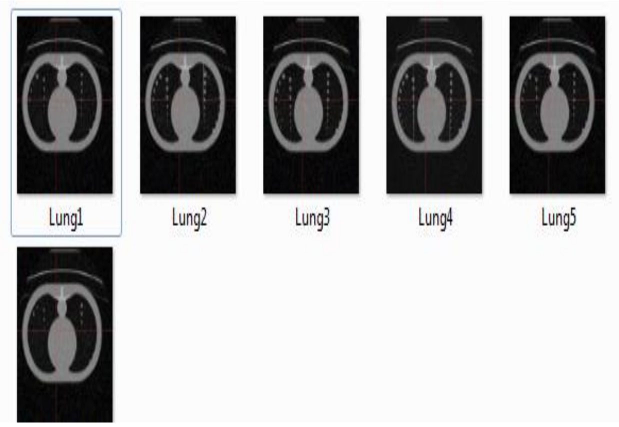

Figure 4.1: Sample Dataset

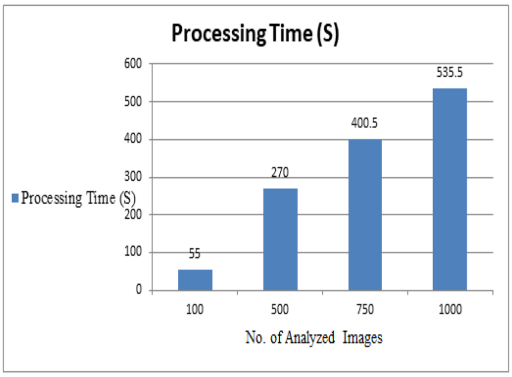

Figure 4.2: Execution time on different dataset dimensions

Considering the execution time, an expert radiologist usually employs between five and ten minutes to analyze a patient acquisition; this time may further increase if the physician is distracted by external factors or if he/she is working from too many hours. Looking at Figure 4.2 and keeping in mind that generally a single acquisition is composed by 200-300 images; the proposed tool is able to analyze the same number of images in less than three minutes. Figure 4.2 reports the implementation period measured on dissimilar dataset dimension (from 100 images to 1000 images). The experimental results show that the computation time goes from 55.00 seconds for 100 images to 536.50 second for 1000 images. Figure 4.2 does not shows the standard deviation since it is not significant: however it ranges between 0.02 and 0.04. Moreover, as it is possible to notice from the same Figure, considering only the pure computation time without the image loading from the disk, execution time is reduced by 20%. In addition, growing the count of images, the execution time grows pretty much linearly, while, on the contrary, a physician will decrease the performance in terms of analysis time with the increase of the number of images analyzed due to his/her tiredness. It is significant to sign that it is promising also to further reduce the total execution time of the proposed tool since it is possible to make some optimizations: as an example, thanks to a statistical analysis, it is possible to discharge a percentage of the acquired slices, which are the ones covering the abdominal region, which is out of the Region of Interest.

Moving on to the assessment of the accuracy some quantitative statistics have been computed.

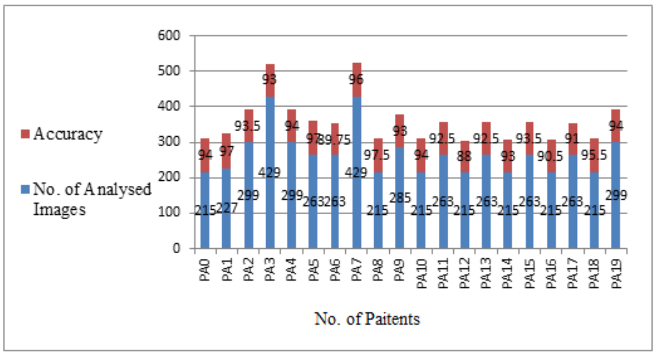

Figure 4.3: Accuracy by considering the number of analyzed images

5. CONCLUSIONS

CLAHE (Contrast Limited Adaptive Histogram Equalization) technique has worthy controllability in picking local histogram mapping task. This technique splits the image into suitable sections and applies histogram equalization to them. It adjusts the intensity standards of the image by retaining a nonlinear practice in direction to make the most of the contrast for all pixels of the image. Being more precise, this work proposed both a methodology and the related software tool that taking as input Digital Imaging of chest and by exploiting the characteristics of them, is capable of automatically identifying the lungs and the possible presence of tumor lesions. The experimental results shown as the obtained accuracy varies between 85-95%. In some cases some slices which do not contain the lungs have been identified as covering a tumor (false positives); in order to increase the accuracy, as future work it is possible to automatically remove a percentage of images decreasing the possibility to create false positive in areas that are certainly not lungs

REFERENCES

[1] Shelda Mohan, M. Ravishankar, "Modified Contrast Limited Adaptive Histogram Equalization Based on Local Contrast Enhancement for Mammogram Images", Springer-Verlag Berlin Heidelberg 2012

[2] A.R. Talebpour, H.R. Hemmati, M. Zarif Hosseinian, "Automatic Lung Nodules Detection In Computed Tomography Images Using Nodule Filtering And Neural Networks", The 22nd Iranian Conference on Electrical Engineering (ICEE 2014), May 20-22, 2014, Shahid Beheshti University

[3] Liguori, C., Frauenfelder, G., Massaroni, C., Saccomandi, P., Giurazza, F., Pitocco, F., Marano, R., and Schena, E.: Emerging clinical applications of computed tomography. Medical Devices (Auckland, NZ), 8:265, 2015

[4] Khan, M. D. H., Ahmed, M., and Bach, C.: Preliminary detection and analysis of lung cancer on ct images using matlab: A cost-effective alternative. Journal of Biomedical Engineering and Medical Imaging, 2(6):01, 2016.

[5] Nisha, Lavina Maheshwari, "Lung Tumor Detection by Using Image Segmentation and Neural Network", International Journal of Enhanced Research in Science, Technology & Engineering, December-2015

[6] B. C. Preethi, Gia Elizabeth Abraham, "Lung Tissue Extraction Using OTSU Thresholding in Lung Nodule Detection from CT Images", International Journal of Current Trends in Engineering & Technology, Volume: 02, Issue: 06, 2016

[7] Adnan Qayyum , Syed Muhammad Anwar, "Medical image retrieval using deep convolutional neural network", Neuro computing 266 (2017) 8–20

[8] Xuechen Li , Linlin Shen and Suhuai Luo, "A Solitary Feature-Based Lung Nodule Detection Approach for Chest X-Ray Radiographs", IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 22, NO. 2, MARCH 2018

[9] K. Senthil Kumar, K. Venkatalakshmi, and K. Karthikeyan, "Lung Cancer Detection Using Image Segmentation by means of Various Evolutionary Algorithms", Hindawi Computational and Mathematical Methods in Medicine
Volume 2019, Article ID 4909846, 16 pages https://doi.org/10.1155/2019/4909846

[10] RuoXi Qin,Zhenzhen Wang,"Fine-Grained Lung Cancer Classification from PET and CT Images Based on Multidimensional Attention Mechanism", Hindawi Complexity Volume 2020, Article ID 6153657, 12 pages https://doi.org/10.1155/2020/6153657