A Review on Lung Cancer Detection Analysis

1Preeti Bala, 2Aashta Mahajan

1Research Student, Department of Computer Science & Enng., Swami Sarvanand Group Of Institutes Of Management and Technology, Punjab, India
2HOD, Department of Computer Science & Enng., Swami Sarvanand Group Of Institutes Of Management And Technology, Punjab, India

Abstract: Recently, image processing techniques are widely used in several medical areas for image improvement in earlier detection and treatment stages, where the time factor is very important to discover the abnormality issues in target images, especially in various cancer tumour such as lung cancer, breast cancer, etc. Image quality and accuracy is the core factors of this research, image quality assessment as well as improvement are depending on the enhancement stage where low pre-processing techniques are used following the segmentation principles, an enhanced region of the object of interest that is used as a basic foundation of feature extraction is obtained generally. Relying on general features, a normality comparison is made. In this research, the main detected features for accurate images comparison are pixels percentage and mask labeling.

Keywords: Cancer Detection, Image processing, Feature extraction, Masking, Segmentation, Classification.

I. INTRODUCTION

Computed Tomography (CT) has outperformed conventional radiography in the screening of lungs because it generates very detailed high-resolution images and can show early stage lesions that are too small to be detected by conventional X-ray. CT has been widely used to detect numerous lung diseases, including pneumoconiosis, pneumonia, pulmonary edema, and lung cancer [1]. Early detection of diseases is very crucial for treatment planning. However, it is considered one of the most challenging tasks performed by radiologists due to the huge amount of data generated by CT scan. Therefore, computer-aided diagnostic (CAD) systems are needed to assist radiologists in the analysis and evaluation of CT scans.

A CAD system analyzes medical images in several steps; first a preprocessing step for noise reduction and enhancing the image quality and then segmentation step to differentiate region of interest (ROI) from other structures in the image. After segmentation, different features such as geometrical, textural, and statistical features are extracted. Finally, a classification/evaluation step is done to evaluate and diagnose the ROI based on extracted features.

Many efforts have been made to provide computer aided diagnosis for lung images. Lung segmentation is a necessary step; it has progressed from manual tracing to semi-automated to fully automated segmentation. Here, some automated lung segmentation studies are presented [2–9]. Other studies present content-based image retrieval (CBIR) systems for lung images [10–15]. Earlier work in classification of lung cancer includes the work of Patil and Kuchanur [16] and Kuruvilla and Gunavathi [17] that used artificial neural networks to classify lung cancer images based on the features extracted from lung segmented images. Nevertheless, Patil and Kuchanur used geometrical features for classification and achieved only 83% accuracy of classification. And Kuruvilla and Gunavathi used statistical parameters as features for classification and achieved accuracy of 93.3%. Another work by Depeursinge et al. [18] classified different lung tissue patterns using discrete wavelet frames combined with gray-level histogram features. However, the main limitation of this work was the lack of resolution in scales with the decomposition.

Lung cancer detection in early stage is the important part and key to cure. Generally, it can be diagnosed with the help of X-rays, Chest films, CT scan, MRI, and with many more techniques. Among all these, there is a vital measure called neurotic determination that analyses examples of needle biopsies that will be acquired from the bodies of subject that is going to diagnosed. In last decades it was very difficult to diagnosed Lung cancer at early stages with the help of image processing and pattern recognition. But with the time new hybrid Lung cancer detection method come into existence and accuracy of diagnosis also get improved. Many methods were proposed that were based on the subtraction between two serial mass chest radiography. This was proposed to detect new lung nodules. Some methods are used to extract and analyzed features of the pulmonary blood vessel region and lung and then diagnosed with different rules. Result obtained from diagnosis was used for the detection of tumour. Until now there is rigorous theory indicate that gives error free approximation. Artificial neural network have already been widely exploited in this research but now it get improved with the generalization ability which include training of several artificial neural networks and combining their predictions. Lung cancer diagnosis using X-ray chest film gives more effective results.
Operators of morphology include morphological operations. Morphological operations are applied on binary images. Binary image is the image with only two colors black and white. They are used to get the information about shape, convex hull /region of images. Next is extraction of features. Which is very important step and that is used to detect and separate the shape of given images. To predict the probability of lung cancer presence, we are using: linearization and GLCM, these two methods works on anatomy of lung and works on computing tomography. Grey Level CoOccurrence Method is a method which is used to show the information in tabular form, it is used to show how different combinations of pixel value present in an image.

Non-small cell lung cancers represent approximately 90% of all kind of lung cancers as well as these are not as much spread in the direction of several other tissues in addition with several organs sluggishly as occurred in small cell cancers. This type of lung cancer occurs more typical in assessment with smaller kind of cells or moveable lung cancer tissues, by accounting just about 85% up to 90% related with lung tissues. The three foremost sorts of non-small cell or transferrable lung cancer tissues are normally mentioned as pertaining to the specific kind of cellular material present in the tumor [9]:

1. Squamous cell or movable epidermis carcinoma: The article every so often would instigate in the inner part of bronchi nearby lung part.

2. Adenocarcinomas: These kind of cellular materials ordinarily would activate on the external boundaries on the lung zone. The aforementioned does not take utmost traditional sort of lung cancer tissue in some specific individuals those in no way being as a smokers.

3. Large cells or movable epidermis carcinomas: They're a small clustering of tissues devouring considerable, anomalous-looking cellular material. Most of these kinds of tumors could probably start from some place in lung region

II. LITERATURE SURVEY

Once the nodules are detected and segmented, it has to be diagnosed into benign or malignant nodule. Diagnosis is based on the shape and appearance of the nodule. Commonly used classifiers are linear discriminant classifiers and neural network classifiers. Features used for classification include Geometrical features, Texture features, Histogram features, Gradient features and spatial features. Geometric features include features such as spherical disproportion, circularity etc. Texture features include features like contrast, energy, entropy, etc. Histogram features are average, standard deviation, skewness, etc. Gradient features describe the average, standard deviation, kurtosis, etc. Spatial features are the location of the nodule. Kawata et al [5] proposed a CAD system that classifies the primary nodules based on curvature index and the relationship of the nodules with their surrounding features. The curvature index specifies the surface type (i.e. ridge, saddle, pit etc.) and the curvedness of the nodule. Curvedness specifies the degree of curvature. The maximum value obtained in their work is only 0.94. McNitty-Gray et al. [6] developed a CAD system which is based on pattern classification approach for determination of the nodules. They used the features shape, size, attenuation, and texture. Stepwise discriminant analysis is used for selecting features. The method showed an accuracy of 90.3%. Shah et al [7] in their CAD system extracted the feature from solid part of the nodule and ground glass components. The features used are nodule attenuation, shape and texture. The Az value obtained is 0.92 for 35 nodules. Classification based on morphological and textural features
is developed by Way et al [8][9]. This system achieved an Az of 0.85 with 152 patients. The initial neural network based CAD system was developed by Gurney and Swensen [10]. They used the nodule border smoothness, speculation, and population as the features for classification. Matsuki et al [11] developed another neural network based system with improved performance included the shape-based parameters such as border definition, speculation, and concavity. This system gives Az as 0.951. Henschke et al. [12] developed their CAD system to automatically define and extract features and also to classify the nodules. Neural network’s clustering technique is used for extracting and grouping the features into clusters. In their classification of 28 benign and malignant nodules, only three benign nodules were misclassified. Lo et al. [13] in their work used a combination of Radio graphical parameters and shape indices. Radio graphical parameters include vascularity, CT density distribution, and shape indices include aspect ratio, circularity, irregularity, extent, compactness and convexity. Their results showed the value of Az as 0.89. Suzuki et al [10] classified the nodules by training the MTANN classifier with benign and malignant nodules. They achieved a Az value of 0.89 with 415 patients. Chen et al. [11] used ANN ensemble method to classify benign and malignant nodules and achieved a Az value of 0.915. Also they have test the performance of different feature set with two separate networks. One network is trained with subjective features such as nodule size, shape, marginal irregularity, speculation, border definition, lobulation and nodule density. Other network is trained with matched features that are automatically obtained from chest radiographs. Matched features include effective diameter, degree of circularity, mean gradient, mean pixel etc. Both networks used shape based features also. A high degree of performance Az =0.854 is achieved with objective features compared to subjective features. Iwano et al [13] CAD system is to automatically classify the nodules based on shape features and the results were as accurate as the radiologist’s classification. A two-step supervised learning scheme was developed by Lee et al [14] based on texture and shape features by combining a genetic algorithm with a random subspace method. They achieved a Az value of 0.889. A 2D approach was proposed by El-Baz et al [15] for diagnosing malignant nodules. The approach is based on representing the spatial distribution of Hounsfield values in detecting the nodules with a 2D rotationally invariant second order MGRF. They also proposed [16] a 3D work in which the nodules are represented with spherical harmonics. Special harmonic analysis approximates the 3D surfaces of the detected nodules the accuracy achieved is Az =0.97.

Table 1: Lung Cancer Staging [17]

<table>
<thead>
<tr>
<th>Stage</th>
<th>T= primary tumor, N= regional lymph nodes, M= distant metastasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 0</td>
<td>Carcinoma in situ</td>
</tr>
<tr>
<td>Stage IA</td>
<td>T1 N0 M0</td>
</tr>
<tr>
<td>Stage IB</td>
<td>T2 N0 M0</td>
</tr>
<tr>
<td>Stage IIA</td>
<td>T1 N1 M0</td>
</tr>
<tr>
<td>Stage IIB</td>
<td>T2 N1 M0</td>
</tr>
<tr>
<td></td>
<td>T3 N0 M0</td>
</tr>
<tr>
<td>Stage IIIA</td>
<td>T3 N1 M0</td>
</tr>
<tr>
<td></td>
<td>T1 N2 M0</td>
</tr>
<tr>
<td></td>
<td>T2 N2 M0</td>
</tr>
<tr>
<td></td>
<td>T3 N2 M0</td>
</tr>
<tr>
<td>Stage IIIB</td>
<td>T4 N0 M0</td>
</tr>
<tr>
<td></td>
<td>T4 N1 M0</td>
</tr>
<tr>
<td></td>
<td>T4 N2 M0</td>
</tr>
<tr>
<td></td>
<td>T1 N3 M0</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
III. CONCLUSION

With enough discussion about the lung segmentation methods, it is decided that applying threshold multiple times on the CT image separates the lung region from the background. Multi-level thresholding is the simplest and efficient techniques for lung segmentation. For nodule segmentation any interactive segmentation gives better results than other segmentation methods. Classification of the nodules is based on shape and texture features. Considering the above discussion, the feature-based approach was found to be better for classification.

REFERENCES


