ABSTRACT

Lung cancer is the leading cause of cancer death. It usually does not cause any symptoms in the early stage of its evolution. Most of the victims have been diagnosed in an advanced stage, where the symptoms become prominent, which results in poor curative treatment and high mortality rate. Screening is looking for early signs of lung cancer before a person has any symptoms. Screening tests are recommended for people with high risk of developing the disease such as long history of smoking, coal miners and long-term exposure to carcinogens. Chest X-ray, Computed Tomography (CT) and Sputum cytology have been studied for a long time as the choices for lung cancer screening. Recently, Low Dose Computed Tomography has become the standard for screening lung cancer which lowers the risk by 20% as compared with chest X-rays.

Lung nodule is an important clinical observation in Computed Tomography images. The probability that a nodule can be malignant is about 40%. Distinguishing between pulmonary vessels and nodule is a challenging task since they share similar shape and intensity characteristics. The complexity of nodule detection process increases when the lung nodules are attached with the blood vessel or near the lung wall. Computerized nodule detection schemes have shown substantial increase in diagnostic accuracy of lung cancer detection.

A typical Computer Aided Diagnosis system consists of the following steps: pre-processing, lung parenchyma segmentation, region of interest detection, feature extraction and nodule classification. The first four steps deal with image processing and final classification step encounters the problem of pattern recognition.

The main objective of this work is to develop a Computer Aided Diagnosis system for the early diagnosis of lung cancer from Low Dose Computed Tomography images. The two important requirement of a CAD system are very high sensitivity and specificity. In order to keep these parameters high, the number of false positives detected by the system should be reduced and the system should not miss any positive candidate.

In this research work, the appropriate method has been proposed to cater the requirements of each processing step of the computer aided lung cancer diagnosis system. An un-supervised Gaussian Mixture Model based lung parenchyma segmentation framework has been proposed. This work presents a relatively simple method based on histogram processing for estimating the initial Gaussian parameters. The segmented lung parenchyma experiences an adaptive morphological filtering to reduce the boundary errors.

The proposed lung segmentation method achieves a highest Dice Similarity Coefficient (DSC) of 0.97 as compared to other methods. Hence, it will be a promising precursor to the Computer Aided Diagnosis (CAD) of lung diseases such as lung density analysis, blood vessel analysis, airway analysis, lung mechanics analysis and so on.

The processing time of the CAD system is proportional to the number of ROIs analysed. Therefore, the number of ROIs detected should be reasonably less meanwhile it would not miss any true positive nodules. ROIs in the Low Dose Computed Tomography image appears to be a bright blob like structure with radiating lines growing outwards. The proposed method first enhances the segmented lung parenchyma by Contrast-Limited Adaptive Histogram Equalization technique and then gray scale thresholding is applied to segment the ROIs. This method produces many false positive ROIs. The false positive regions are reduced by applying a rule-based filter that utilises shape, size and intensity characteristics of the lung nodule.

The proposed method takes the advantages of reducing the computational complexity and the production of fewer ROIs as compared with other techniques. It reduces approximately 87% of the false positive ROIs when compared with other thresholding-based ROI detection techniques.

Volumetric segmentation of lung nodule in a three-dimensional space gives more information about the nodule and facilitates quantitative analysis. Images containing artifacts and abnormal lung region will increase the complexity of the nodule segmentation task. Three dimensional volume growing segmentation technique is proposed. A new cost function is derived based on Directionality Histogram for guiding the segmentation algorithm. The proposed method achieves very low false positives per scan as compared with other automated three dimensional nodule segmentation methods.

The performance of the CAD system is highly influenced by the selection of appropriate features that are completely characterising the lung nodules. Texture analysis is widely adapted for classification of objects in an image. The texture features extracted from the nodule region reflect intratumoural heterogeneity that aids in classifying the nodules.

The proposed feature extraction method is based on Dual Tree Complex Wavelet (DTCWT) Transform and Gray Level Co-Occurrence Matrix (GLCM). Four texture features have been derived from the GLCM matrix. These features are computed for different combinations of the GLCM parameters (distance and orientation) and the results are concatenated to form the feature vector. The SVM classifier with the Radial Basis Function (RBF) has been chosen for the classification of the nodules. The results show that the

proposed Computer Aided Diagnosis (CAD) system for lung cancer achieves a high sensitivity of 97.56% and a very low false positive rate of 1.35 per scan. Therefore, the proposed Computer Aided Diagnosis system will help the physician for diagnosing the early signs of lung cancer from the screening Low Dose Computed Tomography images.