Abstract

Lung diseases are the most common disease which causes mortality worldwide. In this study, the computed tomography images are used for the diagnosis of the lung diseases such as normal, small cell lung carcinoma, large cell lung carcinoma and non small cell lung carcinoma by the effective extraction of the global features of the images and feature selection techniques. The images are recognized with the statistical and the shape based features. The texture based features are extracted by Gabor filtering, the feature outputs are combined by watershed segmentation and the fuzzy C means clustering. Feature selection techniques such as Information Gain, correlation based feature selection are employed with Genetic algorithm which is used as an optimal initialization of the clusters. The dataset of lung diseases for four classes are considered and the training and testing are done by the Naive Bayes and random forest classifier. Results of this work show an accuracy of above 80% for the correlation based feature selection method using naive bayes classifier.
- C. Brambilla and S. Spiro, HIGHLIGHTS IN LUNG CANCER, Copyright #ERS Journals Ltd. 2001 European Respiratory Journal, ISSN 0903-1936.
April 2006, pp. 104-110.

**Index Terms**

Computer Science  
Pattern Recognition

**Keywords**

Global features  
Genetic Algorithm  
Image segmentation.