ABSTRACT

The main reason of cancer-related death is lung cancer. The primary symptoms of lung cancer are persistent cough, shortness of breath, persistent fatigue, cough up blood/blood in sputum, loss of appetite, weight loss, persistent chest pain and pain in bone (back/hips)/shoulder/neck/arm. Spread of cancer into the human body refers to the staging of cancer. The stages of lung cancer are assigned from I to IV. Stages I and II are called early stage. Stages III and IV are called later stage. Depending on the staging, symptoms of cancer will differ. Staging and treatment will be the key for patients’ survival with cancer. Patients have better chances of survival if lung cancer is identified in its initial stage, than at the later stage.

The most recent statistics provided by World Health Organization (WHO) indicates that around 7.6 million deaths worldwide each year because of lung cancer. It is expected to raise around 17 million worldwide deaths in 2030. In India, 51,000 lung cancer deaths were reported in 2012, which include both men and women. The only way to control the lung cancer death rate is diagnosis of cancer in its early stage.

The development of Computer-Aided Diagnosing (CAD) helps the physicians and radiologists to diagnose the lung cancer accurately from the symptoms and CT scan images in its early stage. Various techniques are available for the lung cancer detection, but many approaches provide less accuracy and more false positives. So a new approach to detect cancer in early stage is necessary. The main objectives of this research are: to develop a CAD system for lung cancer detection from the observed symptoms and from the chest CT images of patients and to develop a Three Dimensional (3-D) CAD visualization system to find the location of the detected cancerous nodules.
The main aim of these proposed approaches is to accurately detect the lung cancer by analyzing the observed symptoms and CT images of the patients.

To detect the lung cancer from the symptoms, Cancer Assessment Specific (Stage-wise) Questionnaire for Lung (CASQ-L) was prepared. CASQ-L consists of three parts, namely, patient’s demographic information, Lung Cancer Questionnaire (LCQ) and Lung Cancer Specific (Stage-wise) Questionnaire (LCSQ). The volunteer group has been appointed to collect the data orally from the patients through interview by explaining the questionnaire. The significant symptoms from LCQ and LCSQ were identified with Statistical Package for Social Sciences (SPSS) software. After the identification of the significant symptoms, Hybrid Neuro-Fuzzy System (HNFS) was employed in which the neural network detect the presence or absence of lung cancer (significant symptoms of LCQ) and fuzzy system (fuzzy rules framed from the significant symptoms of LCSQ) make the final decision on the staging of the cancer.

To detect the cancerous nodules at early stage (who have confirmed through biopsy and HNFS study) of patients, suspected clusters from their CT images were segmented using Automatic Region Growing with Morphological Masking (ARGMM) and Fuzzy Auto-seed Cluster Means Morphology (FACMM). The suspected clusters comprised of small tiny non-cancerous clusters, irregular line like structures, blood vessels and calcifications (calcium deposition on lungs). The size/shape features, area and eccentricity criteria were applied on the suspected clusters to keep only the initial suspected nodules. Though the ARGMM segmented more number of suspected clusters (most of them were non-cancerous), the number of initial suspected nodules after applying size/shape features were same. Further, the segmentation time per slice was higher for ARGMM than FACMM. Hence FACMM was utilized for further analysis. The centroid shift and texture
features such as contrast, homogeneity and auto-correlation, extracted for the initial suspected nodules from the consecutive CT slices were fed to the Artificial Neural Network (ANN) and Support Vector Classifier (SVM) to detect the cancerous and non-cancerous (blood vessels and calcifications) nodules.

Treating of lung cancer in the initial stage includes surgical removal of the diseased portion of the lung lobe or treating it with chemotherapy or radio therapy or sometimes combinations of these. For this, the surgeons need to know the exact location and size of the cancer. To visualize the cancerous nodules, the 3-D visualization system was carried out using Materialise’s Interactive Medical Image Control System (MIMICS) software.

The developed HNFS system for the lung cancer detection based on the observed symptoms has achieved the sensitivity, specificity, accuracy, precision and F-measure of 100%, 80% and 95%, respectively for the 217 (56 early stage cancerous, 111 later stage cancerous and 50 non-cancerous with lung infections) subjects, who visited Bharat Scans, Chennai, between the months of June 2014 and December 2014. Therefore, HNFS system could be a potential pre-screening tool to identify the lung cancer from the patients who come with symptoms. Similar study reported by Durai et al. (2011) for the lung cancer detection based on the symptom values using fuzzy rule based inference system yielded a sensitivity and specificity of 89% and 76% respectively. Similarly study conducted by Balachandran and Anitha (2014) using ANN achieved a sensitivity of 90%. A study demonstrated by Abhinav et al. (2011) using feed forward neural network and back propagation algorithm provided a sensitivity, specificity and accuracy of 93%, 78% and 90%, respectively.
The extracted centroid shift and texture features from the consecutive CT slices were trained and tested using ANN classifier, which produced the sensitivity, specificity and accuracy of 91%, 89% and 89%, respectively for the 106 (56 early stage cancerous and 50 non-cancerous with lung infections) subjects of Bharat Scans, Chennai. The False Positives (FP) per patient was computed as 0.6. Therefore the implemented CAD system using FACMM and ANN classifier could be the screening tool to detect the lung cancer from the CT images at the early stage (stages I and II). Later, the extracted features such as centroid shift and texture features from the consecutive slices were trained and tested using SVM classifier, which produced the sensitivity, specificity and accuracy of 100%, 93% and 94%, respectively. Similar study reported by Ozekes et al. (2008) for the lung cancer detection based on the CT images using genetic algorithm and fuzzy rule based system yielded a sensitivity of 100% but FP/scan was 13.4. Similar study reported by Kuruvilla and Gunavathi (2014) using morphological segmentation and ANN classifier produced a sensitivity, specificity and accuracy of 91%, 100% and 93% respectively. A study conducted by Demir and Camurcu (2015) using 2-D and 3-D preprocessing and SVM classifier yielded a sensitivity, specificity and accuracy of 98%, 88% and 90% respectively.

The implemented 3-D CAD visualization system using MIMICS successfully segmented the malignant nodules from the CT images of 10 cancerous patients (in the initial stage) and then superimposed on their lung lobes to show its exact location in 3-D. Further, height, length and breadth of the malignant nodules were measured. Similar study conducted for the 3-D visualization system using AMIRA software (Wei et al. 2013) and YaDiV with MATITK software (Aggarwal et al. 2010) segmented only the lung lobes and cancerous region. However, no information was reported about the height, length, breadth and volume of the segmented malignant nodules.