Identification and Classification of Lung Nodules Using Neural Networks

Dr.S.Gnanavel[1], N.Durai Murugan[2], K.E.Narayana[3], M.Sree Krishna[4]

[1,2,3,4] Assistant Professor (SG)

[1,2,3,4] Department of Computer Science and Engineering, Rajalakshmi Engineering College, India

[1] gnanavel.s@rajalakshmi.edu.in

Abstract— Lung cancer is a serious health concern, which is also one of the major types of cancer that has a profound impact on the overall cancer mortality rates. The detection of lung cancer nodules is quite a challenge as the major challenge is the structure of the cancer nodules; here the cells are imbricated with each other. The prediction and classification of lung cancer is done by applying digital image processing techniques to the acquired input images of the nodules. This methodology also aids early detection which in turns reduces the criticality of the condition and provides scope for early intervention and treatment. The prediction methodology involves extracting several features of the lung cancer cell and then applying pattern-based prediction techniques. In recent times, owing to the fact that the time and execution parameters are very important aspects to detect the abnormality of the fast-spreading cancer cells, digital image processing techniques are being widely deployed. The fundamental factors of this research are the quality of image assessment and the precision of feature extraction. Following our proposed methodology, a clear picture of the region of interest is obtained which acts as a basis for the feature extraction process. Here an overall evaluation of the digital image processing techniques used by previous scholars for the finding and classification of lung cancer nodules have also been emphasised.

Keywords— Lung cancer, Neural Network, Pre-processing, Enhancement, Segmentation, Median Filter

I Introduction

Carcinoma of the lungs, most commonly known as lung cancer is triggered due to the uninhibited growth of cell nodules in the lung tissues.[4] The survival rates of lung cancer are a very low percentage as compared to the other forms of cancer. It is considered a potentially life-threatening disease with a really high deathrate. These figures are constantly increasing over the years, which induces a negative mindset amongst patients that have been diagnosed with lung cancer. In reality, doctors and several scientific studies have shown that lung cancer can be completely treated provided they are detected early [5]. Thus, early detection is the key to decreasing the fatality rates and also improving the treatment methodologies. Lung cancer prognosis that has been detected and intervened at the right time is the key to diminishing the fear surrounding this grave disease. There are a variety of methods that can be deployed in order to detect lung cancer. Some of these methods are Computed Tomography (CT), Magnetic Resonance Imaging (MRI), X-rays and sputum cytology [6]. These methods are considerably expensive and are not accessible to all. Many of the real time occurrences of lung cancer is brought upon by smoking, both active and passive. Active smokers are the ones who directly abuse the substance whereas passive smokers are those that ingest the fumes let out by active smokers. The efficacy of treatment for such a dreaded disease is dependent on the type, classification and stage of cancer i.e., the extent to which the disease has spread into the patient's organs and the comprehensive health status of the affected patient [7]. The investigation of exhaled breath samples is becoming a growing area of study for considering the respiratory framework and function. The samples of exhaled breath contain about 350 chemical components that is carbon monoxide, nitric oxide, and several other volatile organic compounds. Exhaled breath is known to be synthesized in a variety of malignant and non-malignant respiratory disorders such as asthma, bronchiectasis, cystic fibrosis, chronic obstructive pulmonary disease (COPD) and pulmonary fibrosis [8]. Estimation of the Volatile Organic Compounds (VOCs) present in the gaseous phase of the exhaled breath has become a widely expanding methodology of research scripting in detecting malignant growth in the lungs. Some of these methodologies are in early clinical development stages. Other VOC estimation techniques include Solid Phase Microextraction (SPME) which may be a virtual array of surface acoustic wave (SAW) gas sensors. Solid phase microextraction is an advanced, solvent-free sample prep technology that is fast, cheap, and multipurpose. [9]. In this day and age, CT scans of the lung area are known to be constructive in the diagnosis of lung cancer. These scans can only be read by well qualified doctors, who can gauge the invasion of the lung cancer nodules effectively. The Objective of this learning is to assist the diagnosis of lung cancer using image processing

techniques. Themeasurable results obtained from the proposed system can help doctors and physicians make a knowledgeable decision in real time, consuming minimal resources. Thus, this technique effectively helps in detecting cancers at their early stages [10].

II. LITERATURE SURVEY

In [1] Heewon Chung et al, introduced an exclusive lung contour procedure competent of recognizing juxta-pleural nodules. The procedure is implemented on the idea of the CV model then further adapting a Bayesian approach to detect juxta-pleural nodules and false positives are eliminated through concave points recognition and circle/ellipse Hough transform.

In [2] S.A. Abdelrahman et al, proposea novel AGLI method of representation which is subsequently used on the obtained datasets to map image space to the respective symbolic sequences. Aefficientprocedure for genetic mutations classification is built in conjunction with 2DPCA. By combining the 2DPCA with the AGLI demonstration of mutated lung cancer genes, the undetermined gene is mapped to one of the 23 sequences.

In [3] Azzawi, Hasseeb et al, propose a novel GEP-based classifier for the classification and prediction of lung cancer from microarray data. This proposed method achieved a greater correctness in classifying lung cancer on the commonly used datasets, as compared to similar descriptive machine-learning based classifiers. The final results indicated that the GEP proposition improved the prediction accuracy of lung cancer.

In [4] RussulAlanni et al proposed an efficient gene selection algorithm. It can select relevant genes and significantly subtle to the samples in the few genes and less cost time by the algorithm achieved much high prediction accuracy on several public microarray data, it in turn proves the efficiency and effectiveness of the systems.

III. EXISTING SYSTEM

The automatic segmentation and fusing operation accommodate the multi modal input images segregated into regions which are later fused together on the basis of the fusion rule. The system encompasses three major constituents: multichannel super pixel level feature extraction and fusion, kernel sparse representation, and segmentation. A collection of pixels with identical perceptions is referred to as the super pixel. These super pixels strive as the predominant processing unit for efficiency and compact representation. The constituent processes of kernel sparse representation, such as sparse coding and dictionary learning, are both administered in a high-dimensional feature space with the aid of the kernel trick.

Disadvantages
- With each level of decomposition, the artefacts of the source image are conserved as the fusion system intentionally passes this information through the various levels.
- It is difficult to determine whether narrowing of a spinal canal.

IV. PROPOSED SYSTEM

In the proposed system, both qualitative and quantitative methodologies have been employed. Morphology is the qualitative operation and feature extraction is the quantitative operation that is being used in the proposed system. These operations are then combined with image quality assessment techniques to segregate and detect the cancer of the lung in the digitally scanned input data. The segmentation process begins by first employing the masking methodology and filtering out the acquired image. To this extracted filtered image, morphological operations are then applied, which results in a boundary that is drawn around the studied tumor affected region. The portion encompassed by the boundary is then exported to be analyzed separately. Filtering and contrast enhancement processes are applied to the exported image followed by an image quality assessment suite (Mean Square Error, Peak Signal to Noise Ratio etc.). Conclusively, accuracy estimation methodologies are invoked which is preceded by the neural network and segmentation process. Ultimately, the obtained values of feature extraction and image quality assessment are plotted which provides an overview of the extent of invasion of the tumor cells

Advantages
- Efficiently predicative
- Sensitivity of this proposed system to noise, blurring effects and miss registration.
- This methodology constructs an expert knowledge base about the challenge by making use of the efficient trick.
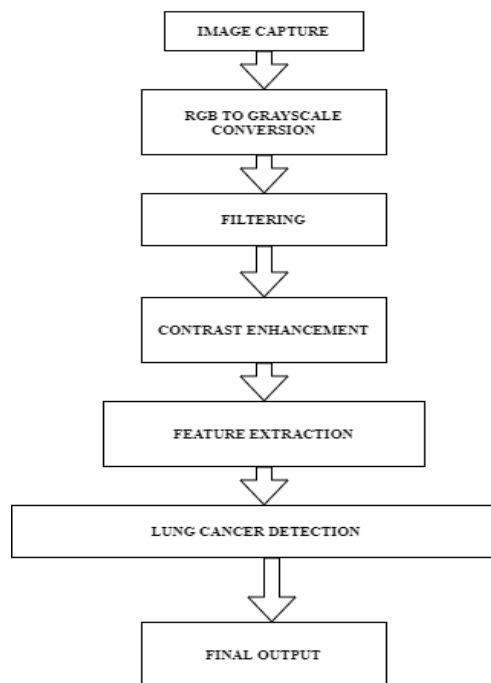
- A considerable work of theory proposes that it should be acceptable to rely on the test error rate as it is actual and accurate.
- Easy to implement and used
- Low cost

BLOCK DIAGRAM
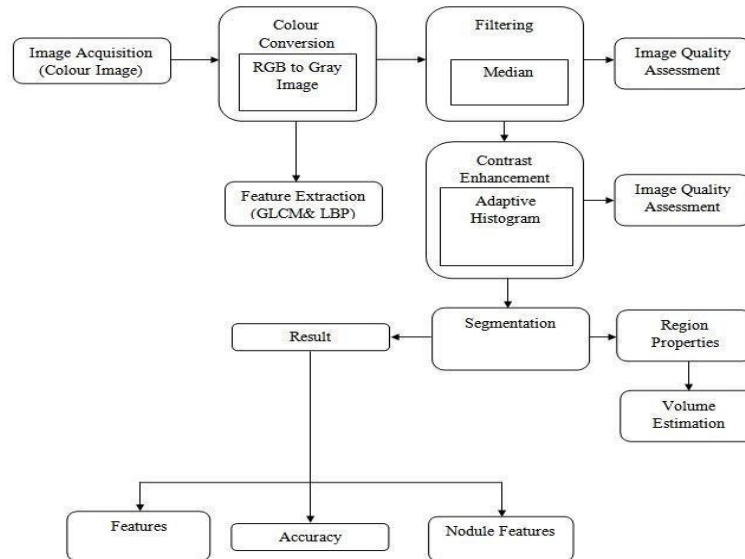
Figure 1. Block diagram – MATLAB unit

MODULE DESCRIPTION

Figure 2. Flowchart – Module Description

A. Input Image

In the RGB colour model, the three lights, red, green and blue respectively, amalgamate in several ways to proliferate a wide range of combined colours. The name of the model is obtained from the first letter of the three supplementary colours. Different devices or gadgets presume or procreate RGB in unprecedented ways as the colours and their reaction to the characteristic red, green and blue colours transpose in multiple ways from one client to another or at times even in the same device at different time intervals as the source of the colours change. Therefore, red green and blue values define different spectrum of colours across different devices. Thus, these characteristics render RGB as a dependent colour model. A colour is manifested by superimposing one each of the red, green and blue light waves. The resultant colour is a product of the three beams, each of them having varying intensities.

Figure 3. Input Image

B. Gray Image

In the digital world, an image conversion type in which the value of each pixel is a single sample, is referred to as a grayscale image. A grayscale image bear only the intensity values of the constituent colours. Since they are a combination of varying shades of grey, from the lowest intensity of black to the strongest intensity of white, these images are also known as black and white images.


Figure 4. Gray Image

C. Filtering

The process of removing unnecessary noise from a signal is known as filtering. Filtering involves the complete or partial suppression of unwanted aspects of the signal that cause distortion or noise in the original intended signal. Such a noise removal protocol is an ideal processing methodology that is carried out in order to improve the accuracy of edge detection. Median filter is employed in the analysis of digital signals to remove noise from the input 2D images. Median filter is used as it carefully distinguishes the edges from the noise. The fundamental intention of the filter is the step by step analysis of every component of the incoming signal, and replace them with the median of the neighbouring entries.

Figure 5. Median Filter

D. Contrast Enhancement

An image processing technique that involves modifying the contrast of the image using the histogram of the image, is referred to as contrast enhancement. This process is carried out by distributing the most intense colour values to a clear blur. Adaptive histogram equalisation (AHE) is an image processing technique that is used to modify the contrast of digital images. The distinguishing factor of adaptive histogram equalisation is the way in which it spreads the lighter values of the colours in the image by drawing parallels with several histograms that it computes.
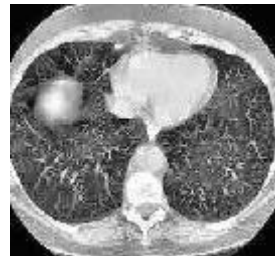
**Figure 6.**Adaptive Histogram Equalization

Feature extraction methods in image processing and, recognition algorithms in machine learning, all sprout from an initial measured dataset. From this initial data, feature extracted values are created that are aimed to be enlightening and educative. Feature extraction is based on the principle of dimensionality depleted quantitative analysis.

Figure 7.GLCM Parameters

IV RESULTS AND DISCUSSION

I selected Suitable classification algorithm from data mining and it's implemented in MATLAB. This algorithm is tested with carcinoma data as input. Examination of the algorithms, the CT images are taken from a NIH/NCI Lung Image Database Consortium (LIDC) dataset. More than 1000 lung images input to this system. From the output results, the analysis rules are formed from those images and these rules fixed to the classifier for the training progression. Subsequently a lung image is proceeded to the proposed system. Then this system is executing its processing and finally it isdetected the input image has cancer or not. This planned CAD system is accomplished of detecting lung nodules with diameter $\geq$ 1.5 mm, which recommends that the system is capable of detecting lung nodules once they're in their early stages. Thus, facilitating early diagnosis would improve the patient's survival rate. This system further categorises the cancer as either small lung or non-small lung cancer.
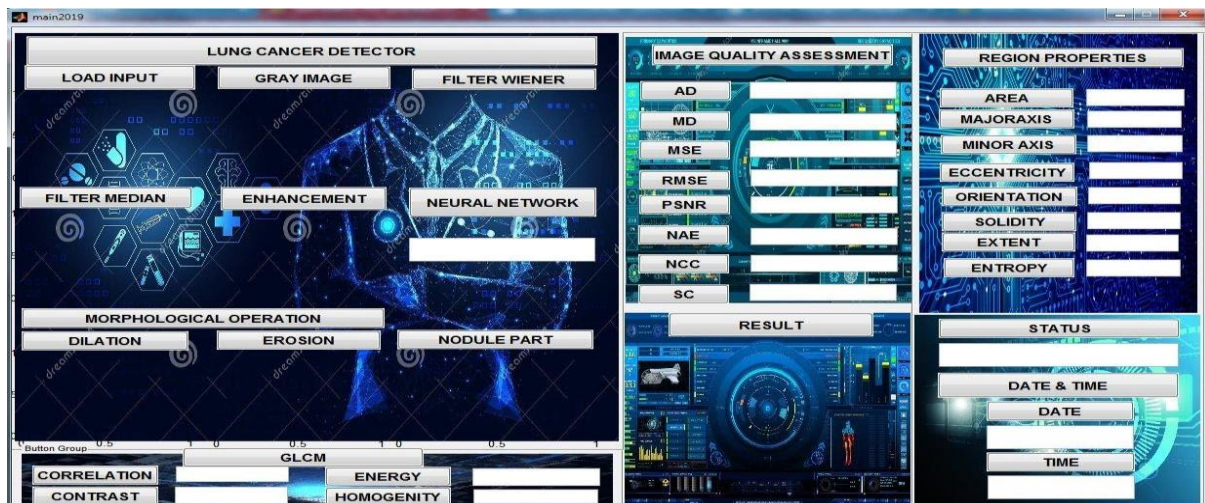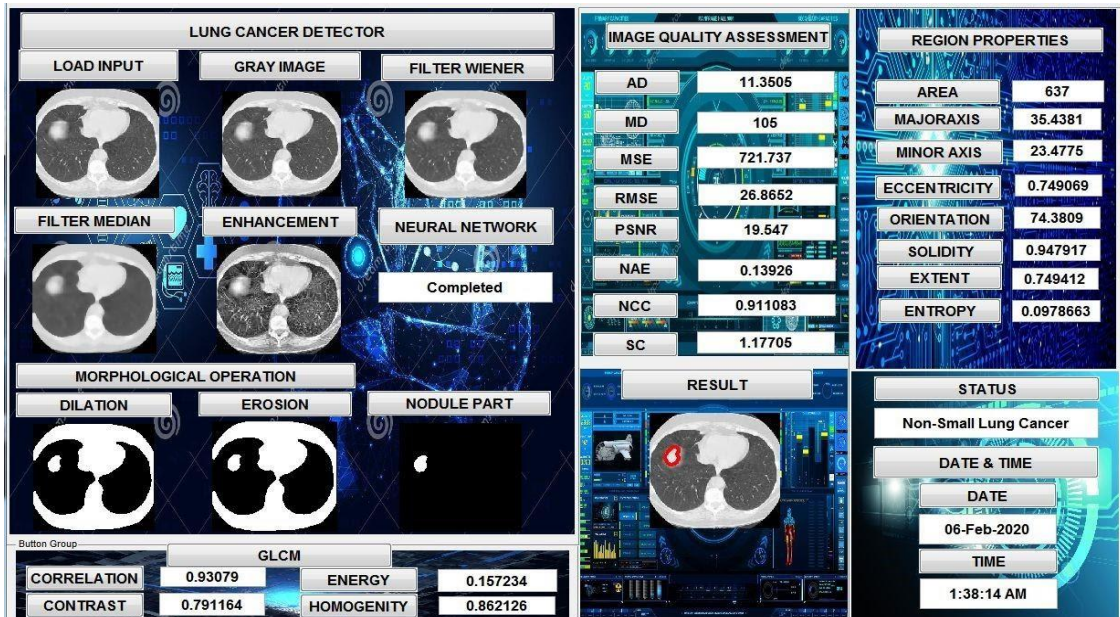

Figure 8. Initial GUI

Figure 9. Final Output

VIII. CONCLUSION

Now a days the area of disease Diagnosis is an uninterruptedly developing and very vigorous field for research and development. The objective of this research is, to predict the status of the patient for initial stage detection of lung cancer. In this work, the Diagnosis of Lung cancer and classification is made by means of Neural Network and Morphological Operation techniques, the segmentation and detection processes are carried out by means of intensity computation, GLCM, image quality assessment and features extraction. Using Median Filter is an effective way of detecting the Lung Cancer at an early stage by enhancing the image by noise reduction. This system is trained by using previously taken data set. It is effectively working and accuracy of the prediction is more. It's found out the lung cancer in the early stages. This tool can be fatherly used and developed by the biomedical department.

REFERENCES

[1] Chung, Heewon, et al. "Automatic lung segmentation with juxta-pleural nodule identification using active contour model and Bayesian approach." *IEEE journal of translational engineering in health and medicine* 6 (2018): 1-13.

[2] Abdelrahman, S. A., and M. M. Abdelwahab. "Accumulated grey-level image representation for classification of lung cancer genetic mutations employing 2D principle component analysis." *Electronics Letters* 54.4 (2018): 194-196.

[3] Azzawi, Hasseeb, et al. "Lung cancer prediction from microarray data by gene expression programming." *IET systems biology* 10.5 (2016): 168-178.

[4] Alanni, Russul, et al. "Deep gene selection method to select genes from microarray datasets for cancer classification." *BMC bioinformatics* 20.1 (2019): 608.

[5] Mrs. V.P.Krishnammal, Firosha S Fathima, Jills P Mathew, M Preethi kalyani, V Shiva Shankari, "Integrated Approach for Bone Tumor Detection From Mri Scan Imagery," *International Journal of Research in Advent Technology (IJRAT)*, VOLUME-7 ISSUE-4, APRIL 2019, pp. 324-325.

[6] Chen, Xuan, et al. "Reworking multilabel brain tumor segmentation: an automated framework using structured kernel sparse representation." *IEEE Systems, Man, and Cybernetics Magazine* 3.2 (2017): 18-22.

[7] H. Jiang, H. Ma, W. Qian, M. Gao and Y. Li, "An Automatic Detection System of Lung Nodule Based on Multigroup Patch-Based Deep Learning Network," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, pp. 1227-1237, July 2018, doi: 10.1109/JBHI.2017.2725903.

[8] Jitender, Mr, and KarambirSheoran. "LUNG CANCER DETECTION USING DIGITAL IMAGE PROCESSING AND ARTIFICIAL NEURAL NETWORKS." *LUNG CANCER* 6.10 (2019).

[9] Mistry, Krupali D., and Bijal J. Talati. "Integrated approach for bone tumor detection from mri scan imagery." *2016 International Conference on Signal and Information Processing (IConSIP)*. IEEE, 2016.

[10] Panpaliya, Neha, et al. "A survey on early detection and prediction of lung cancer." *International Journal of Computer Science and Mobile Computing* 4.1 (2015): 175-184.