"Study On Biomedical Image Processing For Detection of Lung Cancer Tumor"

[1]Shradha C. Fule, [2]Ruhina Quazi
[1]Student Of M-Tech , [2]Assistant Professor,
[1]Electronics And Communication Engineering,
[1]Anjuman College Of Engineering, Nagpur, India

Abstract: For early detection of diseases image processing techniques are most commonly used. A study shows that it can improve accuracy, sensitivity and specificity for early detection of lung cancer through image processing techniques. Collected CT (Computed Tomography) scan image of the lungs are pre-processed and the Region of Interest (ROI) is segmented. The proposed system consist of pre-processing, segmentation, features extraction and image analysis from segmented images. Features such as entropy, co-relation, energy, variance and homogeneity are extracted using a GLCM (Gray level Co-occurrence Matrix) with classification effected by means of an SVM (Support Vector Machine). This Classification identifies whether these CT images are normal or abnormal. The Lung Image Database Consortium dataset (LIDC) has been used for collection of normal and abnormal images of lung for training and testing purpose for this study.

Keywords— Pre-processing; Segmentation ; Thresholding ; Clustering; Features extraction; GLCM (Gray level Co-occurrence Matrix); Normal; Abnormal.

INTRODUCTION:

Lung cancer is one of the most common cause of cancer in both men and women. Due to this large number of people die every year. This disease has different stages whereby it starts from the small tissue and spreads throughout the different areas of the lungs by a process called metastasis. Cancer tumor is the uncontrolled growth of unwanted cells in the lungs [1]. It was found that around 12,203 individuals had lung cancer in 2016, 7130 males and 5073 females; deaths from lung cancer in 2016 were 8839. In Biomedical engineering image processing is the latest emerging tool in medical research used for the early detection of cancers. In Biomedical Engineering image processing techniques can be used in the medical field to diagnosis disease at an early stage. It uses biomedical images such as X-rays, Computed technology and MRIs [2]. The main purpose of image processing in the medical field is to diagnosis the cancer at an early stage, increasing survival rates [3]. This states that prediction of tissues can reduce mortality rate .The time factor plays a critical role for growth of tumors of the brain, lungs and breast. Biomedical Image processing can help to detect these cancerous cells in the early phases of the diseases facilitating, an early treatment process. This image processing technique consists of four basic stages, preprocessing, segmentation, feature extraction and classification. For early detection of these cancerous cells a system is proposed which consist of various stages. Here the CT scan images are pre-processed, segmented and their features are extracted and classified into normal and abnormal.

METHEDOLOGY:

The basic block structure of the proposed system is as shown in figure 1.Here the initial CT scan images are pre-processed. Pre-processing is an initial tool used to improve quality of image, making it suitable for further application. Here noise is removed using enhancement techniques, then improved images are segmented into clusters to identify the affected regions. The segmented images gives the clear identification of region scanned (normal or abnormal)..The segmented ROI is analyzed using various features.

Following features are being calculated:entropy,contrast,correlation,energy,homogeneity,variance,mean,standard deviation, skewness, kurtosis, FCM(fifth central moment),SCM(sixth central moment).

The range of these features classify the data as normal or abnormal.

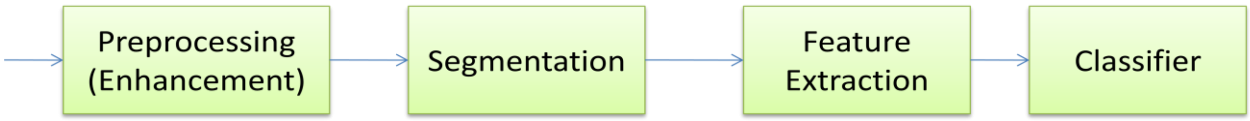

fig. 1 Basic Block Diagram of proposed model

A. PRE-PROCESSING STAGE

In this image Pre-processing stage includes enhancement technique which aims is to improve the interpretability,quality or perception information in images for human visual system or to provide better input for other automated image processing techniques Therefore, normally images undergo several pre-processing steps[4]. Image pre-processing smoothen and removes noise for further processing steps.

i). IMAGE ENHANCEMENT

Image enhancement is done to remove unambiguities or undesired data from images. The images are filter using various filters and depending upon their performance parameters the best filter is selected for pre-processing.

B. IMAGE SEGMENTATION:

The next stage is segmentation which segments out the ROI from pre-processed images. Segmentation can be done using Thresholding and Clustering process[5].Image segmentation is an important process for most image analysis tasks. There are various types of existing techniques for image description and recognition which depend highly on image segmentation results. Thresholding is one of the widely used tools for image segmentation. In this process the segmented image obtained from Thresholding has the following benefit. It has fast processing speed and hence it require lesser space for storage, compared with gray level image which usually contains 256 levels.

Image Segmentation is also facilitated by EK (Enhanced) mean Clustering segmentation where first the enhanced image is clustered through k-mean clustering where after the resulting image is again processed through EK-mean algorithm[5]. In image segmentation the K-mean algorithm is an unsupervised clustering algorithm which divides the input data point into multiple classes which inherit distance from each other. This k-mean algorithm assumes that the data form a vector space.
This k-mean clustering algorithm divides a collection of objects into k groups.
This algorithm has following steps:-
1.It Calculates the mean of each cluster. Calculate the distance of each point in each cluster by calculating its distance from the cluster mean and assign each point to its nearest cluster.

C. FEATURES EXTRACTION:

Features Extraction stage is very important step in image processing there are various algorithms to detect and isolate various required portions or tumor shapes (features) of an image. Feature extraction in biomedical image processing is an essential stage that represents the final results which determines the normality or abnormality of an image[6]. The various features act as the basis for the classification process.
The features extracted are as follow:

a. **Contrast** : Contrast in image is the differences of color and light between the parts of the image and it is used to measure the intensity between pixel and its neighbour pixel.

$$Contrast = \sum_{i}^{N_g} \sum_{j}^{N_g} (i - j)^2 \, p(i, j)$$

b. **Correlation :** Correlation is the process of establishing a relationship or connection between two or more regions. It is used to measure the gray scale linear dependence between pixel for specific position.

$$Correlation = \frac{\sum\sum p(i, j) - \mu_{row}\mu_{col}}{\mu_{row}\mu_{col}}$$

c. **Energy** :Energy in image is used for motivating that particular object to detect and segment and used for minimization.

Energy= $\Sigma_{i,j}\, p(i,j)_2$

d. **Entropy** :Entropy is used for measurement of randomness. Entropy measure reduction in the informations from transmitting signals.

$$Entropy = -\sum (p.*\log 2\ (p))$$

Where p contains the histogram counts returned from imhist.

e. **Mean:** It calculates the mean of all the values in ROI of image. where *p*(*i*, *j*) is the intensity value of the pixel at the point (*i*,*j*). The image is of M by N size.

$$Mean,\ \mu = \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{p(i,j)}{MN}$$

f. **Standard Deviation:** It calculates the average distance between pixels values and the mean. Low value of standard deviation indicates that there is less deviation of the pixels from the mean where the large value indicates the high variation.

$$\sigma = \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} \frac{(p(i,j)-\mu)^2}{MN}}$$

g. **Skewness:** It gives the asymmetric value of the probability distribution about its mean. The skewness of an image can be positive, negative.

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{[p(i,j)-\mu]^3}{(MN)\ \sigma^3}$$

h. **Kurtosis:-**It measures the flatness of the distribution with respect to the normal distribution.

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{[p(i,j)-\mu]^4}{(MN)\ \sigma^4}$$

i. **Fifth and sixth central moment(FCM)** :-These give the deviation from its mean.

Fifth central moment,

$$= \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{[p(i,j)-\mu]^5}{(MN)\ \sigma^5}$$

Sixth central moment

$$= \sum_{i=1}^{M} \sum_{i=1}^{N} \frac{[p(i,j)-\mu]^6}{(MN)\ \sigma^6}$$

j. **Homogeneity:** It gives the distribution value of the closeness of elements of the GLCM to the diagonal of GLCM.It gives the value between the range of 0 and 1.

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|}$$

These 10 features describe the characteristics of the segmented images.

D. CLASSIFICATION:

Classification is also an important step to determine the normality and abnormality of an image and it is carried out using an SVM (Support vector machine), classifying whether the image is normal or abnormal. Boyle in 2011 found the SVM as a classifier defined by a separating hyper plane - a machine learning algorithm[7]. For this support vector machine algorithm, we plot data items in 'n' dimensional space where 'n' is the number of features with the value of the feature being equal to the value of the coordinate and then, we perform classification by finding the hyper plane [7]. SVMs is a supervised learning models which are used to analyze data for classification. Support vector machine use optimum linear separating hyper planes which are further used for classification and regression. Further, according to Kung [8], an optimum hyper plane is used to separate two sets of data in feature space and the optimum hyper plane is produced by distinguishing margins between the two sets [8]. This means, the hyper plane will depend on border training patterns called support vectors. Here, the linear kernel SVM is used to classify the image into normal or cancerous images[14].

CONCLUSION:

Lung cancer is one of the most dangerous and widespread cancer in the universe. Image processing techniques can be used to detect early stage lung cancer in CT scan images . For results we divided our work into four stages: Image Enhancement stage, Image Segmentation Stage and Features Extraction Stage and Classification Stage. Lung Nodule Detection in CT Scans is an active area of research and there are many enhancements that can be included to make more efficient. TheCT scan image is pre-processed followed by segmentation of the ROI of the lung and features are extracted using a GLCM. The results are fed into an SVM classifier to determine if the lung image is cancerous or not. The SVM classifier is evaluated based on an LIDC dataset.

REFERENCES:

[1] J. M. Debois, "proposed a system for the clinics and anatomy of cancer like metastatic and Boston," Kluwer Academic Publishers, 2002.

[2] J.M. Fitzpatrick, & M. Sonka, "Medical imaging 2003: image processing. USA: SPIE Society of Photo-Optical Instrumentation Engineers," 2003.

[3] C. M. Haskell, & J.S. Berek, " Cancer treatment. Philadelphia: W.B. Saunders, 2001.

[4] B.V. Ginneken, B. M. Romeny and M. A. Viergever, "Computer-aided diagnosis in chest radiography: a survey",
IEEE, transactions on medical imaging, vol. 20, no. 12 (2001).

[5] Preeti Panwar, Girdhar Gopal, Rakesh Kumar,proposed a system for "Image Segmentation using K-means clustering and Thresholding", for a IRJET, Volume: 03 | May-2016.

[6] Suzuki K., et al., proposed a system for "False-positive Reduction in Computer-aided Diagnostic Scheme for Detecting Nodules in a patient Chest Radiographs by Means of Massive Training Artificial Neural Network", for a Academic Radiology, 12, No 2, February 2005, pp. 191-201.

[7] B. H. Boyle, proposed a system for "Support vector machines : data analysis, machine learning, and applications," Retrieved from http://public.eblib.com/choice/publicfullrecord.aspx?p=3021500, 2011.

[8] S. Y. Kung, "Kernel methods and machine learning, 2014.

[9] S. Akram, M.Y. Javed, & A. Hussain, " Automated Thresholding of Lung CT Scan for Artificial Neural Network Based Classification of Nodules," RetrievedFrom http://ieeexplore.ieee.org.ezproxy.csu.edu.au/stamp/stamp.jsp?arnumber=7166616, 2015.

[10] G. Vijaya,A.Suhasini,& D.D. Selvi, "Lungs Tumor Detection Using Pixel Value Matching (PVM) Method," AENSI, 2015.

[11] P.B. Sagamihara, & S. Govindaraju, "Lung Tumour Detection and Classification using EK-Mean Clustering," Retrieved from http://ieeexplore.ieee.org.ezproxy.csu.edu.au/stamp/stamp.jsp?arnumber=7566533, 2016.

[12] D. Harini, & D. Bhaskari, "Image retrieval system based on feature extraction and relevance feedback," ACM, 2 Penn Plaza, Suite 701, New York,NY 10121-0701, 2012.

[13] Deep Prakash Kaucha, & S. Sreedharan, proposed a system for "Early Detection Of Lung Cancer Using SVM Classifier in Biomedical Image Processing," .

[14] K.Gopi &Dr.J.Salvakumar, "Lung Tumor Area Recognition and Classification using K-mean Clustering and SVM," .

[15] S. panda, "Color Image Segmentation Using K-means Clustering and Thresholding Technique," *IJESC,* march 2015.