ANFIS Classifier Based Lung Tumor and Lymph Node Differentiation using CT

M.Kavitha[1] T.Raja[2]
[1]PG Scholar [2]Assistant Professor
[1, 2]Centre for Advance Research, Muthayammal Engineering College, Rasipuram.

Abstract--- Analysis of primary lung tumors and disease is important for lung cancer staging and an automated system that can detect both types of abnormalities will be helpful for clinical routine. In this project, we present a new method to automatically detect both tumors and abnormal nodes simultaneously using Computerized Tomography (CT) thoracic images. We perform the detection of abnormalities, and then differentiate between tumors and lymph nodes using ANFIS classifier. Then lung tumors are classified as benign and malignant using this classifier. Moreover compared to several state-of-the-art methods (LapRLS, mcSVM, ESRC, mSRC, CBIR), the proposed ANFIS classifier can achieve significant improvement (mean accuracy of 93%, precision of 88%, recall of 94%, etc).

Keywords: Computerized Tomography (CT), Local Binary Pattern (LBP), Gray Level features, Adaptive Neural Fuzzy Inference System (ANFIS) classifier.

I. INTRODUCTION

According to the World Health Organization (WHO), lung cancer is the most common cause of cancer- related death worldwide. Conventionally, lung cancer can be categorized into four types: squamous carcinoma (SC), adenocarcinoma (AC), small cell cancer (SCC) and nuclear atypia (NA). The diagnosis of lung cancer is made by the pathologist. However, the pathologist's inexperience in clinical practice can cause misdiagnosis. The diagnosing process is time consuming and sometimes tedious when a pathologist is asked to analyze a huge volume of sample images from patients. The misdiagnosis occurs due to the factors as fatigue, for an experienced pathologist also. The accuracy not only depends on the pathologist's knowledge and clinical experience but also his/her physical and psychological conditions. Several computer aided methods for diagnosing cancer imaging as MRI, CT, and PET etc.

To analyze the images which are captured by Computer Tomography thoracic images is one of the most popular method and efficient way to diagnose the lung cancer. To classify the different types of lung cancerous image is still a significant and challenging problem due to the different cancerous images is similar to one another and hence difficult to classify.

II. RELATED WORK

For diagnosing a lung cancer images, many efforts have been contributed in recent years. Zhou et al.[10] developed a lung cancer classification system based on two-level neural network ensemble. In the system, the first-level ensemble determines whether a testing cell is normal or cancerous, and the second level ensemble classifies the type of lung cancer for the suspected cancerous cells to determine in the first level. Zhu et al.[14] proposed an image- level approach: multi-class multi-instance AdaBoost (MCMI-AdaBoost), which predicts the label of an image by incorporating the multi-instance distance measurement (Hausdroff distance) under the AdaBoost framework. However, the similarity measurement ignores local information of cells. Shi et al. in recent times introduced a transconductive cost-sensitive learning method for lung needle biopsy image classification, of which the goal is to achieve the best possible results with only a small number of labeled images. J. Kuhnigk, V. Dicken, L. Bornemann[7],Morphological segmentation and partial volume analysis for volumetric of solid pulmonary lesions in thoracic CT scans, morphological segmentation is used and the Retrieval accuracy rate is 67% only and does not support time cost. S. G. Armato, M. L. Giger and H. MacMahon[6], Automated extraction of lymph nodes from 3-D abdominal CT images using 3-D minimum directional difference filter, Directional difference filter is used and results in low sensitivity and low accuracy. W. Wever, S. Stroobants, J. Coolen, and J. Verschakelen[8], A framework for automated tumor detection in thoracic PET images using texture-based feature which results in low accuracy and high detection time cost.

Unfortunately, the results of classification for different types of lung cancer are still unreliable, and there is still space for us to improve the classification performance.

In addition, lung needle biopsy image classification is single-modal learning methods, which fail to make full use of the disagreement information among different modalities. It is noteworthy that, except for the needle biopsy specimens based lung cancer diagnosis, many works that focus on cell/nodule/image classification in medical image analysis are highly related to our work.

For learning with multi-modal data, also referred to multi-view learning and ensemble learning in machine learning community, many algorithms have been developed recently. Also, multi-modal based methods are promising in the field of medical image analysis since multi-modal information is naturally available in the data acquisition procedures of various clinical tasks, such as Alzheimer's disease diagnosis, prostate cancer prediction, and survival prediction for lung cancer.

The rest of this paper is organized as follows. We first present the framework of our method, block diagram representation and finally conclude the paper.

III. FRAME WORK

Generally, our proposed method to differentiate the normal or abnormal condition of the CT- thoracic images. The input images are captured by the Computed Tomography. Then the

CT thoracic or lung images are preprocessed and then segmenting the images. Then using feature extraction methods we get the desired features of the images. Then using an ANFIS classifier, detects and select the best training features from the extracted trained feature set. Then this classifier differentiates the lung image as the benign and malignant.

IV. METHODOLOGY

Preprocessing of CT lung image is the first step in our proposed technique. Preprocessing of lung image is to reduce the noise and enhance the image for further processing.

A. Step 1: Preprocessing

In preprocessing step first the lung CT-thoracic image is converted to gray scale image and then applying median filter on lung CT image in order to remove the noise.

After enhancing the Lung CT image, the next step of our proposed technique is to segment the Lung tumor region from Lung CT image. Segmentation is done to separate the image in to two or more sub module regions. Segmenting an image also saves the processing time for further operations which has to be applied to the image. We use segmentation using a global threshold in order to segment the tumor region from Lung CT image.
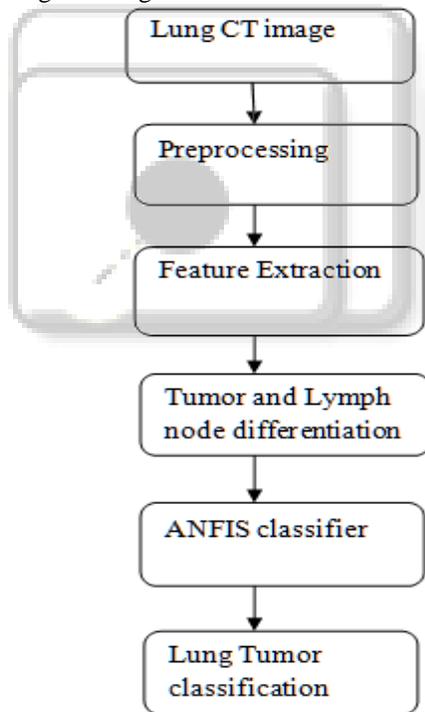

Fig. 1: Stages of Lung cancer detection

B. Step 2: Thresholding operation

Select a global threshold value for the whole CT Lung image. Apply the threshold value to the preprocessed image to convert the image to binary and the threshold image is obtained.

Morphological close operation is applied on the threshold image to fill in holes and small gaps in the image. Reserve the block whose area is the biggest and set the others to zero using 8- connected neighbors. The binary

lung mask is obtained using the above step. Extract the lung boundary by setting a pixel to 0 if its 4- connected neighbors are all 1's, thus leaving only boundary pixels. Multiply the original Lung CT image with the lung masked image to obtain the final segmented lung region with gray level values as those of original image.

C. Step 3: Feature Extraction

Local Binary Pattern:

The local binary pattern operator works in a 3×3 pixel block of an image. The pixels in this block are threshold by its center pixel value, multiplied by powers of two and then summed to obtain a label for the center pixel. As the neighborhood consists of 8 pixels, a total of 28 =256 different labels can be obtained depending on the relative gray values of the center and the pixels in the neighborhood.

$$LBP_{P,R} = \sum_{p=1}^{P} 2^{(p-1)} \times f_1(g_p - g_c)$$

$$f_1(x) = \begin{cases} 1, & x \geq 0 \\ 0, & else \end{cases}$$

Where g_c is the gray value of the centre pixel, g_p is the gray value of its neighbors, P is the number of neighbors and R is the radius of neighborhood. The LBP code may be interpreted as a kernel structure index.
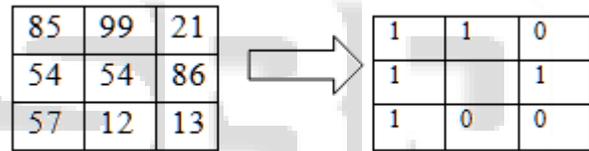
Binary: 11001011 Decimal: 203
Fig. 2:

Gray Level co-occurrence matrix:

The Co-occurrence features can be extracted from each Co-occurrence Matrix. The features are energy, contrast, homogeneity, correlation.

Property	Description	Formula
Contrast	Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image.	$\sum |i-j|^2 . p(i,j)$
Energy	Returns the sum of squared elements in the GLCM.	$\sum p(i,j)^2$
Homogeneity	Returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.	$\sum p(i,j)/[1+|i-j|]$
Correlation	Returns a measure of how correlated a pixel is to its neighbor over the whole image.	$\sum (i-\mu i)(j-\mu j).p(i,j)/[\sigma i.\sigma j]$

Table. 1:

These features are based on the differences between the gray-level in the candidate pixel and a statistical value representative of its surroundings. Let I_H be the preprocessed image.

Let (x,y) be the pixel in I_H.

Let S be the sub-image of I_H.

Let (s,t) be the pixel in S.

For the gray-level based feature extraction,9*9 center point window is used.

$F_1(x,y)= I_H(x,y) - \min\{I_H(s,t)\}$ $F_2(x,y)=$ $\max\{I_H(s,t)\} - I_H(x,y)$ $F_3(x,y)= I_H(x,y)$ $- \text{mean}\{I_H(s,t)\}$ $F_4(x,y)= \text{std}\{I_H(s,t)\}$ $F_5(x,y)= I_H(x,y)$

D. ANFIS classifier:

ANFIS depends on the sizes of the training set and testing set. In this work training and testing set were formed by 120 and 80 data. 120 data (40 normal and 80 abnormal cases) was used for training and the remaining 80 data was used for testing. ANFIS was implemented with MATLAB software. Features from biopsy, area of tumor region, entropy, and Homogeneity properties were given as input to ANFIS. Classification was carried out in two steps using ANFIS:

i) Training
ii) Testing

The Fuzzy Inference System generated creates an initial model for ANFIS training by first applying subtractive clustering on the data. If-then rules were generated from ANFIS and Membership functions are almost consistent for individual parameters before and after training in ANFIS.

The train FIS optimization methods are chosen as hybrid since it includes least square type along with back propagation gradient descent algorithm which trains the membership function parameters to emulate the training data. When the training error goal is achieved, the training process stops and average training error is noted.
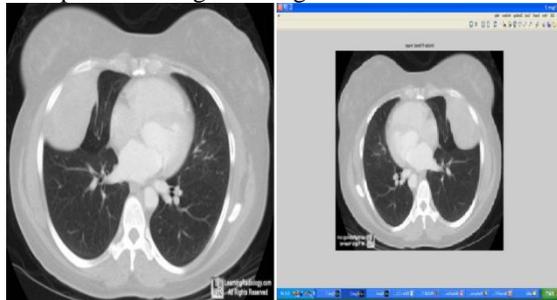

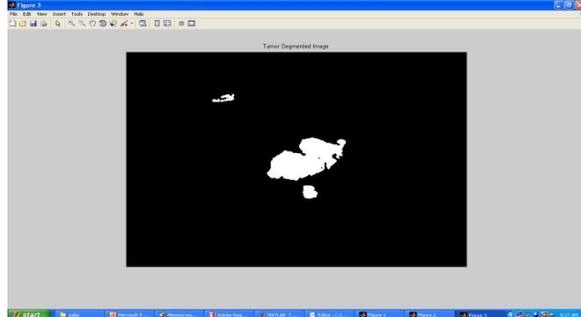Fig. 3: Lung CT image   Fig. 4: Median filtered image

Fig. 5: Tumor segmented region

V. RESULTS AND DISCUSSION

To further evaluate the performance of ANFIS, we also compare ANFIS against other related methods. They are multiclassSVM(mcSVM), MCMI-AdaBoost, LapRLS, and ensemble sparse classification (ESRC) are shown in the flow chart. The experiments are conducted with the help of real time lung images. The dataset consists of 1000's of lung images. Hence, the user develops an automatic system to detect an early stage of lung cancer.
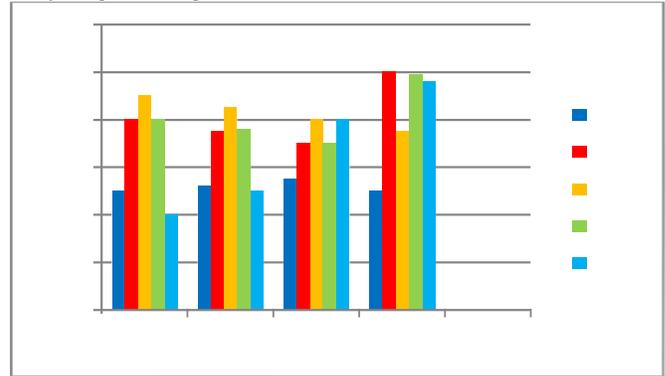

Fig. 6: performance comparison flowchart.

IMAGE	VALUE	CLASSIFICATION
Image 1	1	Malignant
Image 2	0	Benign
Image 3	0	Benign
Image 4	1	Malignant
Image 5	0	Benign
Image 6	1	Malignant
Image 7	1	Malignant

Table. 2:

VI. CONCLUSION

In this paper, we propose a novel method ANFIS classifier for classifying the lung tumor as benign or malignant. The presented ANFIS model combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. ANFIS provides an accuracy of 93%. The first step is preprocessing the image by median filter. The next step is to segment the tumor region by applying a thresholding value. Then the feature extraction is performed by GLCM, LBP. Finally they obtained feature classification is performed to detect the occurrence. We therefore have concluded that the proposed ANFIS model can be used in classifying the tumor by taking into consideration the misclassification rates.

REFERENCES

[1] K. Suzuki, I. Horiba, N. Sugie, and M. Nanki(Nov. 1998), 'Noise reduction of medical X- ray image sequences using a neural filter with spatiotemporal inputs', in Proc Int. Symp. Noise Reduction for Imag.and Comm. Systems, pp. 85–90

[2] S.G. Armato III, and H. MacMahon(2003), 'Automated lung segmentation and computer-aided diagnosis for thoracic CT scans,'International Congress Series pp. 977-982

[3] S. Edge, D. Byrd, C. Compton, A. Fritz, F. Greene,

and A. Trotti, Eds(2010), 'AJCC Cancer Staging Handbook', 7th ed. New York: Springer

[4] I. Kapetanovic, S. Rosenfeld and G. Izmirlian(2004),'Overview of commonly used bioinformatics methods and their applications', Ann N Y AcadSci1020,10–21

[5] M. Kakar and D. R. Olsen, 'Automatic segmentation and recognition of lungs and lesions from CT scans of thorax,' Comput. Med. Imag

[6] S. G. Armato, M. L. Giger and H. MacMahon(2001),'Automated detection of lung nodules in CT scans: Preliminary results', Med. Phys., Vol.28

[7] J. Kuhnigk, V. Dicken, L. Bornemann, A. Bakai, D. Wormanns, S. Krass, and H. Peitgen(Apr.2006), "Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans," IEEE Trans. Med. Imag., vol. 25, no. 4, pp. 417–434,

[8] W. Wever, S. Stroobants, J. Coolen, and J. Verschakelen(2009), 'Integrated PET/CT in the staging of nonsmall cell lung cancer: Technical aspects and clinical integration,'Eur. Respir. J., vol. 33, pp. 201–212

[9] J. Lafferty, A. McCallum, and F. Pereira(2001), 'Conditional random fields: Probabilistic models for segmenting and labeling sequence data,'inProc. ICML, , pp. 282–289

[10] A. Bardu, M. Suehling, X. Xu, D. Liu, S. Zhou, and D. Comaniciu(, 2010), 'Automatic detection and segmentation of axillary lymph nodes', in MICCAI 2010, LNCS, vol. 6361, pp. 28–36

[11] Y. Song, W. Cai, S. Eberl, M. Fulham, and D. Feng(, 2011), "Discriminative pathological context detection in thoracic images based on multi-level inference," in MICCAI 2011, LNCS, vol. 6893, pp. 185–192

[12] Y. Shi, Y. Gao, R. Wang, Y. Zhang, D. Wang. Transductive cost-sensitive lung cancer image classification. Applied Intelligence, vol.38, no.1, p:16- 28, 2013.

[13] L. Zhang, W.D. Zhou, P.C. Chang, J. Liu, Z. Yan, T. Wang, F.Z. Li. Kernel sparse representation- based classifier. IEEE Transactions on Signal Processing, vol.60(4), p:1684-1695, Apr, 2012.

[14] L. Zhu, B. Zhao, Y. Gao. Multi-class multi- instance learning approach for lung cancer cell classification based on bag feature selection. In FSKD,p:487-492, 2008.

[15] YBoykov, O. Veksler, and R. Zabih, 'Efficient approximate energy minimization via graph cuts',IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 12, pp. 1222–1239, Dec. 2001

[16] Rezaul.K.Begg, MarimuthuPalaniswami and Brendan Owen (2005),'Support Vector Machines for Automated Gait Classification', IEEE Transactions on Biomedical Engineering, Vol.52,No.5

[17] Laura Auria1 and Rouslan A. Moro, "Support Vector Machines (SVM) as a Technique for Solvency Analysis", Berlin August 2008

[18] Armato SG III, Altman MB, LaRivière(2003) PJ. Automated detection of lung nodules in CT scans: effect of image reconstruction algorithm. Med Phys; 30:461–472.

[19] Mochamad Yusuf Santoso Lung Cancer Early Detection and Diagnosis System Development Based Adaptive Neuro Fuzzy Inference System.

[20] Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction by Abdulhamit Subasi.