ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Novel Based Detection and Supervised Classification of Lung Nodules

Kranti Malharigol[1], Sandhya Bevoor[2]

PG Student, Department of DECS, Maratha Mandal Engineering College Belgaum, Karnataka, India[1]

Assistant Prof, Department of ECE, Maratha Mandal Engineering College Belgaum, Karnataka, India[2]

ABSTRACT: we propose a novel classification method for the four types of lung nodules, i.e., well-circumscribed, Vascularized, juxta-pleural, and pleural-tail, in low dose computed tomography scans. The automatic detection of lung nodules attached to other pulmonary structures is a useful yet challenging task in lung CAD systems. Our purpose is to develop an efficient Computer Aided Diagnosis (CAD) for detection of lung nodules from parenchyma region of lung. The proposed work was able to detect the lung nodule that falls in close proximity to the lung wall.

KEYWORDS: Classification, feature design, latent semantic Analysis, Patch division, SVM

I. INTRODUCTION

Cancer related deaths in humans due to the cause of lung cancer in worldwide [1]. Lung cancer represent 20% cases with nodules of Lung, therefore the recognition of Lung nodules is necessary for the viewing and analysis of cancer in the Lung [2]-[4]. These are the tiny masses present in the human Lung. They are deformed by adjoining anatomical structures for example vessels and adjacent pleura .Intraparenchymal Lung nodules are extra liable to be malignant than those attached with the adjoining structures and hence Lung nodules are separated into dissimilar types based on their virtual positions.

Fig1. Lung nodule image

In current, Diciotti is the famous approach for the classification and it classifies nodules into four types. First one is well circumscribed (W) with the nodule placed centrally without any link to vasculature .Second type is Vascularized (V) with the nodule placed centrally but nearly linked to adjoining vessels. Third type is Juxta pleural (J) with a huge part of the nodule connected to the surface of the pleural. Finally, Pleural tail closed to pleural surface attached by a thin tail. CT images of four types shown in below figure.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

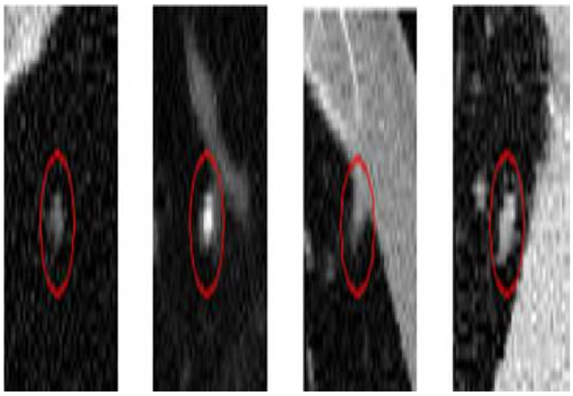


Fig2. Transaxial CT images

Computer Tomography is the mainly imaging modality to get information of anatomical structures. In recent medical custom, however analysis of CT images is difficult for radiologists because of many numbers of cases. The physical analysis can be error prone and the person who reads may fail to see the nodules and thus miss the cancer. Computer aided diagnosis (CAD) systems would be supportive for radiologists by contributing primary viewing for the classification. CADs support interpretation by anatomically computing quantitative measures and are able to viewing the tiny nodules recognized by CT scans.
The primary purpose of our project was to detect and analyze the small masses present in the human lung which are called as lung nodule's images are selected from lung cancer diagnostic research. Hence database contains CT images. For the purpose of training and development methods, database provides resources. CADs provide quantitative measures and are able to analyze numerous small nodules recognized by CT scans.

II. ARCHITECTURE

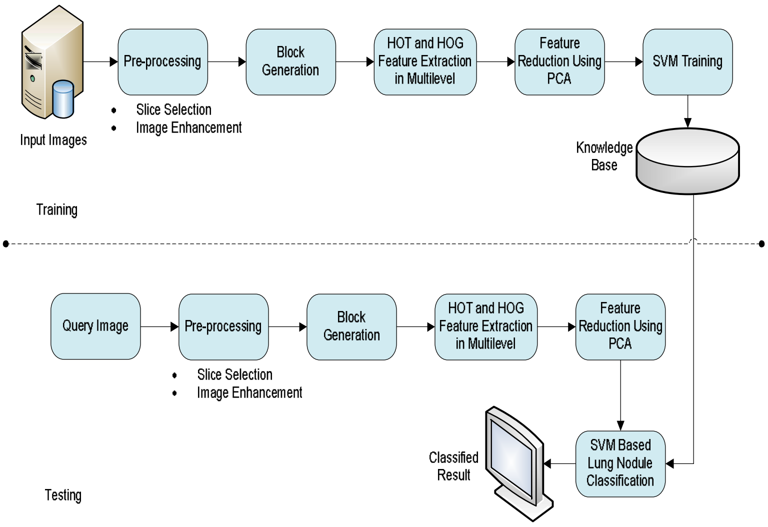


Fig3. Block diagram of proposed system.

This method is mainly divided into two parts training and testing. In training part we are creating a knowledge base where the training images are stored. Training images are nothing but the pre-processed. In testing part the input image or query image is pre-processed & then statistical features of query image are extracted. The extracted features are compared with features of the training images which are stored in the knowledge base & given query image is classified according to the features.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Training part:

1) Pre-processing: If we give input as images of lung it is pre-processed means it is divided into no of segments depends on the size. Images are of 256x256 pixels in size. In this image quality is influenced by variations, noise, non uniform intensity. Next is the slice selection where same types of images are selected (which is in same size). Then it is enhanced to same pixels or amplification done that is called image enhancement. Image intensity values are adjusted in the image processing is the standard solution.

2) Block generation: after getting the same type of images, converted into number of blocks that is put in 8*8, 16*16 or 32*32 format. So numbers of blocks are generated.

3) Hot and Hog feature extraction in multilevel: Now we can select single image from the blocks and converts into multiple slices (planes) that is zooming effect is done. So we get extra levels of 8*8, 16*16 and 32*32 matrix. Apply Hot and Hog features for each levels so we get extra features in multilevel.

4) Feature reduction using PCA: This process is used to reduce the features of Hot and Hog process that is it extracts only useful information.

5) SVM Training: The training is used for the detection result. This process enhances the result accurately. So for the judgement of unlike features, the consequence of training method is used. Support Vector Machine is used for many tasks. In this project SVM is used for Lung nodule comparison.

Testing part:

In this part query image is considered and same process is done. SVM training is applied to classify the lung nodule. So we get the classified result as types of lung nodule detection.

III. LITERATURE REVIEW

Computer aided diagnosis for lung classification attracts the interest of researchers. There was developed many works regarding the classification accurately. The important milestone in lung classification is the image processing. In the image processing, nodule features like size, geometric properties, shape and intensity are considered. Histogram is used to get the gradient information. The images are collected from database which is lung image database. It is easily available in the Google. Lung nodule classification method was proposed by Fan Zhang et.al. This method includes mainly three parts. Patch based division, feature set design and contextual latent semantic analysis for the relevant images. M.H.Hasna et.al proposed the lung nodule classification method for LDCT images. This method overcomes the issue of Lung nodule adjoining of adjacent anatomical structures. In a recent research high performance could be achieved by better feature set design and well advanced classifier to help the classification of overlapping adjacent nodule. In prior research contextual information could be used, but this method involves the complicated segmentation method. For example locality and shape of W and V nodules are similar so it is difficult to identify and distinguish them based on contextual information. Therefore Patch based division can be helpful for such a complication. Various methods can be proposed for lung nodule classification for research based data. In lung nodule detection many classification methods proposed the feature techniques for before the detection of nodules. Hence feature set design and classification is two different stages for the detection of lung nodules in CAD systems.

CAD system was trained and tested on images of lung made available by lung image database. In this development, training, evaluation, testing and comparison of nodules could be done through different approaches. Different CAD methods and algorithms for lung nodule detection were used. Support vector machine as nonparametric classifier are used for the classification.

IV. OUTLINE OF THE PROPOSED METHOD

In light of the above, this paper presents a novel image classification method for the four common types of lung nodules. We suggest that the major contributions of our work are as follows:

i) A patch-based image representation with multilevel concentric partition

ii) A feature set design for image patch description, and

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

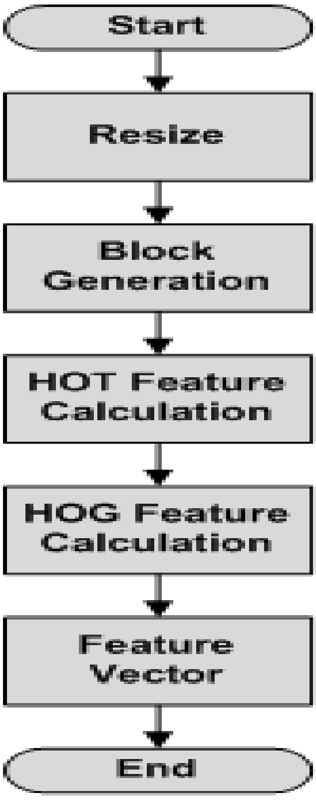
International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

iii) A contextual latent semantic analysis-based classifier to calculate the probabilistic estimations for each lung nodule image. The structure of this paper is as describe the three stages of the proposed method concentric level construction, feature extraction, and context analysis classification.

A.FEATURES OF FLOW CHART



i. PATCH BASED IMAGE REPRESENTATION

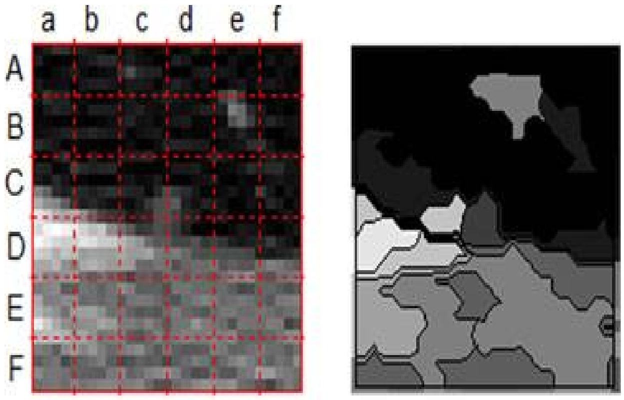
This method is based on a patch based representation of image. Patches with fixed size and shape are basis for this approach, such as partitioning image into the square or circular [7],[8] depends on the radial partitions[5],[6]. The rigid partitions inevitably combine distinct pixels collectively, as shown in figure 4.a. And preferably same patch pixels should share related information for example intensities as shown in figure 4.b

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016



(a) (b)

Fig 4(a) rigid partition 4(b) superpixel formation

SUPERPIXEL FORMATION

It is the process of isolating an image into several segments. It integrates the local spatial information and reduces the noise. This is absolutely robust to requirement of patch partitioning. In this mode seeking algorithm that is quick shift method can be used to segmenting an image into a set of super pixels creating tree of links to adjacent neighbour for which estimate of density increases. Lung nodules are in small sizes, when we apply direct quick shift method, we obtained a poor partition. To overcome this, we make use of quick shift in iterative manner with image amplification and also down sampling.

ii. FEATURE EXTRACTION

Distinction and Invariance are the two characteristics for image feature description. Descriptor desires to acquire the distinct features and be vigorous to capture numerous imaging conditions. Intensity, gradient and texture are the features of various nodules based on the visual study. Hence we designed the feature set of combination of multiorientation of HOG for gradient and HOT feature.

MULTI ORIENTATION HOG DESCRIPTOR FOR GRADIENT

Histogram of gradient is based on SIFT algorithm. For calculation purpose, image is divided into number of blocks. The blocks can extend beyond with other blocks. Every block comprises of four cells. For feature calculation, we considered the cell which is the basic unit. Each pixel contains pixel $I(x, y)$, the orientation as $\theta(x, y)$ and the magnitude can be represented as $m(x, y)$ of the gradient are calculated as

$$dx = \mathrm{I}(x+1, y) - \mathrm{I}(x-1, y) \quad (1)$$

$$dy = \mathrm{I}(x, y+1) - \mathrm{I}(x, y-1) \quad (2)$$

$$m(x, y) = \sqrt{dx^2 + dy^2} \quad (3)$$

$$\theta(x, y) = \tan^{-1}(dy / dx) \quad (4)$$

For each cell histogram is calculated and sum of magnitude of pixels as the length of each bin. And orientations are in the analogous interval. A block can be defined as a 36 dimensional vector because each block contains 2x2 cells.

A vector for each pixel in a sub window can be calculated by COV. The equation is given by

$$[x, y, |I_x|, |I_y|, \sqrt{I_x^2 + I_y^2}, |I_{xx}|, |I_{yy}|, \arctan\frac{|I_x|}{|I_y|}]^T \quad (5)$$

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

In above equation x and y are the pixel locations and remaining terms are intensity derivatives. The last term in the above equation defines the edge orientation. Covariance matrix contains set of vectors .These can be obtained from set of 8- dimensional vectors from each sub region.

HOT FEATURE

HOT feature has more advantage than HOG feature. It uses both gradient and intensity information. In this case pixel level is not extracted directly but extracted from middle level. 3x3 pixel region is the basic unit for middle level. HOT feature extraction is shown above. Higher detection rate can be obtained from HOT feature. It has characteristic of illumination invariant and is more discriminative.

iii. CONTEXT ANALYSIS CLASSIFICATION

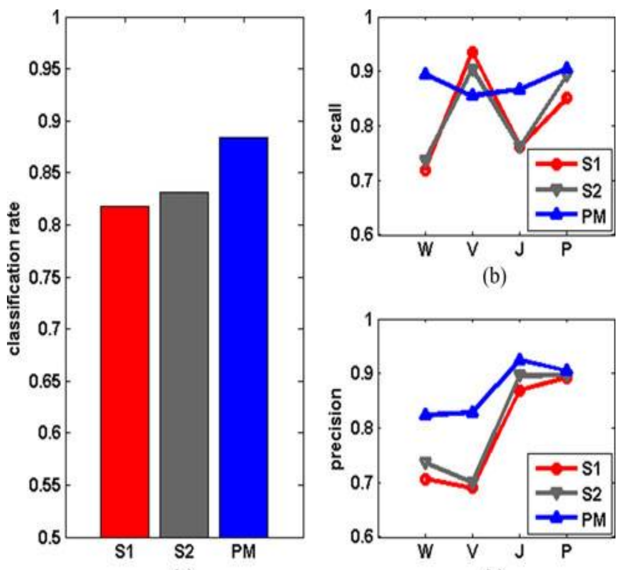
After the concentric level and feature set design, next step is to classify each image with different categories. The structures of lung nodules are same even after feature set design so to help the classification we need contextual information. This project contains SVM classification for lung nodule patches.

The first step is the probability estimation of the feature by using SVM. This is mainly based on lung nodule description. And second step is the contextual probability estimation using the classification method. SVM extracts both explicit and implicit information hidden between the images and their categories. Hence training data was utilised dual and acquires more information.

V. SIMULATION RESULTS

The classification rates of the two SVM-based approaches were quite similar at about 82% and 83%. The relatively small

Enhancement with S2 shows the benefit of context analysis, and also indicates the limit of SVM with its direct classification manner for lung nodule context identification. About7% more of the images were correctly classified using the proposed classifier. The proposed classifier categorizes the lung image based on the concentric level partition and the extracted feature. The level nodule probability and level-context probability are combined to classify the images.



(a) (b) (c)

Fig5.Comparison between other SVM and the proposed classifier: (a) classification rate, (b) recall, and (c) precision.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Comparison table

	W S1/S2/PM	V S1/S2/PM	J S1/S2/PM	P S1/S2/PM
W	71.9/73.7/89.5	**14.0/10.5/3.5**	7.0/7.0/1.8	7.0/8.8/5.3
V	3.2/3.2/4.8	93.5/90.3/85.5	1.6/3.2/3.2	1.6/3.2/6.5
J	7.1/7.1/2.7	**9.7/9.7/4.4**	76.1/76.1/86.7	8.8/7.1/6.2
P	4.8/3.4/3.4	4.8/4.8/2.7	5.4/2.7/3.4	85.0/89.1/90.5

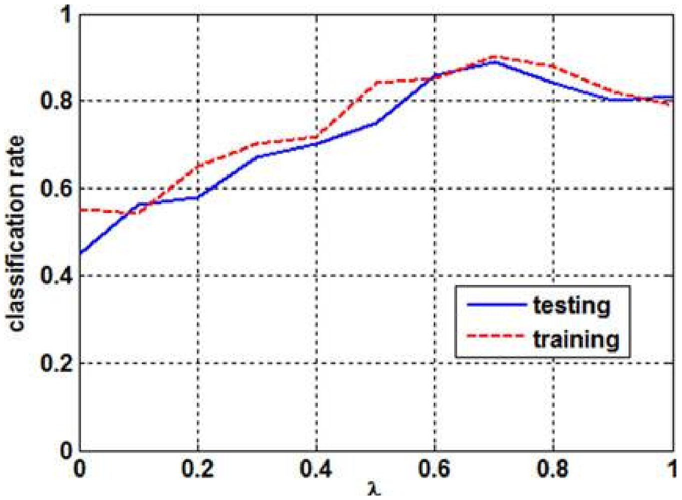


Fig6.Distributions of classification rates on the training and testing datasets.

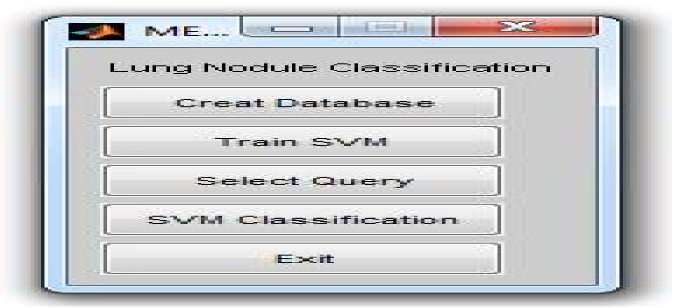


Fig7. Classification block

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Fig8. Query image

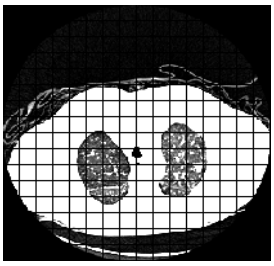


Fig9. Block generation

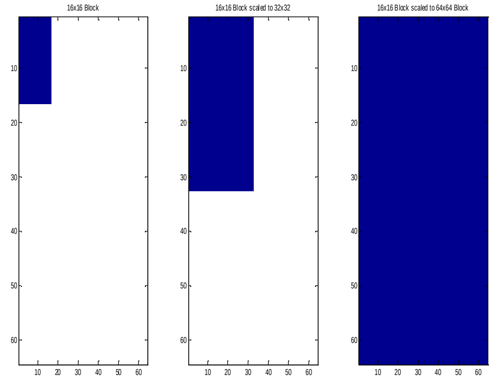


Fig10. Image enhancement

At last it will classify the lung nodule depends on the relative positions. That is well-circumscribed, Vascularized, juxta-pleural, and pleural-tail.

VI. CONCLUSION AND FUTURE WORK

We presented a novel based lung nodule classification method for lung nodule CT images. The four main types of lung nodules well circumscribed, Juxta pleural, Pleural tail and Vascularized objects to be isolated. We developed a NOVEL classification method to conquer the issue of nodule adjoining with the adjacent structures. This method contains mainly three parts concentric level partition, feature extraction and context analysis classification. A concentric level partition formed by superpixel formulation and construction of concentric level partition. Next feature set consists of multiorientation HOG and histogram of template feature was generated. And at last context analysis classification was used to classify the objects, where a supervised SVM classifier is used.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

The proposed method can be helpful for the medical and other general imaging domains. For example superpixel formation could be used as pre-processing stage for the patch based image analysis. Then feature set could be used as feature descriptor for different images. Latent semantic analysis for analysing the image patches.

The use of statistical classifier and feature descriptors helps to provide lung nodule detection and classification. The originality of this project is the detection of the nodules and isolating from the adjoining anatomical structures. It will be essential to train and test the system using the images obtained by CT scans. The result of work is to aware of lung cancer in the nodule which can be identified by radiologists which in turn save the life of human being.

REFERENCES

[1] J. J. Erasmus, J. E. Connolly, H. P. McAdams, and V. L.Roggli, “Solitary pulmonary nodules: Part I. morphologic evaluation for differentiation of benign and malignant lesions,” *Radio graphics*, vol. 20, no. 1, pp. 43–58, 2000.

[2] D. Wu, L. Lu, J. Bi, Y. Shinagawa, K. Boyer, A. Krishnan, and M. Salganicoff, “Stratified learning of local anatomical context for lung nodules in CT images,” in *Proc. Comput. Vis. Pattern Recog.*, 2010, pp. 2791–2798.

[3] R.A. Ochs, J.G. Goldin, F. Abtin, H. J. Kim, K. Brown, P. Batra, D. Roback, M. F. McNitt-Gray, and M. S. Brown, “Automated classification of lung bronchovascular anatomy in CT using adaboost,” *Med. Image Anal.*, vol. 11, no. 3, pp. 315–324, 2007.

[4] A.Farag, S. Elhabian, J. Graham, A. Farag, and R. Falk, “Toward precise pulmonary nodule descriptors for nodule type classification,” in *Proc. Med. Image Comput. Comput.-Assisted Intervention Conf. Lecture Notes Comput. Sci.*, 2010, vol. 13, no. 3, pp. 626–633.

[5] D. Unay and A. Ekin, “Dementia diagnosis using similar and dissimilar retrieval items,” in *Proc. Int. Symp. Biomed. Imag.*, 2011, pp. 1889–1892.

[6] Y. Song,W. Cai,Y. Zhou, L.Wen, and D. Feng, “Pathology-centric medical image retrieval with hierarchical contextual spatial descriptor,” in *Proc.Int. Symp. Biomed. Imag.*, 2013, pp. 202–205.

[7] J. Galaro, A. Judkins, D. Ellison, J. Baccon, and A. Madabhushi, “An integrated texton and bag of words classifier for identifying anaplastic medulloblastomas,” in *Proc. IEEE Annu. Int. Conf. Eng. Med. Biol. Soc.*,
2011, pp. 3443–3446.

[8] A. Cruz-Roa, F. GonzA˜ lez, J. Galaro, A. Judkins, D. Ellison, J. Baccon, A. Madabhushi, and E. Romero, “A visual latent semantic approach for automatic analysis and interpretation of anapaestic medulloblastoma virtual slides,” in *Proc. Med. Image Comput. Comput.-Assisted Intervention.*