Automated Lung Cancer Detection in CT Images

Anitha S.B(M.Tech)[1], Dr. N Manja Naik[2]
[12]*Dept. of Electronics and Communication Engineering*
U.B.D.T college of Engineering, Davangere, Karnataka, India

Abstract- Lung diseases are the most common disease which causes mortality worldwide. In this study, the computed tomography images are used for the diagnosis of the lung diseases such as normal, small cell lung carcinoma, large cell lung carcinoma and non-small cell lung carcinoma by the effective extraction of the global features of the images and feature selection techniques. The images are recognized with the statistical and the shape based features. The texture based features are extracted by Gabor filtering, the feature outputs are combined by watershed segmentation and the fuzzy C means clustering. Feature selection techniques such as Information Gain, correlation based feature selection are employed with Genetic algorithm which is used as an optimal initialization of the clusters. The dataset of lung diseases for four classes are considered and the training and testing are done by the KNN and SVM.

Index Terms- CT, Automatic segmentation, KNN and SVM Classifiers

I. INTRODUCTION

Lung cancer isa disease of abnormal cells multiplying and increasing thetumor. Cancer cells continue to increase abnormal cell and new cells. So finding lung cancer earlier is most important for successful treatment.Diagnosis is mainly based on CT scan images. Cancerous tumor starts in the part of lung is called primary lung cancer. Lung canceris divided into two main types:

1. Small cell cancer
2. Non small cell cancer

A mass of tissue that originates by a slow development of strange cells is known as a tumor. Tumor cells are those cells that develop, despite the fact that when the body does not require them, and besides as typical old cells, they don't lapse.The malignancy cells present in lung causes lung growth illness. These cells discovery is critical issue for medicinal specialists. The odds of a compelling treatment will essentially increment with early recognition.

The Computed Tomographypictures are utilized which are more proficient than Xbeam. Tumor cells are identified in lung malignancy CT pictures by utilizing marker controlled watershed grouping method. Watershed gives better results compared with other segmentation algorithms and calculation time is less in watershed segmentation. Classification algorithms Support Vector Machine and KNN is used. Evaluation would be done on the basis of correctly classified trial data. MATLAB is broadly utilized programming for the investigation of lung disease identification from CT filter pictures. This work concentrates on discovering tumor and its stages. In this Marker-controlled Watershed division is utilized to disengage a lung of a CT picture. A graphical user interface is developed to scan all the images and display the features and cancer stage. This system can help in early detection of lung cancer more accurately. The accuracy of the this paper is obtained by 80-90%.

II. LITERATURE SURVEY

[1] M. Norouzi, M. Ranjbar, and G. Mori, "Heaps of Convolutional Restricted Boltzmann Machines for Shift-Invariant Feature Learning," in IEEE Conference on Computer Vision and Pattern Recognition, 2009.

In this paper display a technique for learning class specific highlights for acknowledgment. As of late an eager layer wise methodology was proposed to instate weights of profound conviction systems, by review each layer as a different Restricted Boltzmann Machine (RBM). We build up the Convolutional RBM (C-RBM), a variation of the RBM show in which weights are shared to regard the spatial structure of pictures. This system takes in an arrangement of highlights that can produce the pictures of a particular protest class. Our component extraction display is a four layer order of rotating sifting and most extreme sub sampling. We learn include parameters of the first and third layers seeing them as independent C-RBMs. The yields of our element extraction chain of command are then nourished as contribution to a discriminative classifier. It is tentatively exhibited that the extricated highlights are compelling for question location, utilizing them to get execution practically identical to the best in class on written by hand digit acknowledgment and person on foot discovery.

[2] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional profound conviction systems for versatile unsupervised learning of various leveled portrayals," in ICML, New York, New York, USA, 2009, pp. 609– 616

There has been much enthusiasm for unsupervised learning of various leveled generative models, for example, profound conviction systems. Scaling such models to full-sized, high-dimensional pictures remains a troublesome issue. To address this issue, we show the convolutional profound conviction organize, a various leveled generative model which scales to practical picture sizes. This model is interpretation invariant and backings productive base up and top-down probabilistic deduction. Key to our approach is probabilistic max-pooling, a novel procedure which contracts the portrayals of higher layers in

a probabilistically solid manner. Our investigations demonstrate that the calculation learns helpful abnormal state visual highlights.

III. OBJECTIVES

The main objective is

- To implement SVM and KNN classifier
- To achieve accuracy of minimum 80%.
-

IV. METHODOLOGY

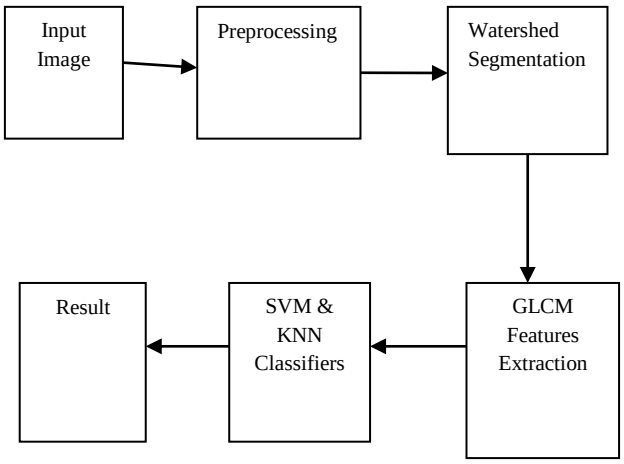

Fig.1: system architecture

A .Image Preprocessing

Preprocess the CT image slice by using the Gaussian filter. Picture improvement procedures in Picture Preparing Tool kit empower to expand the flag to-clamor proportion and emphasize picture includes by changing the hues. Edge identification calculations let distinguish question limits in a picture. These calculations incorporate the Sobel, Prewitt, Roberts, Watchful, and Laplacian of Gaussian strategies. The Watchful strategy can identify genuine feeble edges without being tricked by commotion.

B. Image Segmentation

Segment the preprocessed image using the marker controlled watershed segmentation. This is the separating a picture into different parts. This is commonly used to recognize objects or other important data in computerized pictures. The objective of division is to improve or potentially change the portrayal of a picture into something that is more significant and simpler to examine

Marker Controlled Watershed Segmentation: This count considers the information picture as a topographic surface (where higher pixel regards mean higher height) and reenacts its flooding from specific seed centers or markers. A typical decision forthe markers are the nearby minimaof the slope of the picture, yet the technique takes a shot at a particular marker, either chose physically by the client or decided naturally by another calculation.

Morphological and filtering approach:The Morphological operation is used for the different filtering techniques for a different input image. This operation consists dilation and erosion techniques. Dilation it can perform both binary and grey tone image.Disintegration relies upon Structuring component measure with bigger Structuring components having an increasingly articulated impact and the aftereffect of Erosion with an extensive Structuring component is like the outcome acquired by iterated Erosion utilizing a littler organizing component of a similar shape.

C.Feature Extraction

Extract the feature from the binary image by using the Gray Level Co-Occurrence Matrix techniques.

GrayLevel Co-OccurrenceMatrix (GLCM):This is commonly used techniques to extract texturaldata of Images is GrayLevelCo- occurrenceMatrix. This technique gives sensible surface data of a picture that canbeacquired just fromtwopixels. Dark level co-event frameworks acquainted by Haralick endeavor with portray surface by measurably inspecting how certain dim levels happen in connection to other dim levels.

D.Classification

With the extracted feature identify the stage of cancer by using SVM and KNN classifiers.

K-Nearest Neighbor (KNN): This is simplest method for all machine learning algorithm. KNN is a non-parametric method used for regression and classification. The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. In order to identify neighbors, the objects are represented by position vectors in a multidimensional feature space. It is usual to use the Euclidean distance, though other distance measures, such as the Manhattan distance could in principle be used instead.This algorithm is sensitive to the local structure of the data.

Support Vector Machines (SVM): Used generally for arrangement/Classification(additionally, can beadjusted for relapse and notwithstanding forunsupervisedlearningapplications). It meetsexactness tantamount to Multilayer Perceptions.

V. FLOWCHART

Algorithm workflowisrepresented instepsusing theflowcha rt. Using the various kind of notations the procedure will be represented and the arrowsare used for the connection. The solution to the given problem can be easily found by the flow chat. This are the keypoints which dealwith the perfect identification and formulizing the structure.

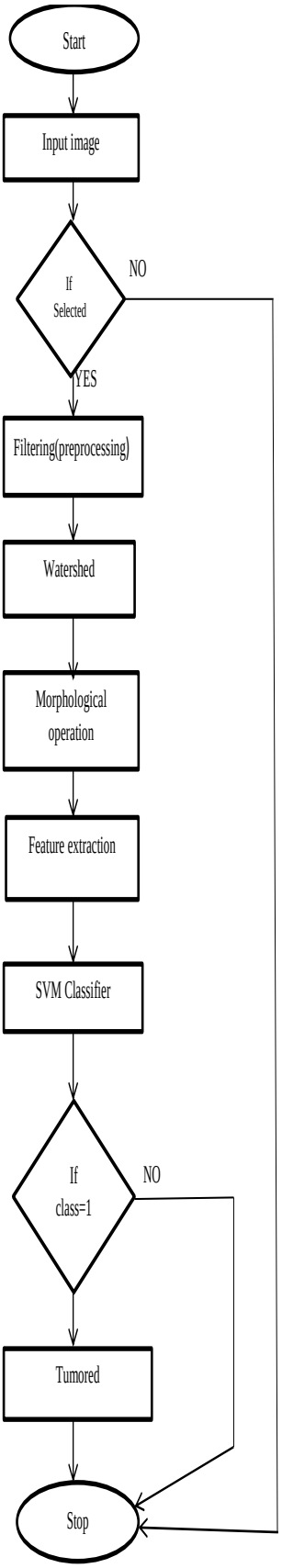

Fig.2: Flowchart

VI. RESULT

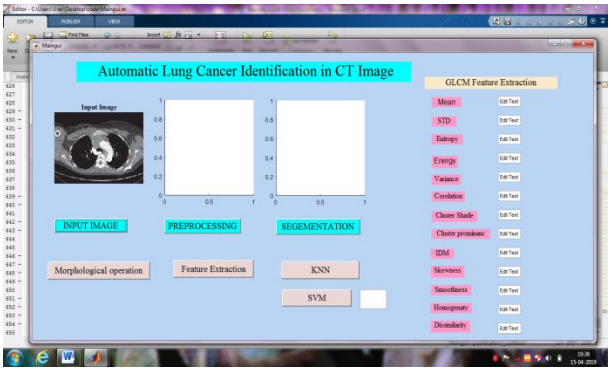

Fig.3: Input Image

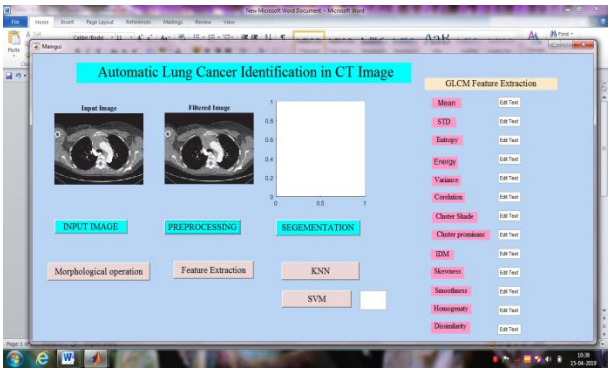

Fig.4: Preprocessing

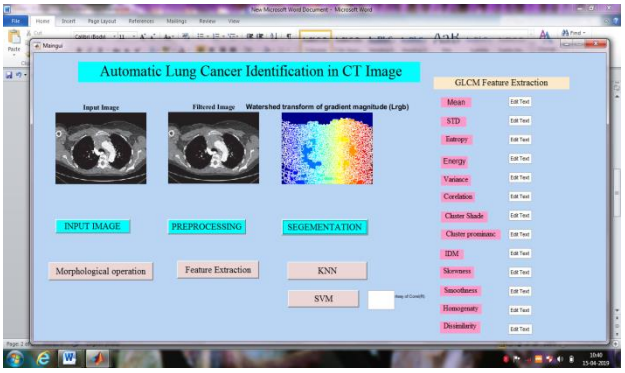

Fig.5: Segmentation

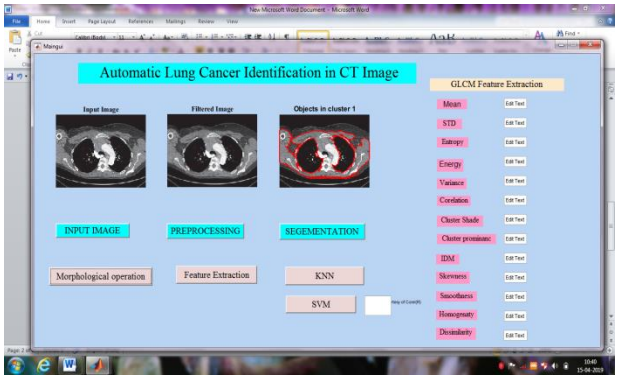

Fig.6: Morphological operation

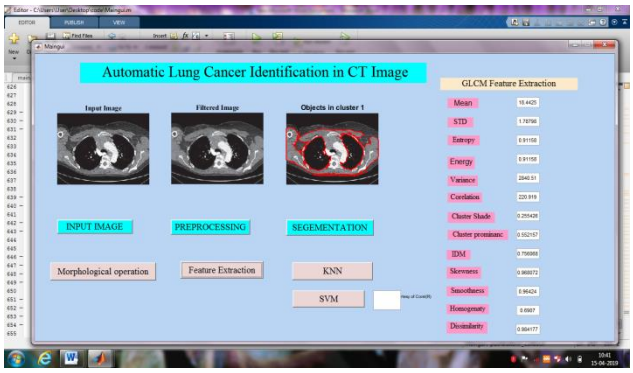

Fig.7: Feature Extraction

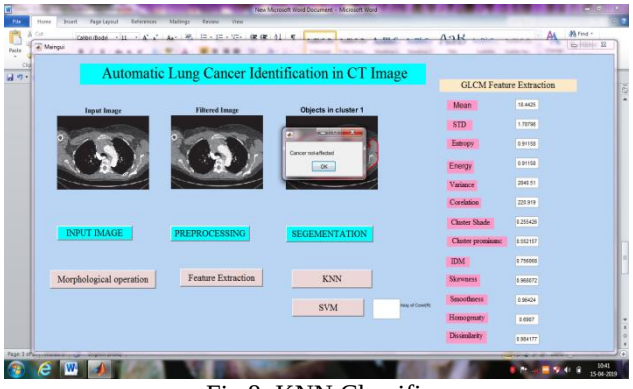

Fig.8: KNN Classifiers

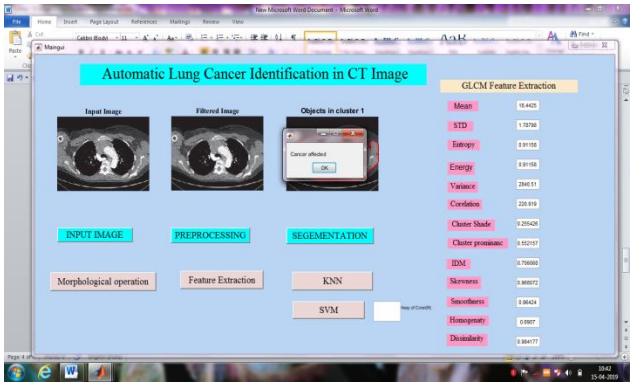

Fig.9: SVM Classifiers

Identifies whether the lung cancer is affected by cancer or not.

V. CONCLUSION

The proposed marker controlled watershed division method isolates the contacting objects in the picture. It gives best recognizable proof of the primary edge of the picture and furthermore stays away from over division.Two differentclassifiers are usedSVM andKNN classifier increasesaccuracy of detectionand reduces falsedetection.

VI. FUTUREWORK

Future work will include the modification of the system to recognize images that have multiple scattered tumor on the lung.the proposed work can be extended by additional range of algorithms on additional range of dataset from huge medical database.the perormance analysis of different classification algorithms can be done to provide better result.

VII. REFERENCES

[1]. Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," Universite de Montr ´ eal, Tech. Rep., ´ 2012.

[2]. H. Larochelle, M. Mandel, R. Pascanu, and Y. Bengio, "Learning Algorithms for the Classification Restricted Boltzmann Machine," Journal of Machine Learning Research, vol. 13, pp. 643–669, Mar. 2012.

[3]. G. Desjardins and Y. Bengio, "Empirical Evaluation of Convolutional RBMs for Vision," Universite de Montr ´ eal, Tech. Rep., 2008. ´

[4]. M. Norouzi, M. Ranjbar, and G. Mori, "Stacks of Convolutional Restricted Boltzmann Machines for Shift-Invariant Feature Learning," in IEEE Conference on Computer Vision and Pattern Recognition, 2009.

[5]. H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in ICML, New York, New York, USA, 2009, pp. 609–616.

[6]. "Unsupervised learning of hierarchical representations with convolutional deep belief networks," Communications of the ACM, vol. 54, no. 10, pp. 95–103, Oct. 2011.

[7]. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[8]. A. Depeursinge, A. Vargas, A. Platon, A. Geissbuhler, P.-A. Poletti, and H. Muller, "Building a reference multimedia database for interstitial ¨ lung diseases," Computerized Medical Imaging and Graphics, vol. 36, no. 3, pp. 227–38, Apr. 2012.

[9]. A. Depeursinge, A. Foncubierta-Rodr´ıguez, D. Van de Ville, and H. Muller, "Lung Texture Classification Using Locally-Oriented Riesz ¨ Components," in MICCAI, 2011, pp. 231–238.

[10]. "Multiscale Lung Texture Signature Learning Using the Riesz Transform," in MICCAI, 2012, pp. 517–524.

[11]. A. Depeursinge, D. Van de Ville, A. Platon, A. Geissbuhler, P.-A. Poletti, and H. Muller, "Near-affine-invariant texture learning for lung ¨ tissue analysis using isotropic wavelet frames." IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society, vol. 16, no. 4, pp. 665–75, Jul. 2012.

[12]. A. Depeursinge, T. Zrimec, S. Busayarat, and H. Muller, "3D lung ¨ image retrieval using localized features," in SPIE Medical Imaging, 2011.

[13]. Y. Song, W. Cai, Y. Zhou, and D. D. Feng, "Feature-Based Image Patch Approximation for Lung Tissue Classification," IEEE Transactions on Medical Imaging, vol. 32, no. 4, pp. 797–808, Apr. 2013.

[14]. Y. Song, W. Cai, H. Huang, Y. Zhou, D. Feng, Y. Wang, M. Fulham, and M. Chen, "Large Margin Local Estimate with Applications to Medical Image Classification," IEEE Transactions on Medical Imaging, vol. 34, no. 6, pp. 1362–1377, Jun. 2015.

[15]. **Manjanaik**, **N**. & Manjunath, R. 2013, Selection Of Intra Prediction Modes For Intra Frame Coding In Advanced Video Coding Standard, Ijret: International Journal Of Research In Engineering

And Technology, Eissn: 2319–1163| Pissn: 2321–7308, Volume: 02

[16]. **N Manjanaik**, RManjunath - 2016 International Conference on …, 2016 - ieeexplore.ieee.org This paper proposes Intra frame coding in Advanced Video Coding Standard/H. 264 to reduce bit rate and control PSNR using new technique is Gaussian pulse. The previous work on Intra frame in Advanced Video Coding Standard/H. 264 using fast mode decision