

Classification of Microscopic Images for fibrosis Recognition Using Multiple classifiers

Sheenum Marwaha

Gurpinder Kaur

Shelza

Mtech Student

HOD CSE

Assistant Professor

Computer & Engineering Department
SVIET Banur, Punjab, India

Computer & Engineering Department
SVIET Banur, Punjab, India

Computer & Engineering Department
SVIET Banur, Punjab, India

Abstract: - Computerized aided diagnosis plays a very important role in image processing especially when we want to get the result of diagnosis with reliability and without error. Microscopic Images and Computerized tomographic (CT) images are diagnosed so that the abnormal area in the image can be easily recognized. For the recognition of dissimilar or diseased area in image, we need an important region of the image rather than the whole image. So, we segregate the required region from the image. Spatial regions in images that are most important to the end user are called regions of interest (ROIs). In this paper we are recognizing the fibrosis in microscopic tissue images on the basis of prominent features used in identifying fibrosis in microscopic tissue images and the usage of computerized aided diagnosis in identifying the disease using digital Image processing. Conventional methods like unnecessary needle biopsies, manual diagnosis used in identifying the abnormal region requires more time, the quality of diagnosis is also bad and having very little reliability. In this paper, interest is on the discussion of the steps followed while doing automated computerized diagnosis and the use of mainly four types of classifiers i.e. Radial basis function (RBF), support vector machine (SVM), Bayesian and K-Nearest neighbour (KNN) on fibrotic tissues dataset for recognition of fibrosis. Performance of these techniques is compared through sensitivity, specificity, accuracy etc. Different steps involved in the methodological analysis adopted are also summed up.

Keywords— Image analysis, Computer Aided Diagnosis (CAD) System, region of interest, Image Segmentation, feature extraction, radial basis function, support vector machine

I. INTRODUCTION

The fibrosis is the scarring response formed in the chronic injury of any cause. Fibrosis is a tissue disease because when the body's connective tissues becomes injured by inflammation so that they come under attack [1]. It is very challenging to recognize fibrosis in tissues because of the lack of cognition of the most outstanding features which help in recognition of the fibrosis in microscopic tissue images. It has received extensive attention last two decades, not only because many people are affected with the disease every year which include infants also but also fibrosis recognition is a typical problem whose solution needs to be explored and would help in many other disease classification problems.

Earlier, pathologists used biopsy samples taken from patients for microscopic examination and then made personal

judgement on them. A pathologist typically values the deviations in the cell structures or the change in the distribution of the cells across the tissue while studying such images. However, this judgement is subjective, and often leads to considerable variability [2-3]. In spite of major attainments in our understanding the identification of fibrosis is one of the major research fields in medical imaging which need to be further explored. So, to ameliorate the reliability and quality of fibrosis recognition it is essential to develop computation tools for automatised fibrosis recognition that facilitates objective mathematical assessment.

This paper proposes a system to make the computer automatically locate the fibrosis location in the tissue image and compute relevant features.

The two key steps involved in the implementation of CAD systems are segmentation and classification of wary parts [4].

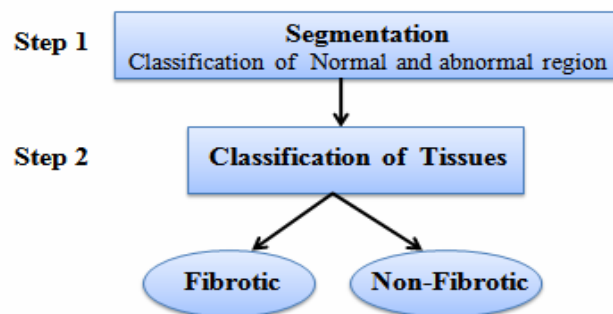


Fig. 1 Key steps involved in the implementation of CAD systems

II. COMPUTERIZED AIDED DIAGNOSIS FOR FIBROSIS RECOGNITION

The CAD machine is provided with functions that automatically examines the acquired data. It help scan digital images, e.g. from microscopic and tomography images, for typical appearances and recognize possible diseases i.e. to highlight blazing sections. CAD is fundamentally based on extremely complex pattern recognition. Usually a few thousand images are needed to optimize the algorithmic rule.

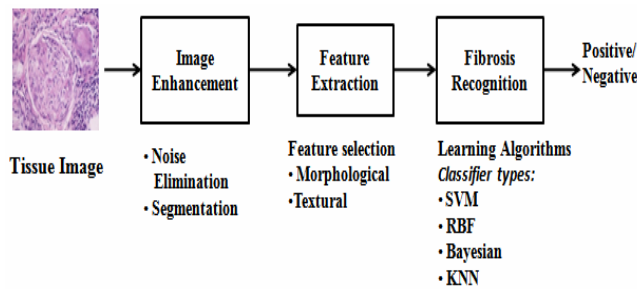


Fig. 2 Overview of the computational steps and techniques used in automated fibrosis recognition

In paper [5], a beneficial illustration of computerized aided diagnosis using real time tissues is shown with the help of experimentation. It is explained that diagnosing rules are generated from those tissue images at the last level with the help of Feature selection and Region of Interest Properties and ultimately classification patterns are made which further passed to the classifiers for the determining procedure. A image is then given to the suggested automatized CAD system and finally processing steps will be followed and finally it will detect whether the provided tissue image is malignant or healthy.

III. ALGORITHM

This section explicates the algorithm and methodological analysis used and basic steps involved in it which are followed for the identification of fibrosis.

There are seven main steps included in the suggested automated computer aided diagnosis algorithm for fibrosis identification which are as follows:

1. The input image is read by using imread function.
2. The quality of the given tissue image is enhanced by using contrast enhancement technique.
3. Then Color-Based Segmentation Using K-Means Clustering is carried out for finding Region of interest. The image is segmented in an automated fashion using the L*a*b* color space and K-means clustering.
4. Finding Region of Interest (ROI) is the next step. In this step, the boundaries of a tissue or the internal nucleus of the tissue is analysed for finding the main regions which are of interest. It may be defined on an image or in a volume, for the purpose of measuring various features like size, area, perimeter etc.
5. In this step, features for tissues are selected and factorization methods are applied. The goal of feature extraction is to obtain representative features that can be used to recognise the defective tissue. The shape-based features are Total area, Perimeter, Diameter, Eccentricity, majoraxis and minoraxis etc.
6. Dataset will be formatted based on various extracted features.
7. Machine algorithm is run to find out the results.

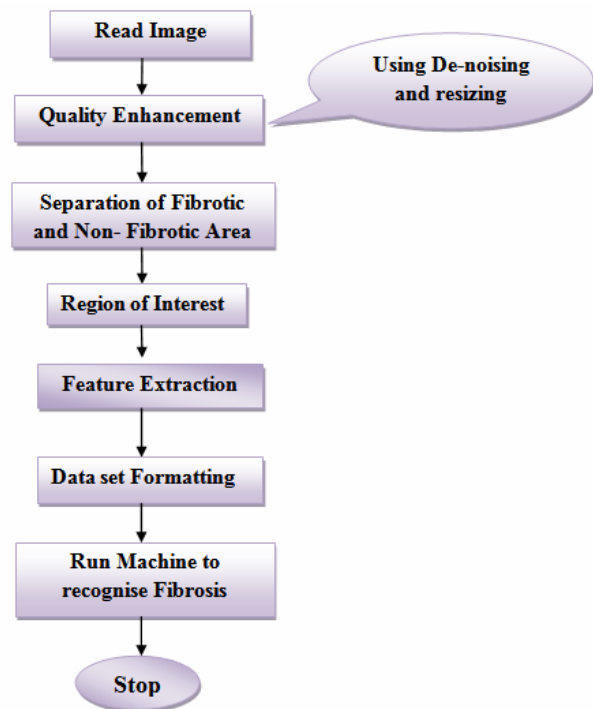


Fig. 3 Flowchart of Methodology used

The steps are detailed below:-

- STEP 1: - Read Image
The Images of tissues are pre-processed and read as matrix having RGB Intensity. All the images are resized before reading because images taken by high-definition cameras can be of large size.
- STEP 2:-Image Enhancement
Image enhancement[6] is the melioration of digital image quality without cognition about the source of abjection. It improves the clarity of images for human vision and helps in removing blurring and noise, increase contrast etc. In [7], it is written that an image might be taken of an endothelial cell, which might be of low contrast and slightly blurry. Reducing the noise, blurring and increasing the contrast range could enhance the image.
- STEP 3:- Segmentation
Using K means[8-9] segmentation is carried out to distinguish between Fibrotic and Non- Fibrotic Areas in the Tissues and also fibrotic and non-fibrotic tissues are separated based on regional descriptors. Images can be segmented into regions by using K Means Clustering Segmentation. It is a method commonly used to automatically partition a data set into k groups.
- STEP 4:- Region of Interest(ROI)

Region of Interest is filtered out after K means. One or more ROI can be determined in a tissue image. The regions can be defined by a range of intensities or they can be geographic in nature, such as polygons that encompass contiguous pixels [10]. This concept is commonly used in medical imaging. For instance, the boundaries of a tissue may be defined on an image or in a volume, for the purpose of measuring its size [11].

- STEP 5:- Feature Extraction

In this step features for tissues are selected and factorization methods are applied. It is the process of transforming the input data into the set of features [12]. The input data will be transformed into a reduced representation set of features when it is too large for an algorithm to be processed and it is suspected to be redundant. The features set would evoke the relevant information if the features extracted are carefully selected.

For example, if the fibroid is of very small size, it can be treated by giving medication, whereas if it is of big size, it should be removed by operation [13]. The shape-based features chosen in the recognition of fibrosis are total area of region affected, perimeter, diameter, eccentricity, major axis and minor axis etc.

- STEP 6:- Dataset Formatting

Dataset will be formatted based on various extracted features and on the basis of these features Radial basis function, support vector machine, Bayesian and KNN classifiers are applied to give results.

The diagnosis principle can be planned to determine the exact candidate region. This diagnosis principle can eliminate the fear of delusive recognition of abnormal part and provides better diagnosis.

- STEP 7:- Run Machine Algorithm to recognise Fibrosis

It is the last step executed in the fibrosis recognition. The machine algorithms are run to recognise check the performance of classifiers to detect fibrosis in the tissues and results are find out. Diagnosis Rules are defined based on Region of interest attributes an malignant regions more accurately detected. It also helps in ignoring all the false positive morbid parts.

- Support Vector Machine:- SVMs have been successfully applied for categorization to an progressively wide variety of biological applications [14]. The main aim of the SVM is to distinguish between two classes by deciding the linear classifier that maximizes the border and it is known as the optimal separating hyperplane [15]. Mathematically, the convention vector of the hyper plane is the weighted sum of the Support Vectors. Also SVM classifier works efficiently in several image classification problems [16-19].

- Radial basis function:- It is neural network function which his applied as an assistant tool for the recognition of the Fibrosis. Other classifiers based on it have been successfully implemented for different purposes such as fault detection, face recognition or medical diagnosis etc.[20-21]. The most basic form of Radial-Basis Function network (RBF) involves three layers which are having entirely different roles. First is Input layer in which source nodes can connect to the network to its environment. Then comes the hidden layer which provides a set of basis function and high dimensionality and the last is output layer which is linear combination of hidden function.

- K-Nearest Neighbour:- It the simplest of all machine learning algorithms where objects are categorized based on snuggest training instances in the feature space. It functions on the supposals that categorization of unknown instances can be done by relating the unknown to the known according to a similarity function. Its accuracy can be severely degraded by the presence of noise or irrelevant features etc. It uses the unsupervised learning algorithm, where the class target is unknown [22]. It is very robust with regard to the search space because classes need not to be linearly separable.

- Bayesian Classifier:- Bayesian network[23-24] or directed acyclic graphical prototype represents a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). It expeditiously encodes the joint probability distribution for a large set of variates. It is considered one of the statistical classifiers. It devolves on how much the tested ROI has a distribution which correlates with each of the two clusters. Based on the size of ROI it is used to classify fibrotic and non-fibrotic Images.

IV. MULTIPLE LEARNING ALGORITHMS USED FOR EVALUATION

Classifier refers to the mathematical mapping, applied by a classification algorithm that maps input data to a particular category. The aimed automated automatized machine learning algorithm is for the classification of occurrence and non occurrence of fibrosis in a tissue. The multiple classifiers that are used for the categorization of fibrotic and non-fibrotic tissues are discussed below:-

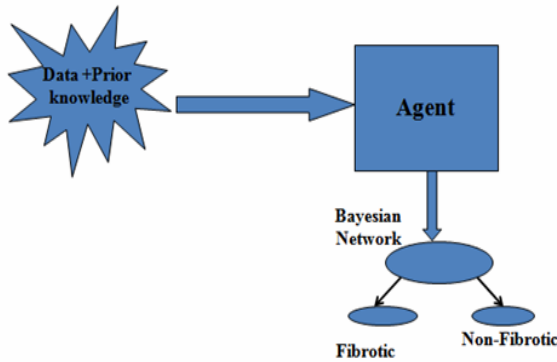


Fig. 4 Bayesian Classifier

V. RESULTS

To recognize fibrosis and test results having high accuracy and to check the performance of the proposed algorithm, some experiments are conducted on several colored and gray scale images. Enhancement, segmentation, finding ROI and dataset formatting are the various steps used in the proposed algorithm for recognition of fibrosis. The quality metrics that are used to evaluate the performance of purposed algorithm are: sensitivity, specificity, accuracy The values for Accuracy should be higher for better results.

TABLE I. SENSITIVITY AND SPECIFICITY MEASURES

Metrics	RBF	SVM	Bayesian	KNN
Sensitivity(SE)	0.85	0.94	0.88	1.00
Specificity(SP)	1.00	0.91	1.00	0.96

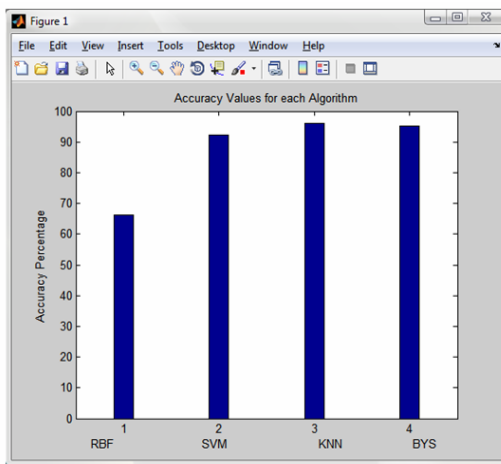


Fig. 5 Graph showing Accuracy

VI. CONCLUSION

In this paper, a computer-assisted based on multiple classifiers i.e SVM ,RBF neural network, KNN and Bayesian is introduced for the recognition of fibrosis in images. The study presented in this paper is quite innovative. Little work

has done on fibrosis recognition of lung microscopic tissue images[25]. Only on computed tomographic and MRI Images lung evaluation has done before [26-28]. Computerized assessment of fibrosis basically on liver is presented earlier in literature [29-31].

This paper also concludes that computerized-aided diagnoses give a lot of information that help doctors in the accurate and consistent diagnosis of medical images and also the time taken in reading an image by traditional methods could be decreased. Also it is provided with functions that automatically analyses data and provides tissue diagnosis automatically to identify the suspected regions from images. Four classifiers have been used for the study . Various quality metrics have been used for evaluating the results to compute these techniques quantitatively. The proposed method gives good visual as well as quantitative results for all these test images.

Further it concludes that knowledge of most salient features for the fibrosis recognition is very useful for doctors. But the procedure adopted in recognition of malady is still dawdling in various aspects. It is still challenging and its accuracy has been an matter of great worry. Therefore, if further new and advanced methods like mass spectroscopic analysis, laser capture micro dissection, microarrays etc are attached with various classifiers it can give more better and accurate results. Also, a number of investigators are investigating new CAD algorithms for recognising fibrosis which are very useful and which can enhance automatized diagnosis abilities to recognise candidate region. Examples should be given to show that the new approach is highly effectual and gives more effective results.

REFERENCES

- [1] Monica Lupşor, Sergiu Nedeveschi, Cristian Vicaş, Radu Badea, Mircea Grigorescu, Simona Tripon, Horia Stăfănescu, Corina Radu, Alexandru Serban, Titus Suteu, " Estimating the fibrosis stage in the human liver tissue using image processing methods on ultrasonographic images", Proceedings of the 3rd International Conference- EMMIT 2007
- [2] A. Andrión, C. Magnani, P.G. Betta, A. Donna, F. Mollo, M. Scelsi, P. Bernardi, M. Botta, B. Terracini, Malignant mesothelioma of the pleura: interobserver variability, J. Clin. Pathol. 48(1995) 856-860.
- [3] S.M. Ismail, A.B. Colclough, J.S. Dinnen, D. Eakins, D.M. Evans, E. Gradwell, J.P. O'Sullivan, J.M. Summerell, R.G. Newcombe, Observer
- [4] Jaya.j. and Thanushkodi K., "Structural Modeling and Analysis of CAD System: A Graph Theoretic Approach." *International Journal of Computer Science and Applications*, Vol. 2, 2009, pp.5-8.
- [5] M.GOMATHI , Dr.P.THANGARAJ," A Computer Aided Diagnosis System For Detection Of Lung Cancer Nodules Using Extreme Learning Machine", IJEST, October 2010
- [6] Raman Maini and Himanshu Aggarwal." A Comprehensive Review of Image Enhancement Techniques", JOURNAL OF COMPUTING, VOLUME 2, ISSUE 3, MARCH 2010, ISSN 2151-9617
- [7] Ambica Verma, Priyanka Gupta, Ramandeep Singh," Enhancement of ultrasound images using fuzzy sets and comparing the results of Radon and Fourier transforms", IJCST Vol. 2, Iss ue 2, June 2011, S N : 2 2 2 9 - 4 3 3 3 (P r i n t) | I S S N : 0 9 7 6 - 8 4 9 1 (O n l i n e)
- [8] Oyelade, O. J, Oladipupo, O. O, Obagbuwa, I. C," Application of k-Means Clustering algorithm for prediction of Students' Academic Performance", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 7, No. 1, 2010
- [9] Tapas Kanungo, Nathan S. Netanyahu," An Efficient k-Means Clustering Algorithm: Analysis and Implementation", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 7, JULY 2002

- [10] V.D.AmbethKumar,Dr. M. Ramakrishnan, G.Jagdeesh, A Discrete correlation of Footprint Image aforementioned to Recognition, National Conference on Future Computing Volume 1, March 2012
- [11] Balpreet Kaur, Deepak Aggarwal, Gurpreet Kaur, Amandeep Kaur, Efficient Image Compression based on Region of Interest, IJCST Vol. 2, Issue 1, March 2011, ISSN : 2229-4333 (Print) | ISSN : 0976-8491
- [12] V.P.Gladis Pushpa Rathil and Dr.S.Palani, "A novel approach for feature extraction and Selection on mri images for brain tumor Classification" David C. Wyld, et al. (Eds): CCSEA, SEA, CLOUD, DKMP, CS & IT 05, pp. 225–234, 2012, © CS & IT-CSCP 2012
- [13] Ratha Jeyalakshmi and Ramar Kadarkarai, Segmentation and feature extraction of fluid-filled uterine fibroid–A knowledge-based approach, Received: 22 February 2010 / Accepted: 28 September 2010 / Published: 5 October 2011, ISSN 1905-7873
- [14] William S Noble, What is a support vector machine?, Primer (Computational Biology), Nature's publishing Group @2006
- [15] Gunn, S., "Support vector machines for classification and regression, Technical paper, 1998.
- [16] G. Camps-Valls and L. L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
- [17] G. M. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification," *Remote Sens. Environ.*, vol. 93, pp. 107–117, 2004.
- [18] Foody, G.M , "A relative evaluation of multiclass image classification by support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1335–1343, Jun. 2004.
- [19] K. I. Kim, J. H. Kim, and K. Jung, "Face recognition using support vector machines with local correlation kernels," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 16, pp. 97–111, 2002.
- [20] J. Víctor Marcos ,Roberto Hornero, Daniel A ´ lvarez, Fe'lix del Campo, Miguel Lo'pez, Carlos Zamarro'n, "Radial basis function classifiers to help in the diagnosis of the obstructive sleep apnoea syndrome from nocturnal oximetry", @International Federation for Medical and Biological Engineering 2007.
- [21] Adrian G. Bors," Introduction to Radial Basis Function networks", Department of Computer science university of York, TO10 5DD, UK.
- [22] Friedman, N., Geiger, D., and Goldszmidt, M., "Bayesian classifier", Machine learning, Vol. 29, 1997, pp.131-163
- [23] Amir Fallahi, Shahram Jafari," An Expert System for Detection of Breast Cancer Using Data Preprocessing and Bayesian Network", International Journal of Advanced Science and Technology Vol. 34, September, 2011
- [24] MONASH UNIVERSITY," CSE5230 Tutorial: The Na'ive Bayes Classifier", David McG. Squire, August 19, 2004
- [25] Ilias Maglogiannis, Haralambos Sarimveis, C. T. Kiranoudis, Aristotelis A.Chatziioannou, Nikos Oikonomou, and Vassilis Aidinis (2008), "Radial basis function neural networks classification for the recognition of idiopathic pulmonary fibrosis in microscopic images, Ieee Transactions On Information Technology In Biomedicine, Vol. 12, No. 1, January 2008
- [26] R. Blechschmidt, R. Werthschutzky, and U. Lorcher, "Automated CT image evaluation of the lung: Amorphology-based concept," *IEEE Trans. Med. Imag.*, vol. 20, no. 5, pp. 434–442, May 2001.
- [27] L. Hsu and D. Schimel, "Computed tomography imaging of lungs in mouse models of human disease: Advancing the computing interfaces with physiology," in *Proc. 17th IEEE Symp. Comput.-Based Med. Syst. (CBMS 2004)*, pp. 385–390.
- [28] Y. Chang, P. C. Chung, and P. H. Lai, "Using a spatiotemporal neural network on dynamic gadolinium-enhanced MR images for diagnosing recurrent nasal papilloma," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 1, pp. 225–238, Feb. 2002.
- [29] T. Caballero, A. P'erez-Milena, M. Maseroli, F. O'Valle, F. J. Salmer'on, R. M. G. Del Moral, and G. S'anchez-Salgado, "Liver fibrosis assessment with semiquantitative indexes and image analysis quantification in sustained-responder and non-responder," *J. Hepatol.*, vol. 34, pp. 740–747, 2001.
- [30] M. Maseroli, T. Caballero, F. O'Valle, R. M. G. Del Moral, A. P'erez- Milena, and R. G. Del Moral, "Automatic quantification of liver fibrosis: Design and validation of a new image analysis method. Comparison with semi-quantitative indexes of fibrosis," *J. Hepatol.*, vol. 32, pp. 453–464, 2000.
- [31] X. Z. Lin, Y. N. Sun, M. H. Horng, and X. Z. Guo, "Computer morphometry for liver fibrosis using an automatic image analysis system," *Eng. Med. Biol. Soc.*, vol. 2, pp. 682–683, 1996.