IJCSE International Journal of Computer Sciences and Engineering **Open Access**

Survey Paper **Volume-5, Issue-3** **E-ISSN: 2347-2693**

A Survey on Classification of Liver Diseases using Image Processing and Data Mining Techniques

R.V. Patil[1*], S.S. Sannakki[2], V. S. Rajpurohit[3]

[1*]Department of Computer Science and Engineering, KLE DR MSSCET, Belagavi, India
[2,3]Department of Computer Science and Engineering, Gogte Institute of Technology, Belagavi, India

[*]*Corresponding Author: raghavendr023@gmail.com*

Available online at: www.ijcseonline.org

Received:22/Jan/2017 Revised: 03/Feb/2017 Accepted: 16/Feb/2017 Published: 31/Mar/2017

***Abstract*- Human soft tissues are diagnosed by different imaging modalities such as Computed tomography CT, Ultrasound US, Magnetic resonance imaging MRI all these imaging modalities are applied depending on the nature of the disease. In the classification of liver related diseases each of these imaging modalities plays important role. Classifying a liver into normal liver and diseased liver (in diseased cirrhotic or fatty liver) depends completely on the texture of the liver. Texture is a combination of repeated patterns with regular or irregular frequency. Texture visualization is easier but very difficult to describe in words. Analyzing liver texture is also difficult. To classify liver into its respective diseases category it is very important to extract the Region of Interest ROI accurately by segmentation, but as liver structure has maximum disparity in intensity texture inside and along boundary which serves as a major problem in its segmentation and classification. There are different textural analysis techniques developed for liver classification over the years some of which work equally well for most of the imaging modalities. Here, an attempt is made to summarize the importance of textural analysis techniques devised for different imaging modalities.**

Keywords* – *Cirrhotic, fatty liver, Ultrasound, Computed tomography, Magnetic resonance imaging, Texture Analysis, Liver Classification.

I. INTRODUCTION

Liver is an important organ of our body, which carries out vital functions. Some important functions of liver are metabolizing drugs, clearing toxins from the blood, and producing blood proteins and bile to aid digestion. Damage to the liver can happen due to several reasons such as alcoholism, obesity, viral hepatitis etc.; some disease would cause serious complications and may lead to liver transplant. If the liver disease is identified at earlier stage then there are chances of saving the liver from serious problems. In this regard many researches have been already made in an attempt to develop a Computer Aided Diagnostic system (CAD) which will help to diagnose the liver into normal liver or affected liver.

Liver images have granular structures called texture. Normal liver usually differs with the diseased one in terms of intensity texture. This variation helps in determining the corresponding disease. A CAD system is a merger of medical imaging and tissue characterization techniques, and is widely used in liver diagnosis. CAD systems are not going to replace the doctors but are going to be an assistant to doctor to provide a second view during diagnosis of the disease and support practitioners' judgment about disease. Basically, a CAD system involves the following steps such as i) Image acquisition ii) Image Pre-processing iii) Region of interest(ROI) selection which can be manual(ROI from liver image is selected by expert radiologist) or automatic(program itself selects ROI for a given liver image as an input) iv) Image segmentation(segmenting involves segmenting ROI from the liver image for feature extraction stage) v) Feature extraction(here Textural features of liver are extracted for analysis) vi) Finally, the classifier would classify the liver into predefined classes of diseases. Feature extraction and liver classification steps of CAD completely depend on segmentation accuracy. Incorrect segmentation would affect CAD systems classification of disease. In this regard, a discussion is made to summarize different techniques used for classifying liver diseases, and an idea has been proposed to classify the liver images into normal and affected liver (in affected category fatty liver, cirrhotic liver). There are several problems associated in choosing the imaging modality for liver images depending on the cost and accuracy for diagnosing. US imaging as shown in fig 1 which is cost effective and is widely used for diagnosing liver disease but on the other hand it is very difficult to analyze the texture features of the liver and then segmenting required ROI would be very difficult. It would require an expert radiologist to identify ROI by marking from the liver image manually and then segmenting the marked portion by the expert radiologist and further perform texture feature analysis on the ROI from segmented image and then classify the disease. CT imaging as shown in fig 2 provides a clear image of the liver from which segmenting and then analyzing the texture feature of the liver can be done in an effective way but CT imaging is costly comparatively with US imaging.

FIGURE 1: US image of Liver

FIGURE 2: CT image of liver

Texture features analysis techniques

Texture analysis is a quantitative method that can be used to quantify and detect structural abnormalities in different tissues. There are different approaches to texture analysis few of them are mentioned below,

- Structural or syntactic
- Statistical
- Model-based

Among the above mentioned categories few of the extensively used techniques [1] are illustrated below:

• Gray Level Difference Statistics (GLDS) or Gray Level Difference Matrix (GLDM):

GLDS is the Probability Density Function (PDF) of pair pixels lying at specific distance and having a particular intensity value difference. Inter pixel gray level values have large variation for fine texture and least variation for coarse texture.

• Spatial Gray level Dependence Matrices (SGLDM) or Gray Level Co occurrence Matrix (GLCM):

SGLDM exploits the fact that spatial relationship between gray levels of an image contributes to overall texture properties of the image. It computes matrix by counting how many times pixels with intensity i and j occur at specified offset.

• Gray level Run length Statistics (RUNL) or Dominant Gray Level Run Length Matrix texture measures (DGLRLM):

RUNL makes use of the fact that there are consecutive points in image having same gray level along a particular direction. Coarse texture contains relatively long runs than short runs. Opposite is true for fine texture.

• Gray Level Histogram:

It employs intensity distribution of image to find out texture parameters.

• Edge Frequency based Texture Features:

These features are inversely related to the autocorrelation function and are based on distance related gradient. Micro-edges and macro-edges can be detected using small and large distance operator respectively.

• First Order Parameters (FOP):

These are independent of spatial relation between pixels and describe only echogenicity and diffuse variation characteristics.

• Laws Texture Energy Measure (TEM):

TEM uses convolution masks of 5x5 to detect various texture types. It works on five basic 1D masks convolved to produce 25 2D masks. Texture image is then filtered with these masks to extract useful features.

• Fourier Power Spectrum (FPS):

This technique is useful for regular wave like patterns with a constant interval. Fourier transformation provides direction and frequency of pattern.

• Wavelet Transform Texture Measures:

These features are derived from wavelet transform of the image or Region of Interest (ROI). Major types are quincunx, Gabor and dyadic. Selection of appropriate textural features plays an important role in success of above mentioned textural analysis schemes. Some important textural features include: Entropy (ENT), Local Homogeneity (LH), Gray Level Distribution (GLD), Run Length Distribution (RLD), Angular Second Moment (ASM), Contrast (CO), Correlation (CORR), Variance (VAR), Inverse Difference Moment (IDM), Standard Deviation (SD), Energy (E), Homogeneity (H), Uniformity (U), Sum Entropy (SENT), Mean (M), Short Run Emphasis (SRE), and Dissimilarity (D).

II. LITERATURE REVIEW

Among the several liver imaging modalities, ultrasound is the least expensive but less precise tool for detecting liver abnormalities. Conversely, CT scan is the most reliable but at the same time costly method of diagnosing liver diseases. A lot of liver texture analysis techniques have been proposed in

the past with major focus on Ultrasound and CT imaging modalities.
Following text gives a brief discussion on different imaging modalities applied for analysis of liver diseases:

Liver diagnosis using CT imaging modality
A considerable percentage of liver texture analysis techniques are based on CT data. Many authors have put their efforts in developing a CAD system to diagnose the liver images.
S.Gr. Mougiakakou, I. Valavanis, K.S. Nikita, A. Nikita, and D. Kelekis [2] used combination of many feature extractors and Neural Network classifiers in the process of classification of liver CT images into normal, hepatic cyst, hemangioma, and HepatoCellular Carcinoma (HCC).Techniques like FOP, SGLDM, GLDS, TEM and FDTA were used for feature extraction from ROI. Genetic Algorithm (GA) was used to reduce dimensionality of feature vector. 5 feature set were considered as an input for 4 class NN and back propogation algorithm was used for training. Outputs of individual NN were combined by majority voting and weighted voting in order to decide liver class.

K. Mala, and V. Sadasivam [3] came up with an approach in which the CT image of the liver was automatically segmented and classified. Following steps were used Firstly morphological operations; secondly complementing the image followed by multiplying the complemented image with original one to segment liver. Thirdly horizontal, vertical and diagonal details were calculated with the help of orthogonal wavelet transform. Eighteen textural features were calculated such as M, SD, ASM, CO, E and ENT for the distance of 4 pixels. Finally, they used features for Probabilistic Neural Network (PNN) training and classification.

K.Mala, and Dr.V.Sadasivam proposed a method [4] for classifying malignant and benign tumor using CT image of liver by amalgamating biorthogonal wavelet transform with Linear Vector Quantization (LVQ) network. First preprocessing was done after that liver was segmented by multiplying and complementing the preprocessed image with the original image. Image was divided into liver, tumor, background by applying FCM. Steps involved are, in classification bi-orthogonal wavelet transform was used to extract horizontal, vertical and diagonal coefficients on tumor region. A Co-occurrence matrix was built in the second step that was also used for computing second order statistical texture features such as ASM, CO, H and ENT in horizontal, vertical and diagonal directions using pixel distance of 1. These features were used for training neural network.

V.S. Bharathi, M.A.L. Vijilious, and L.Ganesan, [5] used Zernike moments and Legendre moments for differentiating between normal and HCC liver using CT liver images. Around 200 ROI were used among these 140 belong to healthy liver class and 60 to HCC. Each ROI was segmented into multiple 8x8 segments. Among these 75 were used for training and the remaining for testing. Accuracy of classification using Zernike features was 92.37% with 5% noise, 85.50% for 10 % noise and 77.86% for 15 % noise. Accuracy in classification for normal liver was 98.60 % and 97.57%, whereas that for HCC was 90.00% and 80.25%.

S. Nawaz, and A.H. Dar, [6] proposed a method for classifying hemangioma, hepatoma, cirrhosis and normal liver using SVM. Liver segmentation from CT images was done by using snakes algorithm. Features were extracted using SGLDM. Further, feature extracted were given as input to a hierarchical SVM which first classifies the liver into diseased and non-diseased tissue, and then diseased image was then fed to another SVM that would classify disease into hemangioma or non-hemangioma. Non-hemangioma image was further used by final SVM to classify it between hepatoma and cirrhosis. 77% classification accuracy was achieved from the method.

Authors of [7] proposed an approach that automatically detects liver tumor in CT images by using region-growing and Support Vector Machine (SVM) which classifies the liver cancer types such as hepatoma, hemangioma and carcinoma. The average error rate and accuracy rate obtained was 0.02 and 0.9.

Liver diagnosis using Ultrasound imaging modality
Texture analysis of ultrasonic liver images is always a challenge for researchers. But, Many well-known liver texture analysis techniques, proposed by researchers, are based on ultrasound images for diagnosing the liver images.
A. Ahmadian, A. Mostafa, M.D. Abolhassani, and Y. Salimpour, proposed a method [8] for categorizing different liver diseases using Gabor wavelet texture feature extraction method and classified ultrasonic liver images into normal, cirrhosis and hepatitis groups. Features were extracted and images were classified into different categories using dyadic

wavelet transform, Gabor wavelet transform and statistical moments.

D. Balasubramanian, P. Srinivasan, and R. Gurupatham [9] came up with an automatic classifying system which classified liver into benign, malignant, cyst and normal liver images using texture features extracted through SGLDM, RUNL, TEM and Gabor wavelets techniques. Eight features were chosen from manually selected features and Principle Component Analysis (PCA) based optimal features. PCA based features and manually selected features were classified by K-means clustering algorithm and BPN respectively. Classification results of BPN were better than K-means.

Authors of [10] focused on improvement of the diagnostic accuracy of focal liver lesions by quantifying the key features of cysts, hemangiomas, and malignant lesions on ultrasound images. The focal liver lesions were divided into 29 cysts, 37 hemangiomas, and 33 malignancies. A total of 42 hybrid textural features that composed of 5 first order statistics, 18 gray level co-occurrence matrices, 18 Law's, and echogenicity were extracted. A total of 29 key features that were selected by principal component analysis were used as a set of inputs for a feed-forward neural network. For each lesion, the performance of the diagnosis was evaluated by using the positive predictive value, negative predictive value, sensitivity, specificity, and accuracy. The results of the experiment indicate that the proposed method exhibits great performance, a high diagnosis accuracy of over 96% among all focal liver lesion groups (cyst vs. hemangioma, cyst vs. malignant, and hemangioma vs. malignant) on ultrasound images. The accuracy was slightly increased when echogenicity was included in the optimal feature set.

Authors of [11] proposed a computer aided diagnosis system for understanding the diseases related to human liver using eleven statistical textural features extracted from US images of 14 fatty and 28 normal livers. Supervised classification was used. For supervised learning input sets were constructed in two alternative methods first, training set containing two third and test set one third of the data secondly, a representative training set was generated using Self Organizing Map where entire data set was treated as test set. Both the inputs were used with and without Principal Component Analysis. The analysis shows that the Multi Layer Perceptron with conventional data set without pre-processing yields better results as compared to other paradigms.

Virmani J, Kumar V, Kalra N, Khandelwal N [15] proposed computer aided diagnosing system using WPT texture descriptors and to extract mean, standard deviation, and energy features for categorizing liver into normal, cirrhotic and hepatocellular carcinoma (HCC). To distinguish among the three classes SVM classifier was used. 88.8% was the accuracy achieved by their method.

Minhas Fu, Sabih D, Hussain M [16] proposed a method for detecting fatty and heterogeneous livers. The required ROI was studied by applying WPT as textural analyzer to obtain the statistical features. For classification among the two predefined classes a multi-class linear SVM was applied. The overall accuracy of their approach led to an Acc of 95.4%

Andrade A, Silva JS, Santos J, Belo-Soares P. [17] came up with a semi-automatic classification approach to identify fatty and normal liver. Different textural analysis techniques were applied such as GLRLM, GLCM, and fractal dimension to extract features for three different classifiers. The accuracy obtained in the three classifiers is 76.92% for ANN, 79.77% for SVM and 74.05% for k-NN.

Huang Y, Han X, Tian X, Zhao Z, Zhao J, Hao D. [18] proposed a method to classify liver into normal and fatty liver. Texture features were analyzed by using GLCM, gray level histogram and GLDS. Images were subjected for de-noising and extracting different statistical features. Probabilistic neural network (PNN) was applied to classify the normal and fatty liver with accuracy rate of 82.5% and 87.5% for normal and fatty classes, respectively.

III. DISCUSSION

Different authors have used different texture feature extraction methods. In some cases usage of feature extraction methods were dependent on the type of liver images i.e. US or CT images. Many other features were also extracted to draw accurate results. Further, different classification algorithms were applied for classification of liver images into predefined classes of liver. Many authors have given the details regarding the rate of accuracy achieved by their proposed method. Some authors have taken help from the expert radiologist in segmenting the required ROI from liver image and few have came up with an automatic segmentation approach. It was found that the entire accuracy of the methods being proposed were dependent on the accuracy of segmenting ROI and texture feature extraction methods used.

Table 1 which gives the details of comparison of different feature extraction techniques, classifiers used applied and the accuracy achieved in US imaging modality by different authors. The table 1 gives details of the work carried out by only few authors.

IV. CONCLUSION

It has been observed that classification of liver diseases is more accurate in CT imaging modality than US imaging modality but at the same time is costlier. CT images of the liver provide a good basis for analyzing the texture of the liver, where as US images impose some difficulties to analyze the structure of the liver and then further analyzing texture is a challenge but US imaging is cost effective. Researches have been made in both the imaging modality to diagnose the diseases related to liver. In case of CT scan, techniques such as statistical moments provide good results and in case of ultrasound scan wavelet-based techniques for feature extraction show good accuracy according to study[14]. There is no perfect method for each of these imaging modalities. A lot can be improved in accuracy of diagnosing the liver diseases based on texture analysis. The accuracy completely depends on the ROI segmentation from liver image and then the texture features extracted and the classifiers applied. The results may improve by applying combination of texture feature extraction methods and classifiers.

TABLE 1: Comparison of techniques used and their results

	Feature Extraction Technique	Classifier(s)	Accuracy Achieved
[14]	WPT	SVM	88.8%
[15]	WPT	Multi-class linear SVM	95.4%
[16]	GLRLM, GLCM, and Fractal Dimension	ANN	76.92%
		SVM	79.77%
		KNN	74.05%
[17]	GLCM, gray level histogram and GLDS	PNN	87.5%

V. FUTURE WORK

Though authors have done a lot of investigation, there is a lot more to be done. A hybrid CAD system can be developed which provides the flexibility for the doctor for choosing the imaging modality depending on the type of the image doctor has for diagnosis for example if the doctor has to diagnose US image then the doctor will have to choose US image for giving input to the CAD system, if doctor has CT image of the liver then doctor will have to choose CT imaging modality for giving the input similarly for MRI, then the CAD system will work accordingly depending on the type of imaging modality chosen by the doctor for instance if the doctor has chosen US imaging then US texture feature extraction techniques would work in the background and then diagnosis would be done and classification of liver disease would be brought as output from CAD system similarly for remaining imaging modalities. Accuracy further can be improved by applying different feature extraction methods and classifiers for the same data set. A lot of improvements can be done in diagnosing any diseases but CAD system will not replace doctors.

REFERENCES

[1] Rathore S, Iftikhar M, Hussain M, Jalil A, *"Texture analysis for liver segmentation and classification: a survey"*, Frontiers of Information Technology (FIT), Islamabad, pp.121–126, 2011, Print ISBN 978-1-4673-0209-8

[2] Mougiakakou S G, Valavanis I, Konstantina S N, Alexandra N and D. Kelekis., *"Characterization of CT liver lesions based on texture features and a multiple Neural Network classification scheme"*, Engineering in Medicine and Biology Society, Mexico, pp.1287 - 1290, 2003, Print ISBN 0-7803-7789-3.

[3] Mala K, Sadasivam V., *"Automatic segmentation and classification of diffused liver diseases using wavelet based texture analysis and Neural Network"*, INDICON 2005 Annual IEEE, NA, pp. 216-219, 2005, Print ISBN 0-7803-9503-4.

[4] Mala K, Sadasivam V., *"Wavelet based texture analysis of Liver tumor from Computed Tomography images for characterization using Linear Vector Quantization Neural Network"*, Advanced Computing and Communications International conference, Surathkal, pp. 267 – 270, 2006, E-ISBN: 1-4244-0716-8.

[5] Subbiah V. B, Vijilious M.A.Leo, and Ganesan L, *"Orthogonal Moments based texture analysis of CT liver images"*, International Conference on Computational Intelligence and Multimedia Applications, Tamilnadu-India, pp.249-253, 2007, Print ISBN: **0-**7695-3050-8

[6] Nawaz S, and Dar A.H., *"Hepatic lesions classification by ensemble of SVMs using statistical features based on co-occurrence matrix"*, International Conference on Emerging Technologies, IEEE-ICET- 2008, Rawalpindi-Pakistan, pp.21-26, 2008, Print ISBN: 978-1-4244-2210-4

[7] Sreeraj R, Raju. G *"Automatic Detection of Liver Tumor in CT Image Using Region Growing and SVM Classifier"* IJCSE Volume-4, Issue-11, 2016

[8] Ahmadian A, Mostafa A, Abolhassani M.D, and Salimpour Y, *"A texture classification method for diffused liver diseases using Gabor wavelets"*, 27th Annual International Conference of the Engineering in Medicine and Biology Society, IEEE-EMBS 2005- Shanghai- China, pp. 1567 – 1570, 2005, Print ISBN: 0-7803-8741-4.

[9] Balasubramanian D, Srinivasan P and Gurupatham R, *"Automatic classification of focal lesions in ultrasound liver images using Principal Component Analysis and Neural Networks"*, 29th Annual International Conference of the IEEE Engineering and Medicine and Biology Society- 2007, Lyon-France, pp.2134-2137, 2007, Print ISBN: 978-1-4244-0787-3.

[10] Hwanga Y N, Lee JH, Kim GY, Jiang YY and Kim SM *"Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network"*, Bio-Medical Materials and Engineering, Shanghai-China , pp. S1599–S1611, 2015.

[11] Neogi N, Adhikari A, Roy M *"Classification of Ultrasonography Images of Human Fatty and Normal Livers using GLCM Textural Features"*, Current Trends in Technology and Science, Vol 3(4), pp.252-259, 2014

[12] Haralik R. M, and Shanmuygam K, *"Textural features for image classification"*, IEEE Transactions on systems, Man, Cybernetics., Vol. SMC-3(6), pp. 610-621, 1973.

[13] Luo S, Qingmao H, Xiangjian H, Jiaming L, Jesse S.Jin, Mira P , *"Automatic liver parenchyma segmentation from abdominal CT images using Support Vector Machines"* , ICME International Conference on Complex Medical Engineering(2009), Tempe-USA, pp.1-5, 2009

[14] Cao G, Shi P, and Hu B, *"Liver fibrosis identification based on ultrasonic images"*, Engineering in Medicine and Biology Society, 2005, Shangai, pp.6317 – 6320, 2005.

[15] Virmani J, Kumar V, Kalra N, Khandelwal N. *"SVM-based characterization of liver ultrasound images using wavelet packet texture descriptors"*, Jounal of Digit Imaging, Vol 26(3), pp.530-543, 2013

[16] Minhas F, Sabih D, Hussain M. *"Automated classification of liver disorders using ultrasound images"*, Journal of Medical Systems, Vol 36(5), pp.3163–3172, 2012.

[17] Andrade A, Silva JS, Santos J, Belo-Soares P. *"Classifier approaches for liver steatosis using ultrasound images"*, Procedia Technology, Vol 5, pp. 763-770,2012.

[18] Huang Y, Han X, Tian X, Zhao Z, Zhao J, Hao D. *"Texture analysis of ultrasonic liver images based on spatial domain methods"*, Image and Signal Processing (CISP)2010, Yantai-China, pp. 562–565, 2010

