Liver tumour classification using average correction higher order local autocorrelation coefficient and legendre moments

Aravinda H.L.*, M.V. Sudhamani2

*Research Scholar, Jain University, Asst. Professor, Department of Telecommunication Engg., Dr.Ambedkar Institute of Technology, Bengaluru, India.
2Professor and Head, Department of Information Science and Engg., RNS Institute of Technology, Bengaluru, India.
*Email: arvindhlait@gmail.com

Abstract

The major reasons for liver carcinoma are cirrhosis and hepatitis. In order to identify carcinoma in the liver abdominal CT images are used. From abdominal CT images, segmentation of liver portion using adaptive region growing, tumor segmentation from extracted liver using Simple Linear Iterative Clustering is already implemented. In this paper, classification of tumors as benign or malignant is accomplished using Rough-set classifier based on texture feature extracted using Average Correction Higher Order Local Autocorrelation Coefficients and Legendre moments. Classification accuracy achieved in proposed scheme is 90%. The results obtained are promising and have been compared with existing methods.

Keywords: Liver tumors, Average Correction Higher order Local Autocorrelation Coefficient (ACHLAC), Legendre Moments (LM)

1. Introduction

The liver has vital role as a part of human body. Liver cancer may lead to the death of human being. Several factors play role in developing the liver cancer. Cancer cells from other parts can accumulate in the liver as the liver functionality is to filter the blood from all body parts; it starts growing and spreads in the liver. Hepatocellular Carcinoma includes major risk factors like chronic hepatitis virus infection, especially, Hepatitis B and C. Cirrhosis is due to hepatitis or alcoholism. The liver cancer diagnosis occurs at the final stages of the diseases when very less number of treatment options are available and there is very poor prognosis of patients in cases of HCC.

Different imaging techniques like MRI, PET and CT are standards used for liver cancer diagnosis, liver pathologies and fulminant hepatic failure. But most of the diagnosticians prefer to use CT images as they give better anatomical information regarding the structures which have been visualized, they have better spatial resolution and high signal to noise ratio. The current interest is as follows: automatic liver cancer detection or other disease with respect to liver; transplantation of liver; its volume measurement. All these kinds of research goals require handling fundamental problem that is liver volume segmentation from CT data. This task is a usually done by expert radiologist who will be tracing the liver contour manually on the CT slice. Manual segmentation of liver cancer is tedious and time consuming. Thus development of semi-automated and automatic techniques using image processing techniques has gained demand. There are many advantages of automated methods that overcome manual or interactive techniques. Human presence is not required in automated techniques and it is also faster. Internal organs are analyzed using CT images. Automatic segmentation is challenging because of indefinite shape and similarities in the intensity between the tumor and other neighborhood tissues.

2. Literature Review

The major classification of the tumor is benign or malignant. Benign is relatively harmless and malignant has ability to spread from liver and is more serious.

Amitha et al [11] used Markov Random Field (MRF) for liver tumor segmentation from CT images. First, liver is required to segment by the MRF embedded level set method. It is fast segmentation and is robust to noise. The ambiguities with respect to the shape of the segmented liver have to be found by techniques used for the shape analysis. Here training set is used for correction. Hepatic tumors are identified by using graph cut technique from the corrected liver segmentation. Then features are extracted using SVM.

Masuda, Yu, et al [12] makes use of EM algorithm for detection of liver tumor. The first step is preprocessing which eliminates the noise. The EM algorithm clusters the liver tissue and shape informatics is used to eliminate the false positives.

Militzer, Arne, et al [13] detects and segments the liver tumor from CT contrast enhanced image. Probabilistic boosting tree is utilized to classify of the liver to parenchyma or lesion. This method is fully automatic which detects and also segments the tumor at a time. Iterative classification is integrated for robust segmentation and also incorporates the knowledge of initial iterations at the later decisions.

G Ismail et al [14] uses hybrid approach of combining fuzzy cluster and grey wolf optimization (GWO) for automatic segmentation of liver. Fast fuzzy c means is utilized for lesion candidate clustering and different features of each candidate are used for classification. GWO helps us in finding optimal values for threshold and improve the results of clustering that were obtained from FCM. The centroids value obtained from FCM are used as initial positions of wolf in GWO technique.
N Ghatway et al [15] uses fractional differential for liver segmentation of enhanced CT images. This is objective of improving the texture and edge features. Than fusion is applied to combine different enhanced images and improves different features so far obtained. Each image is partitioned into non-overlapping blocks of size N x N so that require feature is extracted. At last blocks are classified into tumor and non-tumor by using SVM classifier.

Dirk Smeets et al [16] uses level set method for accurate and reliable segmentation. The initial step is spiral scanning technique which is based on dynamic programming. The evaluation of level set is dependent on speed image, i.e., statistical pixel classification with supervised learning.

3. Methodology

In this work, we are attempting to extract the liver tumor from CT images and classify them into tumor and non-tumor, tumor are further classified into benign and malignant. The following are the steps of liver tumor detection and classification: Image Enhancement, Liver Segmentation, Predicted tumor region identification, tumor validation and tumor classification. Input image is considered to be CT image, because of noise, there is lack proper distribution of intensity and the segmentation becomes difficult. Hence image enhancement is done by applying threshold and top-hat algorithm. Adaptive Region Growing method is applied to segment the liver region.

![Fig 1: Block Diagram of Proposed System](image)

3.1. Region Growing

Region growing is being used also as a learning process for determining homogeneity parameters. This greatly reduces unwanted interactions. The adaptive procedure needs two runs of growing, since these homogeneity parameters are subjected to change during learning process. The first run includes estimation of homogeneity parameters. In the second run these parameters are implemented by making use of the same seed point, for segmenting the region.

Mean and two different standard deviations needs to be computed for gray values of a numerous pixels of the region to learn the homogeneity criterion. The mean value is required to be computed using the median rather the average value of pixels, because the union of two halves of two unique Gaussian functions is never a Gaussian. By independently assessing the gray values of the test pixels which are greater or lesser than the mean, the two standard deviation values can be computed. The present estimation of homogeneity must be applied weakly for fixing membership of region which eliminates early stoppage of the process, during early period of learning. Upper and lower thresholds are calculated using Eq. (1) and Eq. (2) respectively.

\[
T_{\text{upper}} = mgv(n) + [ud(n), w + c(n)]
\]

\[
T_{\text{lower}} = mgv(n) - [ld(n), w + c(n)]
\]

Where, n-number of pixels employed to determine the estimate. Function c(n) is used to compensate erroneousness of the estimates of mgv, ud and ld. This function c(n) is inversely proportion to value i.e. as the value of Fincreases the value of function decreases. In this manner the adaptive region growing algorithm segments the liver part in an efficient manner.

3.2. SLIC Method

In this work, Simple Linear Iterative Clustering algorithm employs cluster of pixels (group of pixel having similar characteristics) called superpixels k which are equal in size. SLIC is a simple algorithm helpful for segmenting image. The images of CIE LAB (Commission Internationale de l’Eclairage) color space represent all colors based on human view are considered in proposed work. SLIC performs clustering by exploiting (x, y) coordinates of pixels and by using CIE LAB color space values in five dimensional spaces. Pixels are clustered based on their color similarity in five dimensional [labxy] spaces to generate superpixels. SLIC receives equal size of k super pixels as input and each superpixel consists of N/K pixels. Thus equal sizes of superpixel are represented as S = \(\frac{N}{k}\). Pixels associated with superpixel cluster lie within 25x25 area centered in (x, y) plane.

For small pixel distance euclidean distance works well in CIE LAB color space, but when pixel distance exceed then euclidean distance is not helpful. Hence for large spatial pixel distance measure \(D_3\) is employed. \(D_3\) represent sum of lab distance in color space and (x, y) coordinate distance. \(D_2\) Identify cluster center nearest to pixel. SLIC algorithm proceeds by sampling pixel clusters and moving the clusters into 3x3 neighborhoods, to avoid probability of selecting distorted pixel. Every pixel i assigned to the cluster whose center is nearest and search region will overlaps the location. This speeds up the algorithm as size limitation of the search region will decrease the amount of distance calculations, and will also result in improvement in speed which is advantage over k-means clustering in which each pixel is compared with other cluster centers.

This is done by introducing the distance measure \(D_3\), expected spatial extent of super pixel is considered to be of region size, the similar pixels are searched in region of size S in the surrounding of super pixel center. As each pixel is assigned to a cluster, the cluster centers are adjusted to the mean vector [labxy] where all pixels are assigned to the cluster. The residual error E is calculated among centers of new cluster and previous cluster locations using L2 norm. The assignment step and updating steps are recursively iterated till error converges. Last step is post-processing in which reassigning of disjoint pixels with nearest super pixels establishes connectivity again.

3.3. Gray-Level Co-Occurrence Matrix (GLCM)

GLCM gives the information regarding the various pixel intensities occurrence in a gray scale image. It is one of the statistical approach in which spatial relationship of the pixels is explored. GLCM computes in the way pixel intensity occurred in horizontal, vertical and diagonal directions to the pixel intensity j. Certain properties with respect to gray intensity spatial relationship in the image are exhibited by the GLCM. The four values considered for \(\theta\) are 0°, 45°, 90°, and 135°.

![Fig 2: Formation of GLCM](image)
These values cover the orientations and common distance length $d = 1$ that different values of $\theta$. For an image I of size $K \times K$, the gray level co-occurrence matrix for vector d is defined in Eq. (3)

$$M_{ed} = \sum_{x=1}^{K} \sum_{y=1}^{K} \begin{cases} 1 & \text{if } I(x, y) = i \text{ and } I(x + d, y + d) = j \\ 0 & \text{otherwise} \end{cases}$$

(3)

3.4. Histogram Features

The histogram features are extracted by below steps. On the channel of intensity I, image pyramid $I(c)[c = 0, 1, 2, 3]$ is created using spatial scales which ranges from 1 to 1:8 in four octaves. In every scale I(c), histogram is computed. Gradient filter [-1, 0, 1] is used to calculate the horizontal $Gx(x, y)$ and vertical $Gy(x, y)$ gradient of $I(C)$. Magnitude $|G(x, y)|$ and $\theta(x, y)$ angle is computed using by following equation:

$$|G(x, y)| = \sqrt{Gx(x, y)^2 + Gy(x, y)^2}$$

(4)

$$\theta(x, y) = \arctan \frac{Gx(x, y)}{Gy(x, y)}$$

(5)

Histogram is computed using b orientation bins in $0^\circ$ to $180^\circ$. Magnitude $|G(x, y)|$ having angle $\theta(x, y)$ related to similar bin will get added as this bin value. The histogram is normalized using L2 norm.

3.5. Hamming Distance

The curse of dimensionality contains an express manner on resemblance penetrating in high proportions in sense that can elevates the concern of whether or not adjacent neighbor penetrating is still significant in such a domain. One cause is most of us that are not mainly proficient at imagine the high-dimensional information. In particular, d indicate a distance purpose that require not essentially be a metric, the adjacent neighbor penetrating is not significant when the relation of difference of distance between two random points $p$ and $q$, then drawn as of the information and query distribution, and usual distance among them meets to zero as the aspect $k$ goes to infinity, limit $k \rightarrow \infty$. Variance $[d (p, q)]$ Expected $[d (p, q)] \approx 0$. Otherwise, the distance is to nearest neighbor and distance to furthest neighbor lean to meet as the dimension enlarges.

3.6. ACHLAC

The traditional HLAC cannot reflect image information comprehensively. For example, suppose in the template, the gray values of neighboring pixels are 2 and 125, 15 and 16, and result of multiplication are 250 and 240 respectively. The result of multiplication is similar while the distribution of gray pixel value is distinct. The former local variance is greater than later local variance. The abnormal liver pathological local variance is high and hence formal has higher chances of recognizing abnormal. Difference between former and later HLAC is recognized much. So while doing direct multiplication, there is chance of losing some information, therefore, average correction HLAC (ACHLAC) is proposed. In this method, average gray value is calculated for whole image. Absolute value of difference between current pixel and average pixel is computed for each pixel and current pixel is updated. The updated image may reflect better gray value variation. The HLAC features are calculated with respected to the updated image. Only local information is considered by the traditional HLAC while improved ACHLAC considers both local gray value and average gray value which reflects the local information and also shows local gray value changes.

From image size, differences result calculation is avoided by normalizing process for extracting the ACHLAC features. First calculate the max $x_1, x_2, x_3, ..., x_{25}$. Then divide every element by the max to normalize the values in the range [0, 1]. Here the order 2 is considered and offset limit is considered to be 3x3 region. The ACHLAC pseudo code is given below:

**Algorithm 1: ACHLAC Feature Extraction**

**Input:** Liver tumor image

**Output:** ACHLAC features: $x_1, x_2, x_3, ..., x_{25}$

Step 1: Traverse liver image pixels and calculate the average of gray pixel value.

Step 2: Traverse liver image pixels and calculate the absolute difference between every pixel and average value, then value of every pixel is updated.

Step 3: After updation, traverse each pixel.

Step 4: For $i = 1:25$

Step 5: Using template $P_i$, match local autocorrelation. The value $x_i$ is calculated and updated

Step 6: End for

Step 7: Maximum value is calculated: $(x_1, x_2, ..., x_{25})$

Step 8: For $i = 1:25$

Step 9: $x_i = \frac{x_i}{\max}$

Step 10: End for

3.7. Legendre Moments

Legendre moments are of type continuous orthogonal moments which are dependent on Legendre polynomial. For a digital image of size $N \times N$, Legendre moments is defined based on the orthogonality principles given in Eq. (6)

$$L_{pq} = \frac{(2p + 1)(2q + 1)}{(N - 1)(N - 1)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_p(x) P_q(y) f(x, y)$$

(6)

Where $P_p(x)$ and $P_q(y)$ denotes the order $p$ and $q$ Legendre polynomials respectively. Normalized coordinates are denoted by $x_i$ and $y_j$ which ranges within [-1, 1]. The transformation of image coordinates into is done as follows,

$$x_i = \frac{2x_i}{N} - 1;$$

(7)

$$y_j = \frac{2y_j}{N} - 1;$$

(8)

With defining Legendre polynomial as,

$$P_p(x_i) = \sum_{k=0}^{p} \frac{(-1)^{(p-k)}}{2^p} \binom{p}{k} x_i^k$$

(9)

Where $|x_i| \leq 1$ and $(p-k)$ are even.

3.8. Rough set based classifier

Rough sets are observed as a novel arithmetical advance to imprecision. In calculation, the viewpoint of rough set can found on statement that each purpose of universe of conversation is connected by several data. The equation of theory defines that: Eq. (10) depicts Lower approximation, Eq. (11) depicts upper approximation, and Eq. (12) depicts boundary region. The set $X$ represents that rough with respect to information in B if boundary region were non-empty.

In Eq. (13), if $\mu_B(A) = 1$, it is definable; otherwise it is indefinable, where $0 \leq \mu_B(A) \leq 1$.
\[ B_1 = \{ x| x \in X \} \]
\[ \overline{B}_1 = \{ x| x \not\in B \} \]
\[ BN_B(x) = B_1 - \overline{B}_1 \]
\[ \mu_B(A) = \frac{\text{card}(B_1)}{\text{card}(U)} \]

Rough sets process is utilized to analyze the system during a sequence of logical supposition processes. This can be considered as a choice table indicated by IS = (U, A, C, D), in which U is universe of discourse; A is a set of primitive attributes (characters or variables).

C, D ⊂ A are two subsets of features, supposing that A = C \cup D and C \cap D = \emptyset, where C represents the condition quality and D represents the decision attribute. The evaluate describing the imprecision of estimate classification known as the worth of imprecision of estimate classification known as the worth of D

The classification accuracy can be defined as below in Eq. (14):

\[ \gamma_B(A) = \frac{\text{card}(B_1)}{\text{card}(U)} \]

4. Experimental Results

The results of above proposed work is presented in this section. CT liver image is taken as input images is shown in Fig.3 (a) and Fig.4 (a). It can be observed that in figure 3 (a) input image is abdominal image of liver (Haemangiomata) then applying adaptive region growing algorithm image for segmentation liver extracted image is obtained in Fig 3(b). Similarly in figure 4, input image is abdominal image of (Hepatocellular Carcinoma) shown in Fig 4(a) which is segmented using adaptive region growing algorithm to obtain liver extracted image shown in Fig 4(b). The preprocessing step is done by using the median filter so we get the enhanced image depicted in Fig 3(c) and Fig 4(c). SLIC method will segment the tumor part of the liver is show in figure 3 (d), (e), (f), (g), and in figure 4 (d), (e).

Then apply the feature extraction techniques to extract the features from segmented image. Methods used for feature extraction are ACHLAC and Legendre Moments. After extracting the features apply classification technique that is rough set classifier to classify the type of tumor. Figure 3(l) and Figure 4 (j) shows that tumor is classified as Benign and Malignant.

The following parameters are considered for the liver classification when the efficiency is evaluated. Mathematical formula of accuracy, recall, specificity and precision is given below in terms of \( T_p, T_N, F_p, F_N \). Where \( T_p \) indicate positive samples classified as positive. \( T_N \) shows negative samples classified as negative. \( F_p \) defines negative samples classified as positive. \( F_N \) represents positive samples classified as negative. Precision is the ratio of number of positive samples correctly classified to the total number of Samples in a class. Recall, is the ratio of positive samples correctly classified to the total number of samples classified as positive. Accuracy obtained for proposed method is 90%. Sensitivity 90% with Specificity 88%

\[ \text{Sensitivity} = \frac{T_P}{T_P + F_N} \]
\[ \text{Specificity} = \frac{T_N}{T_N + F_P} \]
\[ \text{Accuracy} = \frac{T_N + T_P}{T_N + F_N + F_P + T_P} \]

Table 1 represents the comparison table for accuracy of existing and proposed system and Figure 5 depicts the comparison graph for the accuracy of proposed and existing system. This will represents that our proposed system is improved than the previous methods.

```
<table>
<thead>
<tr>
<th>S.No</th>
<th>Paper</th>
<th>Classifier</th>
<th>Method</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ch. Srinivasa Rao et.al [17]</td>
<td>Support Vector Machine</td>
<td>Legendre Moments</td>
<td>80%</td>
</tr>
<tr>
<td>2</td>
<td>R. Bock et.al [18]</td>
<td>Support Vector Machine</td>
<td>Principal Component Analysis</td>
<td>80%</td>
</tr>
<tr>
<td>4</td>
<td>Abdolhossein Sarrafzadeh et.al [19]</td>
<td>Support Vector Machine</td>
<td>Exact Legendre Moments</td>
<td>81.71%</td>
</tr>
<tr>
<td>5</td>
<td>Proposed Method</td>
<td>Rough set Classifier</td>
<td>Legendre Moments and ACHLAC</td>
<td>90%</td>
</tr>
</tbody>
</table>
```

The classification Accuracy defines the total number of samples correctly classified to the total number of samples classified. Results are demonstrated by considering 120 samples of CT images with \( T_P, T_N \) score are 101 and 8. \( F_p \) is 1, \( F_N \) is 10
Using classification accuracy formula from confusion matrix given in Eq. (17) accuracy, sensitivity is calculated.

![Performance Analysis Graph]

**Fig. 5:** Performance Analysis Graph

### 5. Conclusion

This paper presents our work on tumor segmentation using Simple Linear Iterative Clustering method and classification of tumours into benign and malignant based on rough-set classifier using texture feature extracted from Auto Correction Higher Order Local Auto Correlation Coefficient technique and Legendre moments. The proposed method has been implemented in the working platform MATLAB R2010a. The results obtained are discussed in the paper and we could achieve 90% of accuracy in classification. This has been compared with existing method and we have got the improvement in classification by 10%. In future we can extend the same method for classifying more number of disease classes.

### References


[22] Aravinda H.L and Dr.M.V.Sudhamani, “Simple Linear Iterative Clustering based Tumour Segmentation in Liver region of Abdominal CT-scan” is presented in International IEEE Conference on Recent Advances in Electronics and Communication Technology, held during 16-17 March 2017 at SJBIT, Bangalore.