

Performance Analysis of Classification of Liver Tumors using Support Vector Machine and Rough Set based Classifiers

Aravinda H L, Sudhamani M V

Abstract: In recent years the medical diagnosis is majorly done based on the medical images captured using various imaging modalities. The medical doctors and radiologists use these medical images to identify the pathological problems or diseases and suggest the patient about further treatment. In this process, medical doctors and radiologists often prefer to make use of software which can assist in taking the decision. Such an approach is called as Computer Aided Diagnosis (CAD). The CAD systems usually comprise of many phases like segmentation of portion corresponding to a particular organ or region under consideration, detecting the pathology bearing area in that and further classifying into various disease classes. Here, Accuracy of classifiers majorly decides the effectiveness of the diagnosis. In this paper, classification of liver tumors into benign and malignant is considered as a case study. Implementation of two different classifiers namely Support Vector Machine and Rough Set based classifier is carried out using set of features extracted such as Texture features using Average Correction Higher order Local Autocorrelation Coefficients and shape features using Legendre moments. Comparison of performance of both the classifiers is made and tabulated. Here, Rough Set based classifier has performed better when compared with Support Vector Machine.

Keywords – Liver tumor, Average Correction Higher order Local Autocorrelation Coefficients, Legendre Moments, Support Vector Machine Classifier, Rough Set based classifier

I. INTRODUCTION

The design of Computer Aided Diagnosis (CAD) system has become an emerging topic of research since it is playing a vital role in identifying the disease. Most of the CAD systems use the medical images captured using various medical imaging modalities [1]. Medical images give pathological details of affected region which helps doctors or radiologists for diagnosis.

In general to categorize the images the Classifiers are used. The classifiers are modeled and are trained such that they will be able to distinguish the objects based on the features extracted. Such classifiers are used in the CAD systems to distinguish various diseases.

Revised Manuscript Received on December 12, 2019.

* **Mr. Aravinda H L**, Research Scholar, Jain (Deemed to be University), Asst. Professor, Department of Telecommunication Engg., Dr. Ambedkar Institute of Technology, Bengaluru, India. arvindhlait@gmail.com

Dr. M V Sudhamani Professor & HoD, Department of Information Science & Engg., RNS Institute of Technology, Bengaluru, India. mvsudha_raj@hotmail.com

In human body, liver is one of the vital organs and tumors or lesions, both cancerous and non-cancerous can develop in the liver. Some of the types of benign lesions are Hepatic cyst, Haemangioma, Hepatocellular Adenoma and Focal Nodular Hyperplasia. Hepatocellular Carcinoma, Portal vein thrombosis and Cholangio carcinoma are the types of malignant lesions [3, 4].

This paper is organized as follows: section I gives the introduction, section II provides review of literature, section III explains the methodologies used, section IV provides experimental results and contribution is concluded in the V section.

II. LITERATURE SURVEY

There are various classification methods such as Support Vector Machine, Convolution Neural Network based, Fuzzy c-means clustering etc.,. Survey of classification methods employed for categorization of tumors into benign and malignant are considered here.

In [5], Back Propagation Neural Network approach is discussed and a classification accuracy of 73.23% is achieved. Data set with 583 abdominal CT images is considered. In paper [6], a Multi-Layer Perception classifier has been proposed.

In [7], classifier using Multilayer Perceptrons is implemented and an accuracy of 89.15% is achieved.

In [8], authors have presented their work on Convolutional Neural Networks and transfer learning has been employed for classification. 147 ultrasound images are considered and an accuracy of 90.6 % is found.

In [9], authors have employed SVM classifier in their work and have achieved an accuracy of 92.5 %. Use of K-nearest Neighbor (KNN) Classifier for medical image retrieval is discussed in [10].

III. METHODOLOGY

In this work, two classifiers namely Support Vector Machine classifier and Rough Set based Classifier are implemented for verification of their appropriateness in classification of liver tumors.



Feature extraction from the identified liver tumors is done using Average Correction Higher order Local Autocorrelation Coefficient (ACHLAC) and Legendre moments. The ACHLAC extracts texture features and Legendre moments extract shape features. Both the classifiers implemented here are designed to use the same features.

Architecture of classification and decision making is shown in Figure 1. The entire process for classification is divided into training and testing phases. In this work, the abdominal CT images are taken from publicly available repository [3, 4] and are detected tumors in the liver portion

of abdominal CT are considered. All the features extracted from the tumors are stored in repository along with ground truth images. In training phase, the input images are taken as 50 x 50 cropped images of the classified tumors and the tumors identified by the doctor or radiologists are considered as ground truth. By learning both the features, a training model is generated and fed into the Classifiers. The patches of tumors are given as input to classification phase by a doctor or radiologist. Here, based on trained data, test data is classified as either benign or malignant.

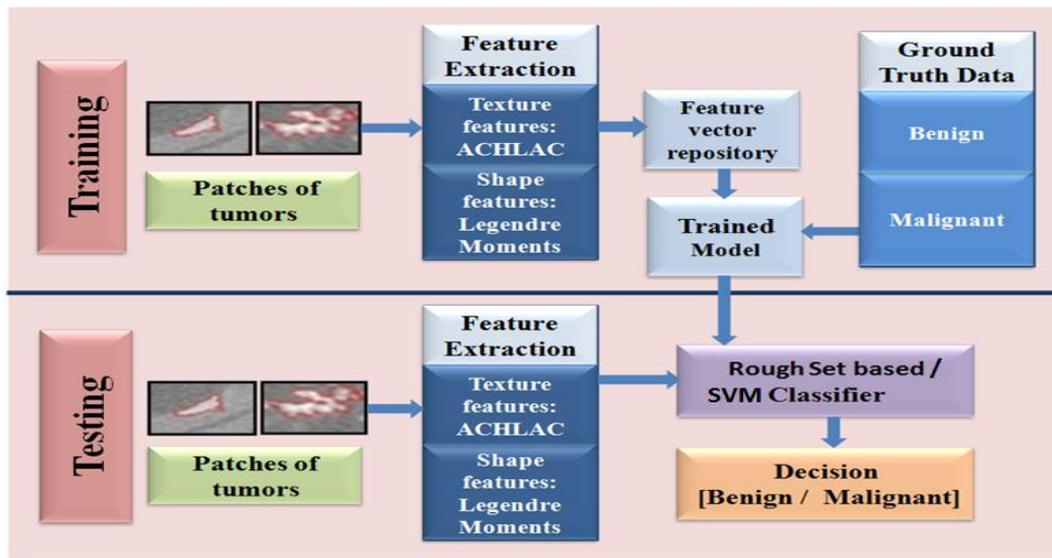


Figure 1. Architecture of the proposed system

3.3 ACHLAC

Higher order Local Autocorrelation Coefficients (HLAC) are used here to extract higher-order statistical features of images by a process called as template matching. The HLAC features have the property of spatial invariance and have shown good performance in recognition of objects in an image [11].

In conventional HLAC, gray levels in the image are multiplied directly. This can lead to the possibility that two pairs of pixels, with varied correlation resulting in the same feature vector. Due to this reason, conventional HLAC cannot indicate the details of the image comprehensively. To mitigate this problem, Auto Correction Higher order Local Autocorrelation Coefficients (ACHLAC) approach is employed.

In ACHLAC method, first the average gray value of the entire region of a tumor is calculated. Later, for every pixel, the magnitude of the difference between gray value that particular pixel and average gray value is calculated, resultant image will be formed by replacing each pixel with the corresponding difference values. Further, conventional HLAC is applied on the resultant image. The ACHLAC method provides more comprehensive information about the tumor in the form of texture features.

3.4 Legendre Moments

The moment of an image is a weighted average of intensities of image pixels. The Legendre moments are continuous orthogonal moments and are normally use in extracting the shape features. The Legendre moments for a digital image of size $N \times N$ are as given by equation (1).

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

where $P_p(x_i)$ and $P_q(y_j)$ denote the order p and q Legendre polynomials respectively. Normalized coordinates are denoted by x_i and y_j will be in the range of $[-1, 1]$. The transformation of image coordinates into normalized coordinates is given in equation (2) and (3),

$$x_i = \left(\frac{2x}{N}\right) - 1; \quad (2)$$

$$y_j = \left(\frac{2y}{N}\right) - 1; \quad (3)$$

The Legendre polynomial is defined as in equation (4),

$$P_p(x_i) = \sum_{k=0}^p \frac{1}{2^p} \frac{(-1)^{\frac{p-1}{2}} (p+k)! x_i^k}{k! (\frac{p+k}{2})! (\frac{p-k}{2})!} \quad (4)$$

where $|x_i| \leq 1$ and $(p - k)$ is even.

The shape features extracted from Legendre moments are used here for classification along with texture features.

3.5 Rough Set based classifier

The Rough sets offer an effective approach of managing uncertainties, reducing redundancies and can be used for the purpose of feature space dimensionality reduction by selecting only the relevant features. They are also used for recognizing and classifying objects. The Rough Sets are usually used for classification of various features which in turn classify the objects or tumors. The Rough sets require no separate parameters and make use of the information present in image itself.

3.6 Support Vector Machine

Support vector machine is a classification technique which learns by examples to assign labels to objects. It is a supervised learning method to classify the objects based on certain feature vectors. It is a binary classifier. A support vector machine constructs a hyperplane which is used for classification, regression and other processes.

A good separation is achieved by the hyperplane that has the largest distance to the nearest training data points of any class.

IV. EXPERIMENTAL RESULTS

The performance of classification is dependent on how many of the benign tumors are properly classified as benign and how many of the malignant tumors are properly classified as malignant. The performance of classifier can be computed from the Confusion matrix. Various performance measures can be computed from the Confusion matrix. Considering Benign and Malignant as two classes and prediction of malignant tumor as Positive (presence of cancer) and prediction of benign tumor as negative (absence of cancer), Confusion Matrix is shown as in Table I:

TABLE I: Confusion Matrix

	Predicted by classifier as Malignant	Predicted by classifier as Benign
Malignant Tumor	True Positive (TP)	False Negative (FN)
Benign Tumor	False Positive (FP)	True Negative (TN)

Based on the confusion matrix, several parameters such as sensitivity, specificity and accuracy can be computed using confusion matrix.

Sensitivity of the classifier can be defined as ability to correctly identify the cases with cancer disease. Sensitivity is mathematically expressed as given in equation (5).

$$\text{sensitivity} = \frac{TP}{TP+FN} \quad (5)$$

Specificity is defined as the ability of the system to correctly identify the cases without diseases and is expressed in equation (6).

$$\text{specificity} = \frac{TN}{TN+FP} \quad (6)$$

Accuracy of classifier is the ability to correctly differentiate both the classes and is defined in equation (7).

$$\text{accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (7)$$

For binary classification, accuracy can be computed as expressed equation (8).

$$\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Experimentation to evaluate both the classifiers was performed on 50 images among which 25 are benign and 25 are malignant. Table II shows the results of Rough Set based classifier and Table III shows the results of SVM classifier.

Table II: Results of Rough Set based Classifier

Ground Truth	Predicted by classifier as Malignant	Predicted by classifier as Benign
Number of Malignant tumors 25	23	02
Number of Benign tumors 25	01	24

From the Table II when Rough Set classifier is used, TP = 23, FN=02, FP=01, TN=24.

Table III: Results of SVM Classifier

Ground Truth	Predicted by classifier as Malignant	Predicted by classifier as Benign
Number of Malignant tumors 25	21	04
Number of Benign tumors 25	03	22

From the Table III, when SVM classifier is used, TP = 21, FN=04, FP=03, TN=22

Considering the results in Table II and III, performance measurement parameters of both the classifiers are computed. Table IV provides the comparison of performance parameters of both the classification methods.



Table IV: Comparison of Performance measurement parameters

Classifier type	Sensitivity	Specificity	Accuracy of the classifier
Rough Set based	0.92	0.96	0.94
Support Vector Machine	0.84	0.88	0.86

The Table IV shows that Rough Set based classifier has shown better sensitivity, specificity and accuracy. Graph indicating comparison of classifier performance is shown in Figure 2.

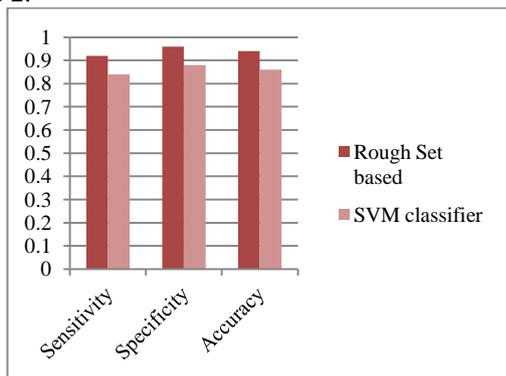


Figure 2. Comparison of performance of classifiers

V. CONCLUSION

In this paper, comparison of performance of two classifiers namely SVM and Rough Set based classifier for tumor classification in medical images is considered. Both the classifiers are implemented using MATLAB R2010a. Liver tumors identified by the radiologists or medical doctors from the abdominal CT scan images of publicly available data repositories are considered for classification. Tumor region features are then extracted using ACHLAC and Legendre Moments. Based on the results of experimentation it is inferred that Rough Set based classifiers outperform the SVM classifiers in terms of classification accuracy by 8%.

REFERENCES

- Han-bin Ouyang, Shuai Liu, Li You, Wen-hu Huang, Shi-zhen Zhong, "Study on the new design of Computer Aided Diagnosis system", IEEE International Symposium on IT in Medicine & Education, 2009.
- <https://pages.cs.wise.edu/~beechung/docs/papers.html>
- <https://pubs.rsna.org>
- <https://radiopaedia.org/articles/liverlesions>
- Poonam Devi, Poonam Dabas, "Liver Tumor Detection using Artificial Neural Networks for Medical Images", International Journal for Innovative Research in Science & Technology, Volume 2, Issue 03, 2015.
- Zhen Liang, "Image Pre-Classification Based on Saliency Map for Image Retrieval", Conference on Information, Communications and Signal Processing, 2009.
- Amita Das, Priti Das, S S Panda, Sukanta Sabut, "Detection of Liver cancer using modified clustering and decision tree classifier in CT images", Pattern Recognition Image Analysis, 2019.
- Santosh Reddy D, R Bharath, P Rajalakshmi, "A novel computer-aided diagnosis framework using Deep Learning for classification of Fatty Liver disease in Ultrasound Imaging", IEEE 20th International Conference on e-Health Networking, Applications and Services, 2018.

- Anju Krishna, Deepesh Edwin, S Hariharan, "Classification of liver Tumor using SFTA based Naïve Bayes classifier and support vector machine", International Conference on Intelligent Computing, Instrumentation and Control Technologies, 2017.
- P. Bharathi, "Medical Image Retrieval Based on LBP Histogram Fourier Features and KNN Classifier", IEEE International Conference on Advances in Engineering & Technology Research, 2014.
- Takahiro Toyoda, Osamu Hasigawa, "Extension of higher order local autocorrelation features", Pattern Recognition, Volume 40, Issue 5, 2007.

AUTHORS PROFILE



Mr. Aravinda H L, is working as Assistant Professor in the Department of Telecommunication Engg., Dr. Ambedkar Institute of Technology, Bengaluru, India and has 15 years of teaching experience. His research area is Image Processing. Content Based Medical Image Retrieval. He is currently pursuing Ph.D from Jain (Deemed to be University).



Dr. M V Sudhamani, currently working as Dean-R&D, Professor and HoD, Dept. of ISE, RNSIT. She is having Teaching and Industrial experience of 25 years. She has specialization in Image Processing, Content-based Image Retrieval, Advanced Algorithms and Databases. Guided and guiding candidates for Ph. D degree. She has carried out two research projects from VTU and AICTE. She has served as member of Board of Examiners (BOE) and Board of Study (BOS) member in VTU and other autonomous institutions across India. She has organized two international conferences ICDECS 2011 and 2015