Performance Analysis of classifiers in Detection of Abnormalities from Ultrasonic Images-A Review

R Harikumar¹, R Karthikamani²

¹ Professor, Bannari Amman Institute of Technology, Department of Electronics and Communication Engineering, Sathyamaagalam, Tamilnadu, India.
² Assistant Professor, Bannari Amman Institute of Technology, Department of Electronics and Instrumentation Engineering, Sathyamaagalam, Tamilnadu, India.
harikumarr@bitsathy.ac.in, karthikamani@bitsathy.ac.in

Abstract. In this paper the ultrasonic abdominal diseases are classified using various classifiers and their performances are analysed. Early diagnosis of abdominal disease is very important in deciding the proper treatment process. To carry out any research work without the comparison of the proposed one with the other already existing method is not effective to give better result. Objective of this paper is to review the classification methods based on the standard parameters like, Sensitivity, Specificity, Accuracy. In this review four classification algorithms Naive Bays Classifier, Support Vector Machine (SVM), k-Nearest Neighbor (kNN) ,Artificial Neural Networks(ANN) are compared.The SVM classifier attained higher accuracy of 98.33 % when compared with all the classifiers.

Keywords: GLCM, SVM, kNN, ANN

1. INTRODUCTION
In the last three decades the abdominal abnormalities of the patients can be easily diagnosed by imaging techniques. The imaging is the noninvasive technique to find the diseases of a human body. There are various imaging methods are available such as radiography, Ultrasonography, Computed Tomography, Magnet resonance imaging. Among these Ultrasonography is widely used because of accessibility, cost effectiveness and non-ionization radiation. The Ultrasonography is used to identify the abdomen abnormalities in liver, gallbladder, spleen, pancreas, kidneys, bladder and uterus. Classification of these diseases are very challenging task because of its similar appearances. There are four phases are involved in the diagnosis and detection of abnormalities in ultrasound images [1].

2. MATERIALS AND METHODS
There are four phases are involved in the diagnosis and detection of abnormalities in ultrasound images. They are Image acquisition, preprocessing, Feature extraction, classification shown in figure 1.

Figure 1: General Block Diagram for classification
2.1 Image Acquisition:

The first and important step in image processing is image acquisition. The B mode ultrasonography is used mostly to capture 2 dimensional anatomic images.

2.2 Image Pre-processing:

The second step in medical image processing is removal of noise. Main aim of a CAD system is to provide useful diagnostic information by reducing the noise. Undesirable data that can diminish the contrast, modify the shape and edges and reduce the fine details of the images is named as noise. Noises are introduced in the images because of physical nature of the system, image acquisition device condition and environmental problems. There are two types of noises mostly affect the images. First one is additive noise and second one multiplicative noise. Additive noises are removed easily but the multiplicative noises are difficult to remove. Ultrasonic images are mostly affected by multiplicative noises. Mathematical model for Multiplicative noise model is

\[ x[m] = s[m]v[m] \]  

Where \( x[m] \) is captured image, \( s[m] \) is noiseless image and \( v[m] \) is multiplicative noise [2]. This type of noise is also called as speckle noise. The speckle noise degrades the contrast of the ultrasonic image and tough to perform edge detection, segmentation.

2.3 Feature Extraction:

In medical image processing feature extraction plays an important role. After preprocessing the feature extraction should be done for further classification. The behavior of an image is defined by its features. The commonly used feature extraction method for classifications is structural features, statistical features, and textural features [3].

3 DETECTION OF ABNORMALITIES THROUGH THE CLASSIFIERS

At present Machine learning is one of the newest topics that facilitate machines to learn from data and make prediction without human intervention. Supervised learning and unsupervised learning is the two main classes of machine learning algorithm. In supervised learning all the data’s are named and the algorithm should predict the output from the input, where as in unsupervised learning the data are not labeled and algorithm learn from inherent grouping of input data. The classification algorithm used in this review are

1. 3.1. Naïve Bays Classifier

It is a neural network based on feed forward concept that belongs to the probability classifier family and can be used to resolve problems with classification and regression. [4]. The Bays rule is applied to assign a higher posterior probability class to new input data from the estimated probability of a new input data class.

\[ P(d / y) = \frac{P(y / d)P(d)}{P(y)} \]  

Posterior probability is represented as \( P(d / y) \), likelihood function is \( P(y / d) \), \( P(d) \) is class earlier probability, \( P(y) \) is predictor former probability. The features used in this classifier are self-governing of each other and varying the value of one feature will not affect or modify the value of other features utilized in this algorithm [5]. This algorithm is mostly used to resolve text classification problem. Naïve bays algorithm provides fast processing of learning and testing and it is generative model based classifier. NB is an effective algorithm and it is commonly used in filtering spam, classifying documents, sentiment prediction etc.
2. **Support Vector Machine (SVM)**

SVM is a fast and accurate supervised learning algorithm used to classify two classifications of groups and is used for problems of regression and classification. [6]. It is a differential classifier that is described by a hyper plane that separates. A line separating plane is known as hyper plane and it uses two dimensional space to separate each class that is it separate the classes using a linear boundary. In the SVM classifier n-dimensional space is used to plot each data, where n is the amount of features designated in feature extraction phase. The hyper plane dimensionality is equivalent to the amount of input features minus one. The equation for decision surface separating the classes in hyper plane is

\[ f(x) = w^T x + b \]  

(3)

Where ‘w’ is the weight factor,’x’is input vector,’b’ termed as bias function.

\[ b = \sum_{i \in I} a_i y_i \phi(x_i) + y_i, \forall I \]

(4)

The margin between two points are represented as \( \frac{2}{\|w\|} \), to maximize this distance denominator value \( \|w\| \) should be minimized. While building the SVM classifier the kernel function and parameters are selected mostly.

**Figure 3** Maximum margin Hyper plane

3.3. **k-Nearest Neighbor Classifier**

**Figure 4** KNN classifier
KNN works based on the principle of similar things being near to each other. In kNN all data's are trained while classification and there is no need of training phase. Based on how closely it matches the points in the training data set, the new data point value will be estimated. While implementing the algorithm we should follow some steps. Initially the training and test data set will be loaded then the distance among test data and training data will be calculated using Euclidian or Hamming distance. Now sort them in ascending order based on the distance value. Next top K rows will be sorted and assign a class to the test points [7]. In this algorithm Euclidian distance metrics are used to locate the nearest neighbor values. The space between two points \((x, y)\) are calculated using the Euclidian distance formula [9].

\[
\text{Distance } d(x, y) = \sum_{i=1}^{N} \sqrt{(x_i - y_i)^2}
\]  

Number of features are denoted by \(N\) and \(x\) and \(y\). The classification accuracy and standard deviation are used to define the classification performance.

3. 3.4. Artificial Neural Network

An ANN is a supervised algorithm for machine learning based on the human neuron model. The ANN has multiple nodes which mimic biological neurons of human brain. The ANN is mostly used in non-linear signal processing, optimization and classification. The ANN is multi-layer neural network which is shown in figure 6.

![Figure 5 Layers of Neural Network](image)

ANN consists of 3 layers such as input layer, hidden layer (multiple) and an output layer [8]. In human brain neurons are interconnected similarly in ANN one layer in connected to each other node in the next year. The input layer receives input in different formats given by the user. The hidden layer is located in between the input and output layer and it has summation and activation function. The final result will be obtained at the output layer. The ANN is used for pattern reorganization. The sensory data is interpreted using machine perceptron or clustering.

The four classifiers accuracies are compared in Table 1.
| Authors               | Data Base                                                                 | Algorithm/Metho dology                                                                 | Accuracy                      | Remarks                                                                                                                                 |
|----------------------|---------------------------------------------------------------------------|---------------------------------------------------------------------------------------|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| AlalehAlivar et al[1] | Ultrasonic images are acquired on a Toshiba SSA 550 digital ultrasound machine with a convex probe. | Feature Extraction: Spatial Domain: GLCM Transform domain Features: Gobar Filter, 2D Wavelet Packet Transform Classifier: KNN | KNN (with combination of the three feature space)-96.1 % | The KNN classifier with Gobar filter provide accuracy of 77.63 % , 2D Wavelet packet transform with KNN gave 85.52 % and KNN with GLCM feature extraction method gave 88.18 % accuracy. By combining these 3 feature space the over all accuracy is increased to 96.1% |
| Yi-ming Lei et al[10]| Ultrasonic Image database is collected from Department of hepatobiliarysurgery | Feature Extraction Methods: PCA, Uniform LBP, GLCM and combination of PCA&U-LBP Classifier: SVM | PCA with SVM-68.75 % Uniform LBP with SVM- 86.25 % PCA+U-LBP with SVM-83.75 % GLCM with SVM-78.75 % | Uniform Local Binary Pattern is rarely used for classification of cirrhotic liver image classification. |
| R.Suganya et al [11] | The US image database is collected from GEM Hospital, Coimbatore and Raj scan center, and Meenakshi Mission Hospital at Madurai | Feature Extraction: GLCM &Haralick features Classifier: SVM | The classification accuracy of SVM with 5 Haralick features is 81.7 % | The accuracy obtained by hybrid classifiers (SVM and Random forest ) was 72.1 %.So the Classification accuracy is improved by using SVM with Haralick features |
| Prajith C.A et al [12]| Ultrasonic liver images were captured from Voluson E6 GE health care | Feature Extraction: First Order Statistical features and GLCM Classifiers: SVM, Gaussian Mixture Model(GMM), | GMM-86 % ANN-88 % SVM- 94 % | GMM is unsupervised learning algorithm and ANN,SVM are supervised learning algorithms. |
| Researcher                                | Data Collection Details                                                                 | Feature Extraction/Selection | Classifier       | Classification Performance | Comments                                                                 |
|------------------------------------------|----------------------------------------------------------------------------------------|------------------------------|------------------|-----------------------------|--------------------------------------------------------------------------|
| Tarek M. Hassan et al [13]               | The US image database is collected from Egyptian Liver Research Institute and the Sherbin Central Research Hospital, Egypt | Feature Extraction/Selection: GLCM, GLRLM, IH /PCA | Multi SVM, KNN, and Naïve Bays | Multi SVM-96.5 % KNN-93.6 % Naïve Bays-95.6 | Compared with KNN and Naïve Bays algorithm, the overall Accuracy is high for Multi-SVM classifier |
| Ali A. Sakr et al [14]                   | Image database is collected from research group                                         | Feature Extraction: Haralick and Histogram feature extraction method | Multi SVM-96.11 % KNN-93.3 % | Better Accuracy for Multi SVM classifier                                  |
| Jitendra Virm ani et al [15]             | Ultrasonic image database is collected from the Department of Radio diagnosis and Imaging at Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India | Feature Extraction/Selection: GLCM-M, GLCM-R /SVD | SVM classifier | GLCM-M, GLCM-R /SVD, SVM classifier the accuracy is 97.5 % GLCM R/SVD with SVM-98.33 % | SVM classifier is mostly used because it is less susceptible to over fitting and provide better performance without dimensionality reduction |
| Abraham Amole et al [16]                | The image database is collected from St. Gregory’s Specialist Clinic and Ultrasound Diagnostic Service, Adeoyo road, Yemetu, Ibadan, LAUTECH Teaching Hospital, Ogbomoso and Several Radiological centre. | Feature Extraction: Second order statistical and GLCM | Fuzzy MLP, ANN | Fuzzy multilayer perceptron (Fuzzy MLP) – 92.73 % and ANN- 81.82 | The classification rules are framed by using fuzzy MLP easily because the network will be trained with enough number of sample images. |
4. DISCUSSION AND CONCLUSION

In machine learning the supervised learning algorithm is widely used. The algorithm will be selected depending on the problem in our hand. Generally SVM and neural networks will provide better result when dealing with multidimensional and continuous features and it require large samples to reach better prediction accuracy but Naive Bays used when small data set is available. The Table 1 shows the classification accuracy obtained for various feature selection and classification algorithms. The classification accuracy is better for GLCM R/SVD with SVM algorithm is 98.33 %. Next is GLCM-M, GLCM-R- with SVM classifier the accuracy is 97.5 %. The accuracy of GLCM features with kNN, Haralick with SVM, GLCM with ANN and Naïve Bayes are 88.18 %, 81.7 %, 88%, 95.6 %. In past studies the authors implemented single classifier algorithm for classification of abdominal ultrasonic images. In future the hybrid classifier algorithm will be used to provide better accuracy instead of single algorithms.

References

[1] AlalehAlivar, HabibollahDaniali, Mohammad SadeghHelfroush, ’Classification of liver diseases using ultrasound images based on feature combination’ IEEE Conference 2014.

[2] Ahmed S.Bafaraj, “Performance Analysis of Best Speckle Filters for Noise Reduction in Ultrasound Medical Images” International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 6 (2019).

[3] ByungEun Park, Won Seuk Jang, Sun Kook Yoo,” Texture Analysis of Supraspinatus Ultrasound Image for Computer Aided Diagnostic System”, Healthc Inform Res. 2016 October;22(4):299-304.

[4] PouriaKaviani, SunitaDhotre,” Short Survey on Naive Bayes Algorithm”, International Journal of Advance Engineering and Research Development, Volume 4, Issue 11, November -2017.

[5] ShaymAAndanAbdulrahman, “Comparative study for 8 computational intelligence algorithms for human identification”, Computer Science Review, May 2020.

[6] RajkumarPalaniappan, Kenneth Sundaraj and Sebastian Sundaraj , “A comparative study of the SVM and k-NN machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals” BME Bioinformatics 2014.

[7] ShaymAAndanAbdulrahman, “Comparative study for 8 computational intelligence algorithms for human identification”, Computer Science Review, May 2020.

[8] K. Ogawa, ‘Computer-aided diagnostic system for diffuse liver diseases with ultrasound images using neural networks’, IEEE Transactions on Nuclear Science, Vol. 45, No. 6, December 1998.

[9] Yi-ming Lei, ‘Cirrhosis Recognition of Liver Ultrasound Images Based on SVM and Uniform LBP Feature’ IEEE 2015.

[10] R. Suganya, S. Rajaram, ’Feature Extraction and Classification of Ultrasound Liver Images using Haralick Texture-Primitive Features: Application of SVM Classifier’, IEEE 2013.

[11] Prajith C.A, SukeshKumar.A, Harish Kareem, ’Supervised Classification and Prediction of Fibrosis Seriousness Using Ultrasound Images’, IEEE 2016.

[12] Tarek M. Hassan, Mohammed Elmogy, ElsayedSallam,"A Classification Framework for Diagnosis of Focal Liver Diseases” IEEE Conference 2015.

[13] Ali.A.Sakr,Magdi Elias Fares,Mai Ramadan,"Automated Focal Liver Lesion Staging Classification based on Haralick Texture Features and Multi-SVM", International Journal of Computer Applications (0975 – 8887),Volume 91 – No.8, April 2014.

[14] JitendraVirmani, Vinod Kumar, Naveen Kalra, NiranjanKhandelwal,"SVM-based characterization of liver cirrhosis by singular value decomposition of GLCM matrix “, in International Journal of Artificial Intelligence and Soft Computing · January 2013.

[15] Abraham Amole, Ojo, John. A,Amole, Abraham. O, "Application Of Fuzzy-MLP Model to Ultrasonic Liver Image Classification”, in Article in European Scientific Journal · April 2014.