Research Article

Artificial Neural Network Application in the Diagnosis of Disease Conditions with Liver Ultrasound Images

Karthik Kalyan,[1] Binal Jakhia,[1] Ramachandra Dattatraya Lele,[2,3] Mukund Joshi,[4] and Abhay Chowdhary[1]

[1] *Systems Biomedicine Division, Haffkine Institute for Training Research and Testing, Parel, Mumbai, Maharashtra 400012, India*
[2] *Research Advisory Council, Haffkine Institute for Training Research and Testing, Parel, Mumbai, Maharashtra 400012, India*
[3] *Nuclear Medicine Department, Jaslok Hospital and Research Centre, Pedder Road, Mumbai, Maharashtra 400026, India*
[4] *Ultrasound Department, Jaslok Hospital and Research Centre, Pedder Road, Mumbai, Maharashtra 400026, India*

Correspondence should be addressed to Ramachandra Dattatraya Lele; drlele@vsnl.net

The preliminary study presented within this paper shows a comparative study of various texture features extracted from liver ultrasonic images by employing Multilayer Perceptron (MLP), a type of artificial neural network, to study the presence of disease conditions. An ultrasound (US) image shows echo-texture patterns, which defines the organ characteristics. Ultrasound images of liver disease conditions such as "fatty liver," "cirrhosis," and "hepatomegaly" produce distinctive echo patterns. However, various ultrasound imaging artifacts and speckle noise make these echo-texture patterns difficult to identify and often hard to distinguish visually. Here, based on the extracted features from the ultrasonic images, we employed an artificial neural network for the diagnosis of disease conditions in liver and finding of the best classifier that distinguishes between abnormal and normal conditions of the liver. Comparison of the overall performance of all the feature classifiers concluded that "mixed feature set" is the best feature set. It showed an excellent rate of accuracy for the training data set. The gray level run length matrix (GLRLM) feature shows better results when the network was tested against unknown data.

1. Introduction

Ultrasound imaging modality is quite popular and most widely used modality for visualizing and studying the liver for any disease conditions without causing any pain or discomfort to the patient. Ultrasound liver imaging is widely used due to its noninvasive nature and low cost as compared to other imaging modalities. The diagnosis of various diseases is performed on the basis of various image features such as the echogenicity, legion shape, and echo texture. Liver imaging is one of the best techniques of early detection of liver diseases and early detection is very important because it saves patients from further ailments such as enlarged stomach filled with ascites fluid, bleeding varices, and encephalopathy or sometimes jaundice. Liver disease conditions such as fatty liver, cirrhosis, and hepatomegaly are known for producing distinctive echo patterns during US imaging as shown in Figure 1; however these images are also known to be visually

challenging for interpreting them because of their imaging artifacts and speckle noise. As a result of it, the sonographers have to rely upon additional pathological tests [1–3].

A visual measure for diagnosing a liver is done by evaluating the liver echogenicity and the granular structure and surface echo-texture of liver. Texture analysis presents various image features, which characterize different liver conditions including normal and abnormal conditions. Texture analysis also provides some important information that cannot be obtained from visual examination of ultrasonic images. A texture analysis of each liver disease condition differs from other disease conditions as well as from the normal liver image. Normal liver ultrasound image is described as pyramidal with smooth surface and no lumps. The normal liver parenchyma is of homogeneous echogenicity [2]. Fatty liver ultrasound image shows liver parenchyma of hyperechogenicity. Fatty liver is reversible in its early stages; therefore an early detection is very essential [4, 5]. The ultrasound image of liver
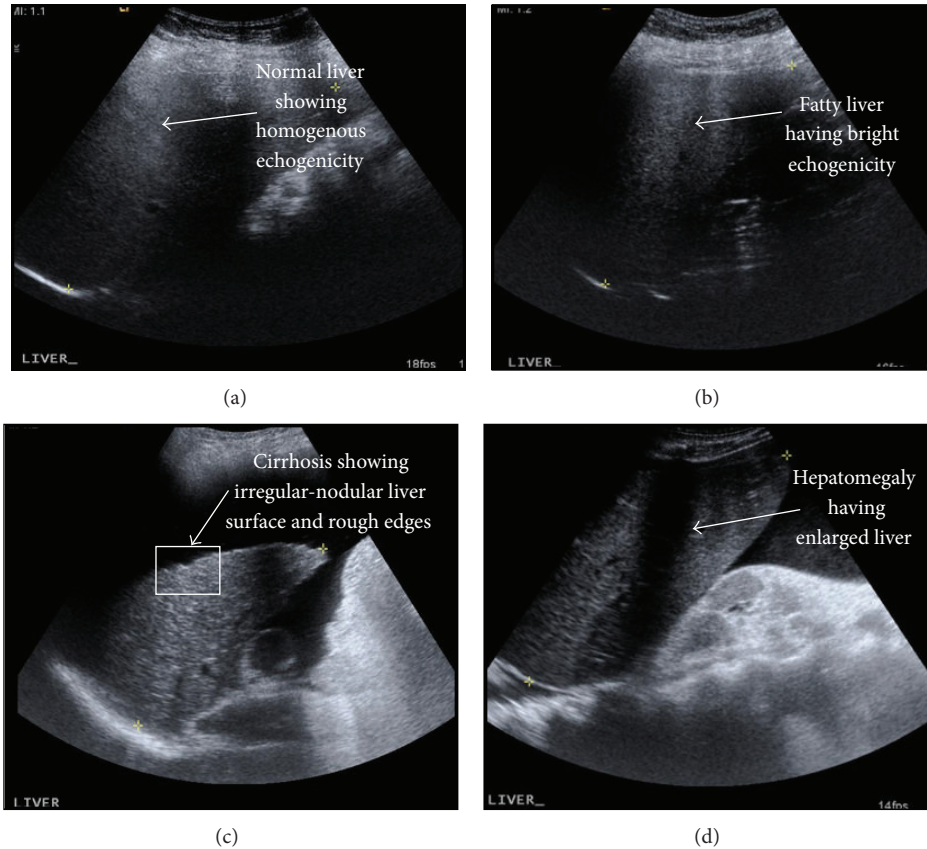
FIGURE 1: Ultrasound image of (a) normal liver, (b) fatty liver, (c) cirrhosis, and (d) hepatomegaly.

cirrhosis shows inhomogeneous echo texture and irregular-nodular liver surface [5]. Most gray levels of cirrhotic tissue appear darker than the normal tissue [6].

The granular structure of the tissue area can be examined to characterize it. This specific granular pattern of normal liver, cirrhotic liver, and fatty liver can be described as texture and thus "texture analysis" for tissue characterization may be used to study and correlate the physiological changes in the liver. Moreover this approach provides some important information that may not be obtained through visual interpretation of ultrasound images. These echo-patterns extracted from the images can be studied and processed for characterization of liver diseases [2].

Picture Archive and Communication Systems (PACS) are most widely used systems for medical image storage and retrieval. PACS are comprehensive networks of digital devices designed for digital image data management, image acquisition, data transmission, storage, image display, and management of diagnostic imaging studies, interfaces to printers and portable media, and communication routes to other electronic systems. PACS are usually based on DICOM (Digital Imaging and Communications in Medicine) standards and are comprehensive management systems for diagnostic imaging studies that are increasingly used in hospitals and other health care systems [7, 8].

An artificial neural network (ANN) is a nonlinear, computational, and mathematical model, comprised of densely interconnected simple processing elements called neurons. Artificial neural networks are inspired by information processing simulation in human brain by biological neurons. The main characteristics of neural networks are their ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data. A key benefit of neural networks is that a model of the system can be built from the available data [9, 10].

The aim of this study was to investigate the image feature classifiers and to find the best feature classifier for the diagnosis of liver disease conditions using artificial neural network. The objectives included quantification of various features extracted from the ultrasound image and utilize them as an input to artificial neural networks towards the liver disease classification.

Previous studies conducted by Kadah et al. and Jeon et al. compared the analysis of statistical classifiers and neural network based classifiers generated using tissue characterization parameters from liver images [11, 12] by various feature extraction algorithms [11]. Another study conducted by Lee et al. established the classification of liver lesions such as cyst, hemangioma, and malignancies using the multiple regions of interests (ROIs) based feature selection methods. Here the liver lesion classification from the US images is heavily dependent upon certain characteristics (traits) such as internal echo, morphology, edge, echogenicity, and posterior echo enhancement. The proposed method seems to have

achieved the enhanced and stable classification regardless of features used alongside outperforming the existing classification methods that are designed for focal liver lesions [13]. A study by Plesea-Condratovici et al. established the evaluating ability of a neural network based tool in the prediction of steatosis of liver, where the data has been collected from 100 patients and a data matrix is generated from that data and 10 variables are dedicated for the purpose. Out of the 10 variables 7 are input and 3 are output parameters. The results are validated against another study where the level of steatosis is known in the patients [14]. Study published by Gletsos et al. proposes the use of CAD system to classify the hepatic lesions such as hepatic cysts, hemangiomas, and hepatocellular carcinoma from CT images [15].

The rest of this paper is organized as follows. Section 2 gives an outline of the overall methodology. Then, the processes of image acquisition, image preprocessing, and image processing (i.e., both feature extraction and feature selection processes) are explained under Sections 3, 4, and 5, respectively. We talk about the concept of artificial neural networks (in particular about validation, implementation, and testing aspects) under Section 6. We then talk about the results and discussion part under Section 7. Finally, the conclusion is given under Section 8.

2. Outline of Methodology

In this study, we acquired normal and diseased (i.e., abnormal) liver ultrasound images from "The Ultrasound Department at Jaslok Hospital and Research Centre." Using MATLAB along with "image processing toolbox" these images are then subjected to three different image preprocessing techniques, namely, "cropping," "edge detection," and "background subtraction," in order to accentuate the region of interest from the acquired images (i.e., liver). After image preprocessing, textural features such as "intensity histogram (IH)" [16, 17], "gray-level co-occurrence matrix (GLCM)" [18], "gray-level run length matrix (GLRLM)" [19], and "invariant moments (IM)" [20] were extracted from the preprocessed ultrasound images to calculate the adequate texture features. The features extracted from feature extraction phase were further processed (i.e., narrowed) utilizing "feature selection" method to obtain most significant and optimal features that represent the liver characteristics. WEKA software [21] was utilized within the feature selection phase to provide the selected significant features. These optimal features were then provided as an input to the neural network for classification. Artificial neural network employing a back-propagation algorithm [22] was utilized to classify the normal and abnormal liver disease conditions and to determine which feature classifier is best for classification. The performance of the neural network based classifier was determined using "confusion matrix" and "receiver operating characteristics (ROC)" curve analysis. These techniques as shown in Figure 2 are further elaborated in following subsections.

3. Image Acquisition

All B-mode ultrasound images were acquired from Jaslok Hospital and Research Centre, Mumbai. The patients' data report and their ultrasound images were initially analyzed and the images were shortlisted on the basis of the most prevalent disease condition amongst the acquired data, age of the patient, and severity of the disease. Here, the images of the patients containing ascites along with the liver disease conditions were eliminated from shortlisting as the selection parameters only focused on liver disease conditions. Overall, a total of 60 liver ultrasonic images were selected for this study. These included 30 cases of normal and 30 cases of abnormal liver sonograms. The 30 cases of abnormal liver are comprised of 10 cases of cirrhosis, 10 cases of fatty liver, and 10 cases of hepatomegaly. The images were acquired from both male and female subjects with a mean age group of 53 (±15) years. During image acquisition, the operator examined the whole liver area from different orientations and saved a single frame.

4. Image Preprocessing

Image preprocessing techniques are used to select and enhance the region of interest and to eliminate erroneous data, which is of no interest from the acquired images. The images are subjected to three types of image preprocessing techniques such as cropping, edge detection, and background subtraction as shown in Figure 3.

4.1. Cropping. It is an operation, which is performed on acquired images to accentuate the region of interest (i.e., the liver) and to remove all the unwanted artifacts such as written labels and background noise from them. The cropping operation was done on all images by cutting out the rest of the area, which did not contain the liver, leaving a rectangular region consisting of only the region of interest. Cropping operation was performed on all images in MATLAB.

4.2. Edge Detection. It is one of the vital steps of image preprocessing. Although cropping operation extracts liver from the original image, it does not crop along the boundary of liver. Cropping is limited to a rectangular frame only whereas the shape as well as structure of liver is pyramidal; hence edge detection operation is necessary to define the edges of liver. Edge detection technique outlines the liver boundary in the image. In this work the segmentation is achieved by "Active Snake Contour" model, providing a contour over liver boundary. The active snake contour model is a semiautomatic model, where initializing a curve or a contour close to the boundary of the region of interest is carried out manually by the user and the model functions by deforming the contour through number of iterations to fit to the boundary of region of interest [23]. In this study, the edge detection was performed in "ImageJ" software using an "Absnake" plugin.

4.3. Background Subtraction. It is a method to eliminate unwanted intensity values which are outside the contour (edge) in the images (rectangular images with dimension of
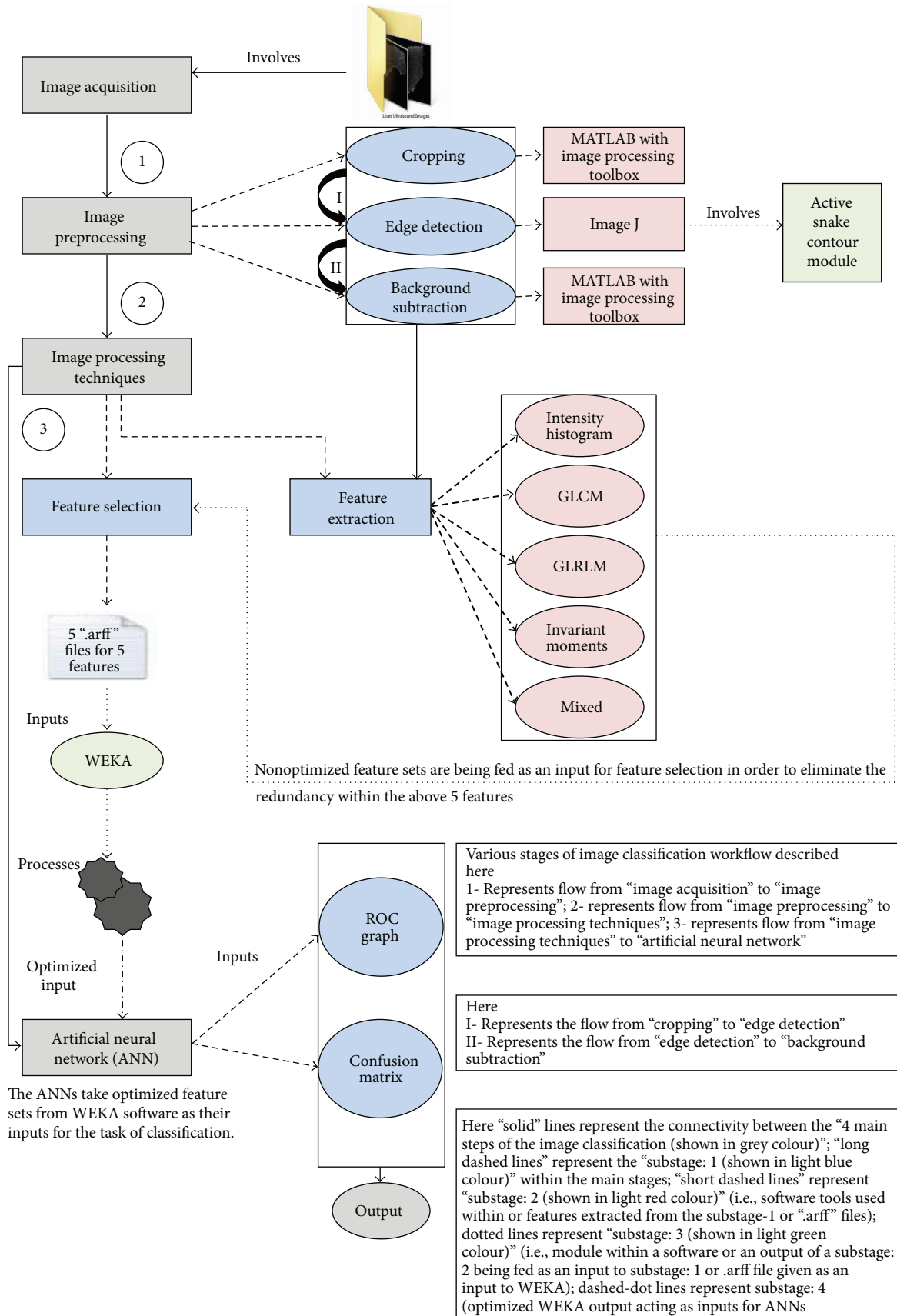
Figure 2: Image classification workflow utilising ANNs.
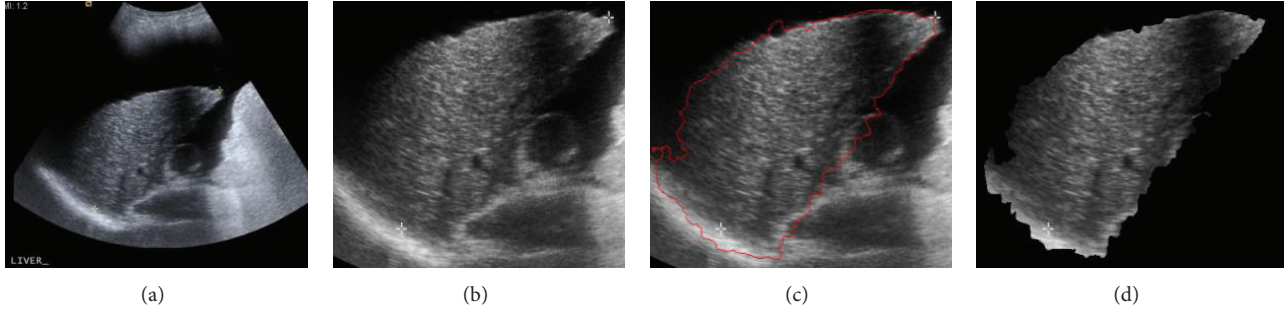
<center>(a)   (b)   (c)   (d)</center>

FIGURE 3: Workflow of image preprocessing step, (a) original ultrasound image, (b) image after cropping operation, (c) image after edge detection, and (d) image after background subtraction.

$m \times n$ that includes pyramidal liver) to avoid the calculation of these unwanted intensities that will be incorporated during extraction of feature parameters. Background subtraction is performed to avoid calculation of pixel intensities, which are outside the region of interest. Background subtraction was performed in MATLAB.

5. Image Processing

Texture is an image feature that provides important characteristics for surface and object identification from an image. Texture is characterized by the spatial distribution of gray levels in a neighborhood in an image. In texture analysis, the most difficult as well as important aspect is to define a set of meaningful features that explores the characteristics of the texture [24, 25]. Image processing techniques involve identifying these sets of essential features and extracting these features from the ultrasound image for further processing. "Feature extraction" and "feature selection" are two most vital steps of image processing and are explained under Sections 5.1 and 5.2 in detail.

5.1. Feature Extraction. Feature extraction is a critical step for ultrasonic liver classification. Feature extraction methodologies analyze the preprocessed images to extract the most prominent features that represent various sets of features based on their pixel intensity relationship and statistics. A set of four features (i.e., statistical texture features), namely, intensity histogram, gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), and invariant moments, were extracted from each of the total 60 images in MATLAB using respective modules, which calculated parameters belonging to each set.

Each feature set comprises individual image parameters. Features derived from *intensity histogram* features include moments such as mean, standard deviation, average energy, entropy, skewness, and kurtosis [16, 17]. *GLCM* features include autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity energy, entropy, homogeneity (1), homogeneity (2), maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation (1), information measure of correlation (2), and inverse difference normalized [18]. The *GLRLM* features

TABLE 1: Features corresponding to intensity histogram.

S/number	Name	Equation
1	Mean (μ)	$(\mu) = \sum_{i=0}^{N-1} i p(i)$
2	Energy (E)	$(E) = \sum_{i=0}^{N-1} p(i)^2$
3	Variance (σ^2)	$(\sigma^2) = \sum_{i=0}^{N-1} (i - \mu)^2 p(i)$
4	Entropy (H)	$H = \sum_{i=0}^{N-1} p(i) \log_2 p(i)$
5	Skewness (μ_3)	$\mu_3 = \sigma^{-3} \sum_{i=0}^{N-1} (i - \mu)^3 p(i)$
6	Kurtosis (μ_4)	$\mu_4 = \sigma^{-4} \sum_{i=0}^{N-1} (i - \mu)^4 p(i) - 3$

include short run emphasis (SRE), long run emphasis (LRE), run length nonuniformity (RLN), gray level nonuniformity (GLN), run percentage (RP), low gray-level run emphasis (LGRE), high gray-level run emphasis (HGRE), short run low gray-level emphasis (SRLGE), short run high gray-level emphasis (SRHGE), long run low gray-level emphasis (LRLGE), and long run high gray-level emphasis (LRHGE) [19]. The *Invariant moments* are I_1, I_2, I_3, I_4, I_5, I_6, and I_7 and these moments are invariant under translation, changes in scale, and also rotation. So it describes the image despite its location, size, and rotation [20].

5.1.1. Intensity Histogram Features. The intensity-level histogram is a function showing (for each intensity level) the number of pixels in the whole image, which have this intensity. Here p represents the "pixel intensity" and "$p(i)$" represents the pixel intensity of I value. The function $f(x, y)$ can take discrete values $i = 0, 1, \ldots, N-1$, where N is the total number of intensity levels in the image. We can calculate the individual features under this feature extraction technique utilizing the formulas as shown in Table 1.

5.1.2. GLCM Features. Also termed as "spatial gray level dependency matrices," it is one of the most widely used statistical tools for extracting texture information from images. The GLCM of a 2D (i.e., $N_x \times N_y$) image containing pixels

TABLE 2: Features of GLCM.

S/number	Name	Equation
1	Mean (μ_x, μ_y)	$\mu_x = \sum_i \sum_j i \cdot p(i,j);$ $\mu_y = \sum_i \sum_j j \cdot p(i,j)$
2	Standard deviations (σ_x, σ_y)	$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 \cdot p(i,j);$ $\sigma_y = \sum_i \sum_j (i - \mu_y)^2 \cdot p(i,j)$
2.1	Autocorrelation (f_1)	$f_1 = \sum_i \sum_j (ij) \cdot p(i,j)$
2.2	Contrast (f_2)	$f_2 = \sum_{n=0}^{N_9-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_G} p(i,j) \mid \lvert i-j \rvert = n \right\}$
2.3	Correlation (f_3)	$f_3 = \dfrac{\left[\sum_i \sum_j (ij)\, p(i,j) - \mu_x \mu_y \right]}{\left[\sigma_x \sigma_y \right]}$
2.4	Cluster prominence (f_4)	$f_4 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i,j)$
2.5	Cluster shade (f_5)	$f_5 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i,j)$
2.6	Dissimilarity (f_6)	$f_6 = \sum_i \sum_j \lvert i - j \rvert \cdot p(i,j)$
2.7	Energy (f_7)	$f_7 = \sum_i \sum_j p(i,j)^2$
2.8	Entropy (f_8)	$f_8 = -\sum_i \sum_j p(i,j) \log(p(i,j))$
2.9	Homogenecity (f_9)	$f_9 = \sum_i \sum_j \dfrac{1}{1 + (i+j)^2} p(i,j)$
2.10	Maximum probability (f_{10})	$f_{10} = \max_{i,j} p(i,j)$
2.11	Sum of squares (f_{11})	$f_{11} = \sum_i \sum_j (i - \mu)^2 p(i,j)$
2.12	Sum average (f_{12})	$f_{12} = \sum_{i=2}^{2N_g} i p_{x+y}(i)$
2.13	Sum variances (f_{13})	$f_{13} = \sum_{i=2}^{2N_g} (i - f_{14})^2 i p_{x+y}(i)$
2.14	Sum entropy (f_{14})	$f_{14} = \sum_{i=2}^{2N_g} p_{x+y}(i) \log \{ p_{x+y}(i) \}$
2.15	Difference variance (f_{15})	$f_{15} = \text{Variance of } p_{x-y}$
2.16	Difference entropy (f_{16})	$f_{16} = -\sum_{n=0}^{N_9-1} p_{x-y}(i) \log \{ p_{x-y}(i) \}$
2.17 2.18	Information measures of correlation-1 (f_{17}) Information measures of correlation-2 (f_{18})	$f_{17} = \dfrac{HXY - HXY1}{\max\{HX, HY\}}$ $f_{18} = (1 - \exp[-2.0\,(HXY2 - HXY)])^{1/2}$ *where HX and HY are entropies of P_X and P_Y* $HXY = \sum_I \sum_J p(i,j) \log(p(i,j))$ $HXY1 = \sum_I \sum_J p(i,j) \log \{ p_x(i)\, p_y(j) \}$ $HXY2 = \sum_I \sum_J p_x(i)\, p_y(j) \log \{ p_x(i)\, p_y(j) \}$
2.19	Inverse difference (f_{19})	Same as homogenecity

TABLE 2: Continued.

S/number	Name	Equation
2.20	Inverse difference normalized [INN] (f_{20})	C_{ij} = the co-occurrence probability between grey levels i and j is defined as $$C_{ij} = \frac{P_{ij}}{\sum_{i,j=1}^{G} p_{ij}}$$ $$f_{20} = \sum_{i,j=1}^{G} P_{IJ} \frac{C_{ij}}{1 + |i-j|^2 / G^2}$$
2.21	Inverse difference moment normalized (f_{21})	$$f_{21} = \sum_{i,j=1}^{G} P_{IJ} \frac{C_{ij}}{1 + (i-j)^2 / G^2}$$

with gray levels $(0, 1, \ldots, G-1)$ is also a 2D matrix "$P(i, j)$", where each and every matrix element depicts the probability of joint occurrence of intensity levels "k" and "1" at a certain distance "d" and an angle "θ". Here "$p(i, j)$" is the (i, j)th entry in a normalized GLCM. The mean (μ_x, μ_y) and standard deviation (σ_x, σ_y) for the rows and columns of the 2D matrix can be calculated using the formulas that are shown in Table 2. We can also calculate the various individual features under this feature extraction technique utilizing the formulas provided in Table 2.

5.1.3. GLRLM Features. Grey-level run-length matrix (GLRLM) is a matrix from which the texture features can be extracted for texture analysis. For a given 2D image, GLRLM is a 2D matrix in which element "$p(i, j)$" gives the total number of consecutive runs of length "j" at grey level "i". Here "M" represents the number of gray levels and "N" represents the maximum run length (here a run length is considered to be a number of neighboring pixels that possess the same grey intensity in a particular direction). From this matrix almost 11 scalar parameters can be computed which analyze the image texture [26] and these parameters are provided in Table 3.

5.1.4. Invariant Moments Features. The idea of utilizing moments within shape recognition became popular in 1962 when Hu utilised algebraic invariants to derive a set of invariants. Hu's 7 moment invariants are invariant under translation, changes in scale, and also rotation. So it describes the image despite its location, size, and rotation. The moment invariants are generally specified in terms of normalized central moments. Here, the central moments are depicted by μ_{pq}, the raw moments are defined by μ_{ii} [27], and these invariants are shown in Table 4.

5.1.5. Mixed Features. Mixed features are a combined set of all features, that is, intensity histogram features, GLCM features, GLRLM features, and invariant moment features. Mixed features are a combination of these features including a total of 47 features including all 46 features of intensity histogram features, GLCM features, GLRLM features, and invariant moment features and length of the liver. Mixed features set is created to check significance of all attributes within all features and to serve as an input to neural network.

5.2. Feature Selection. While feature extraction techniques are applied to extract as many image parameters as possible that identify liver characteristics, a feature selection algorithm is necessary to select few of those extracted features which are most significant and which describe the liver characteristics the best.

A total of 46 features were extracted in feature extraction process from each image but all of these features cannot be supplied to the neural network because the number of features is high. Although each feature is important in classification only few of these features are very significant in classifying and identification of the disease conditions. Therefore instead of using all of these features as input, only those features, which have high significance, were selected.

There are many feature selection algorithms used in this process and each performed search in their unique way. Most of the feature selection algorithms involve a search method throughout the whole space. Many search methods calculate individual feature's significance and rank them accordingly. These methods also provide best features from a given set. Waikato Environment for Knowledge Analysis (WEKA) software gives a variety of feature selection options that include "heuristic," "genetic," and "Bayesian" algorithms. Here, feature selection was performed using WEKA software version 3.6.9 [21].

WEKA is compatible with and recognizes only Attribute-Relation file format (i.e., ".arff file format"); therefore ".arff file" was generated containing feature information of all images (normal as well as abnormal). Along with ".arff file" consisting of four features, one more ".arff file" called mixed features was created containing all the 46 features and the length of liver parameter. Feature selection process incorporated methods like genetic search method, random search, rank search, and so forth to select the best attributes amongst a large set of features. These methods select features on the basis of their ability to correctly identify the pattern in training. WEKA provides option of varied search methods for selection of attributes, out of which we used only two search methods, namely, "RandomSearch" which performs a random search in the space of attribute subsets and "GeneticSearch" which performs a search using the simple genetic algorithm. The feature selection algorithms such as "RandomSearch" and "GeneticSearch" of WEKA software were used to generate the results as shown in Table 5.

TABLE 3: Features of GLRLM.

S/number	Name	Equation
1	Gray level run-length pixel number matrix	$P_P(i,j) = p(i,j) \cdot j$
2	Gray-level run-number vector	$P_g(i) = \sum_{j=1}^{N} p(i,j)$
3	Run-length run-number vector	$P_g(i) = \sum_{j=1}^{N} p(i,j)$
3.1	Short run emphasis [SRE] (f_1)	$f_1 = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \dfrac{p(i,j)}{j^2} = \dfrac{1}{n_r} \sum_{j=1}^{N} \dfrac{p_r(j)}{j^2}$
3.2	Long run emphasis [LRE] (f_2)	$f_2 = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(p(i,j) \cdot j^2 \right) = \dfrac{1}{n_r} \sum_{j=1}^{N} p_r(j) \cdot j^2$
3.3	Gray-level nonuniformity [GLN] (f_3)	$f_3 = \dfrac{1}{n_r} \sum_{i=1}^{M} \left(\sum_{j=1}^{N} (p(i,j)) \right)^2 = \dfrac{1}{n_r} \sum_{i=1}^{N} p_g(i)^2$
3.4	Run length nonuniformity [RLN] (f_4)	$f_4 = \dfrac{1}{n_r} \sum_{i=1}^{M} \left(\sum_{j=1}^{N} (p(i,j)) \right)^2 = \dfrac{1}{n_r} \sum_{i=1}^{N} p_r(i)^2$
3.5	Run percentage [RP] (f_5)	$f_5 = \dfrac{n_r}{n_p}$
3.6	Low gray-level run emphasis [LGRE] (f_6)	$f_6 = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \dfrac{p(i,j)}{i^2} = \dfrac{1}{n_r} \sum_{i=1}^{N} \dfrac{p_g(i)}{i^2}$
3.7	High gray-level run emphasis [HGRE] (f_7)	$f_7 = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(p(i,j) \cdot i^2 \right) = \dfrac{1}{n_r} \sum_{i=1}^{N} p_r(j) \cdot i^2$
3.8	Short run low gray-level emphasis [SRLGE] (f_8)	$f_8 = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \dfrac{p(i,j)}{i^2 \cdot j^2}$
3.9	Short run high gray-level emphasis [SRHGE] (f_9)	$f_9 = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \dfrac{p(i,j) \cdot i^2}{j^2}$
3.10	Long run low gray-level emphasis [LRLGE] (f_{10})	$f_{10} = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} \dfrac{p(i,j) \cdot j^2}{i^2}$
3.11	Long run high gray-level emphasis [LRHGE] (f_{11})	$f_{11} = \dfrac{1}{n_r} \sum_{i=1}^{M} \sum_{j=1}^{N} p(i,j) \cdot i^2 \cdot j^2$

TABLE 4: Features corresponding to invariant moments.

S/number	Name	Equation
1	Raw moments (M_{ij})	$M_{ij} = \sum_x \sum_y x^i y^j p(x,y)$
2	Central moments (μ_{pq})	$f(x,y)$ is a digital image; p, q are raw moments; $\overline{X} = M_{10}/M_{00}$ and $\overline{Y} = M_{01}/M_{00}$ are the components of the centroid: $\mu_{pq} = \sum_{m}^{p} \sum_{m}^{q} \binom{p}{m}\binom{q}{n} (-\overline{x})^{p-m} (-\overline{y})^{q-n} M_{mn}$
3	Invariant moments	$\eta_{ij} = \dfrac{\mu_{ij}}{\mu_{00}^{(1+(i+j)/2)}}$ $I_1 = \eta_{20} + \eta_{02}$, $I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$, $I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - 3\eta_{03})^2$, $I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2$, $I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + 3\eta_{12})^2 - (3\eta_{21} + 3\eta_{03})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]$ $I_6 = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$, $I_7 = (3\eta_{21} - 3\eta_{03})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - (3\eta_{21} + 3\eta_{03})^2 \right] - (\eta_{30} - 3\eta_{21})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right]$

TABLE 5: Features selected by WEKA software.

S/number	Feature category	Number of features before feature selection	Number of features after feature selection	Selected features
1	Intensity histogram	6	4/6	Variance, Skewness, Kurtosis, and Entropy
2	GLCM	22	11/22	Contrast, correlation-1, correlation-2, cluster shade, homogeneity, maximum probability, sum of squares: variance, sum variance, difference entropy, information measure of correlation-1, and information measure of correlation-2
3	GLRLM	11	6/11	Short run emphasis, gray-level nonuniformity, low gray-level run emphasis, high gray-level run emphasis, short run high gray-level emphasis, and long run high gray-level emphasis
4	Invariant moments	7	4/7	I_2, I_3, I_4, and I_6
5	Mixed features	47	9/47	Homogeneity, sum of average, difference variance, information measure of correlation-1, information measure of correlation-2, inverse difference normalized, short run emphasis, short run high gray-level emphasis, and length
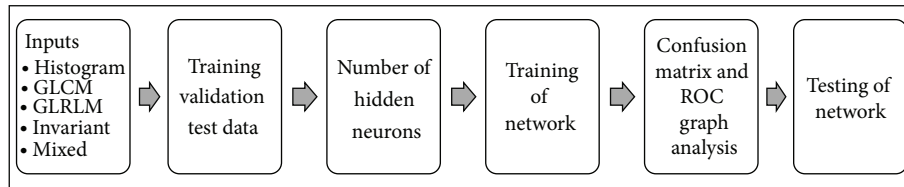
FIGURE 4: Workflow of implementation of artificial neural network.

These optimal features were selected by WEKA software for serving as an input in neural network.

6. Artificial Neural Network

An artificial neural network builds a model of existing system and learns from the previous or known samples and trains the network to achieve target with minimum error [22].

6.1. Implementation, Validation, and Testing. The features extracted from the images after the feature selection process act as an input to the neural network as shown in Figure 4. To perform the analysis for image classification, the back propagation algorithm has been shortlisted and is implemented using MATLAB's Neural Network Pattern Recognition Tool (nprtool). The back propagation algorithm was chosen due to the networks ability to learn and store immense amounts of mapping relations of input-output model without the need for prior disclosure of mathematical equations pertaining to these mapping relations. The algorithm also regulates the network's weight and threshold values in order to obtain minimum error sum of square [28]. The designed neural network classifier used a two-layer feed-forward back propagation network. Two-layer feed-forward network can be best defined as a network with sigmoid hidden and output neurons.

The network was trained with scaled conjugate gradient back propagation [22, 29].

To train the network, the input data and target data need to be fed into the network. The network then divides the input sample data into three different samples, which are training, validation, and testing samples. The training samples are used to train the network, and the network is adjusted according to its error. The validation samples are used to measure network generalization and to halt the training when generalization stops improving. Testing samples are then used to provide an independent measure of the network performance during and after training. If the error of the network is still large, the network can be retrained back as to get more accurate and efficient result [22, 29]. From the training dataset, we utilized 80% of data for training, 10% of data for validation, and 10% of data for testing purposes, while we utilized 20 samples of normal condition and 20 samples of abnormal condition to create the testing dataset to test the efficiency of the artificial neural network.

Confusion matrix as well as ROC graph depicts the overall classification rate and accuracy of the network. If the overall classification rate and the accuracy are high, it signifies that the network was successful in correctly classifying the two classes. After training the network for sufficient number of epochs till the network is perfectly trained having low MSE and less misclassifications, confusion matrix and ROC
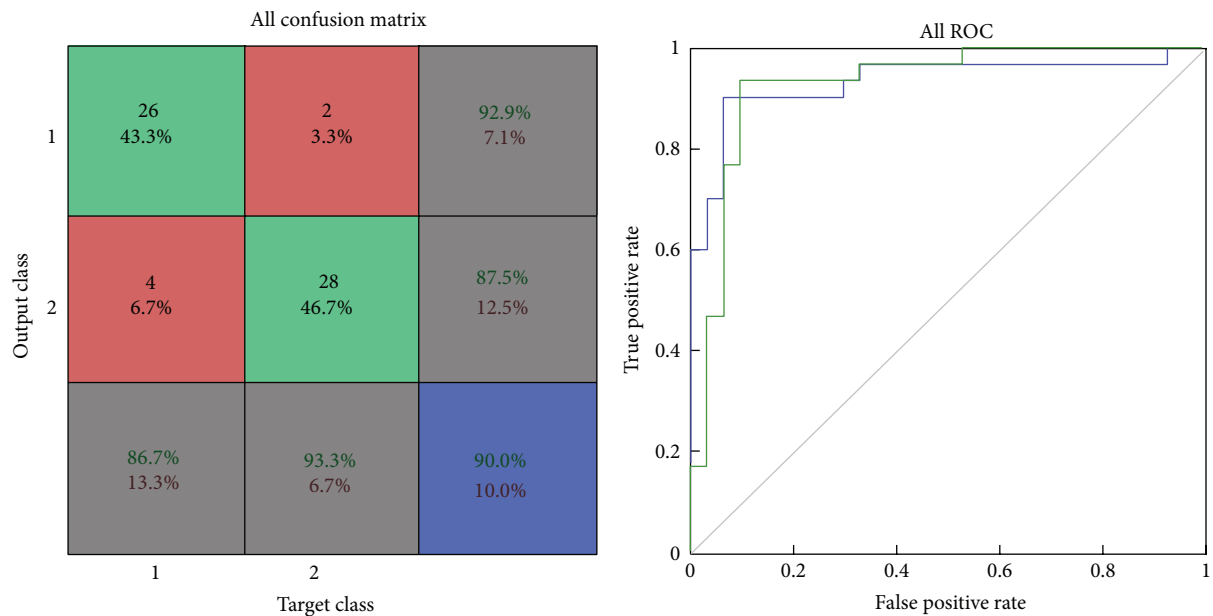
All confusion matrix

FIGURE 5: Confusion matrix and ROC plot of GLRLM training data.

graph were plotted to measure the true positive rate, that is, sensitivity, true negative rate, that is, specificity, false positive rate, false negative rate, and accuracy of the network [29].

Analysis of ROC graph and confusion matrix of the trained network are generally more than enough for evaluating the designed neural network classifier's accuracy. There is an additional option to test network on more data and then decide the quality of the network's performance. Additional tests were performed on test input dataset, comprised of a smaller sample set to evaluate the network's performance on test data. "MSE" and "percent error" provided the mean squared error and difference between the output and target test data, respectively.

7. Results and Discussion

The performance of the designed neural network classifier is measured in terms of accuracy. This term refers to the ability of the model to correctly predict the class of new unseen data. Classification accuracy is calculated by determining the percentage of cases in which the test sets are correctly classified. The performance of the neural network was calculated by analysis of confusion matrix and the receiver operator characteristic curve (ROC).

Confusion Matrix. The diagonal cells show the number of classes that were correctly classified and the off diagonal cells show the misclassified cases. The blue cell in the bottom right shows the total percent of correctly classified cases (in green) and the total percent of misclassified cases (in red) [29].

ROC Graph. The colored lines in this graph represent the ROC curves for each of the two output categories. The ROC curve is a plot of the true positive rate (sensitivity) versus the false positive rate (1-specificity) as the threshold is varied.

A perfect test would show points in the upper-left corner, with 100% sensitivity and 100% specificity [29].

Results of the training data show that selected GLRLM features yield an accuracy of 90%, sensitivity of 86.7%, specificity of 93.3%, false positive rate computed of 6.7%, false negative rate computed of 13.3%, and misclassification rate of 10%. Size of the input dataset loaded in the network was of 60 samples. Of total 60 samples, 54 samples were correctly classified and 6 samples were misclassified by this network as shown in Figure 5. ROC graph shows the plotting of true positive rate against false positive rate (1-specificity). ROC graph of this network shows an excellent classification between the two categories as the curves lie between the diagonal and the upper-left corner but mainly towards the upper-left corner as shown in Figure 5.

The results of the testing data show that GLRLM histogram features yield an accuracy of 95%, sensitivity of 95%, specificity of 95%, false positive rate computed of 5%, false negative rate computed of 5%, and misclassification rate of 5%. Size of the input dataset loaded in the network was of 40 samples. Out of the total 40 samples, 38 samples were correctly classified and 2 samples were misclassified by this network as shown in Figure 6. ROC graph shows the plotting of true positive rate against false positive rate (1-specificity). ROC graph of this network shows a perfect classification between the two categories as the curves lie in the region of upper-left corner and they are very far from the diagonal as shown in Figure 6.

The results of the training data show that selected mixed features yield an accuracy of 91.67%, sensitivity of 93.33%, specificity of 90%, false positive rate computed of 10%, false negative rate computed of 6.7%, and misclassification rate of 8.33%. Size of the input dataset loaded in the network was of 60 samples. Out of the total 60 samples, 55 samples were
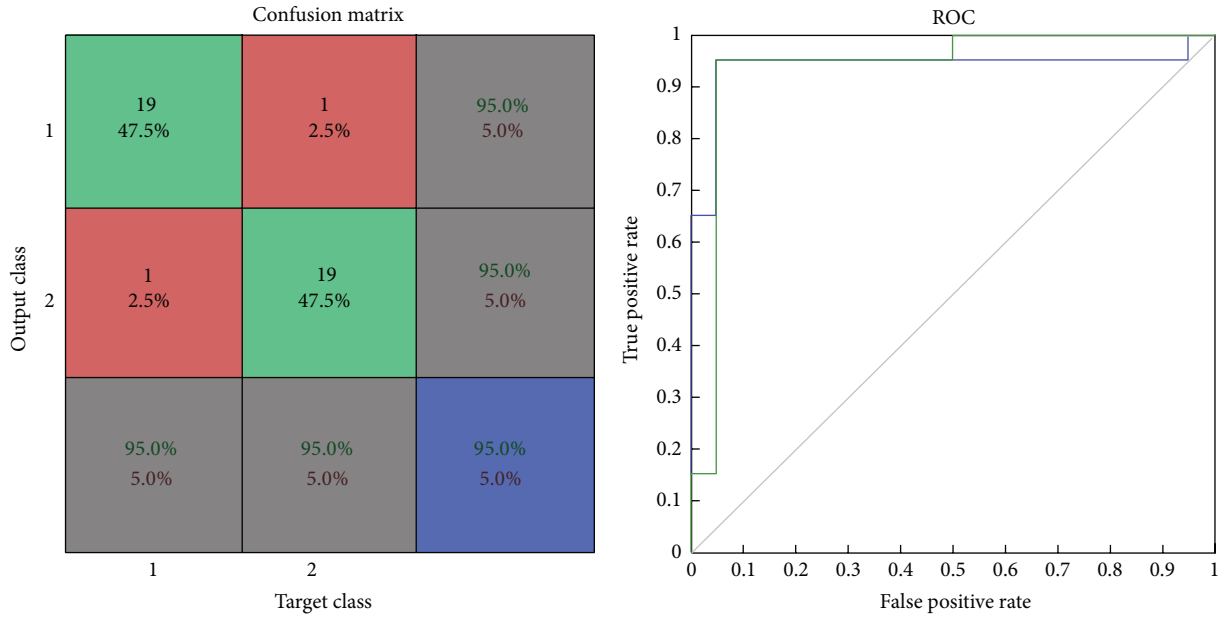
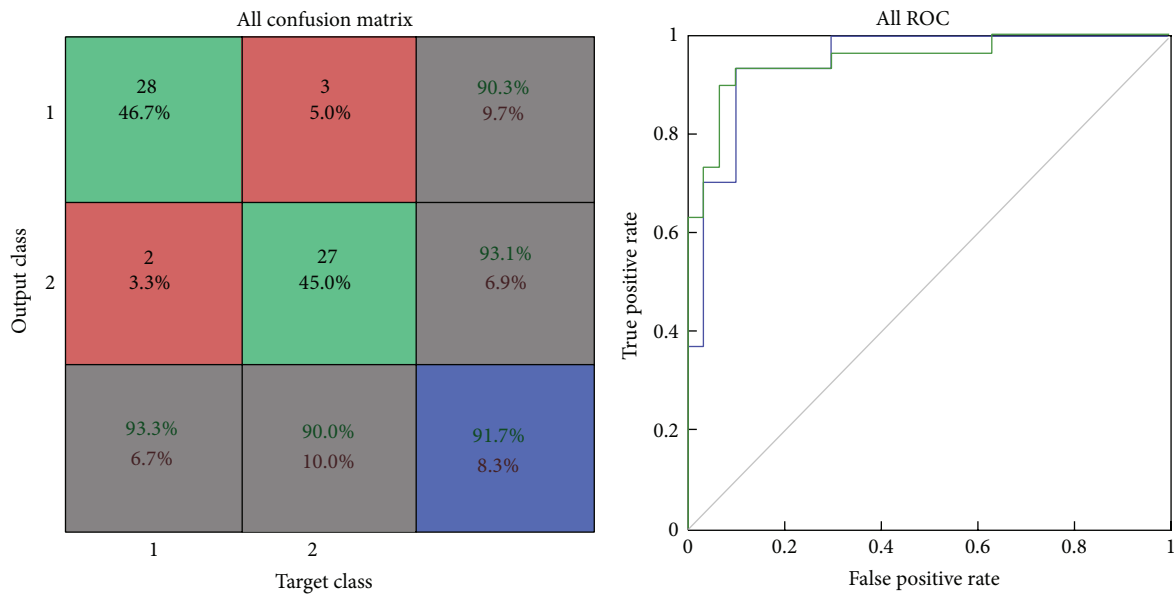FIGURE 6: Confusion matrix and ROC plot of GLRLM testing data.

FIGURE 7: Confusion matrix and ROC plot of mixed features training data.

correctly classified and 5 samples were misclassified by this network as shown in Figure 7. ROC graph shows the plotting of true positive rate against false positive rate (1-specificity). ROC graph of this network shows a perfect and effective classification between the two categories as the curves lie towards the upper-left corner as shown in Figure 7.

The results of the testing data show that mixed features yield an accuracy of 92.5%, sensitivity of 95%, specificity of 90%, false positive rate computed of 10%, false negative rate computed of 5%, and misclassification rate of 7.5%. Size of the input dataset loaded in the network was of 40 samples. Out of the total 40 samples, 37 samples were correctly classified and 3 samples were misclassified by this network as shown in Figure 8. ROC graph shows the plotting of true positive

rate against false positive rate (1-specificity). ROC graph of this network shows a perfect classification between the two categories, as the curves lie in the region of upper-left as shown in Figure 8.

The results showed that the selected mixed features yielded an accuracy of around 91.67% on the training set as compared to GLRLM features and GLCM features, which yielded an accuracy of around 90% and 86.7%, respectively, on the training set. The histogram feature gives around 77.5% and the invariant moments give around 65% of accuracy on training set as shown in Table 6.

These results indicate that the highest accuracy is achieved by the "mixed features set" with the accuracy being 91.7% by the "training dataset" and 92.5% by the "testing
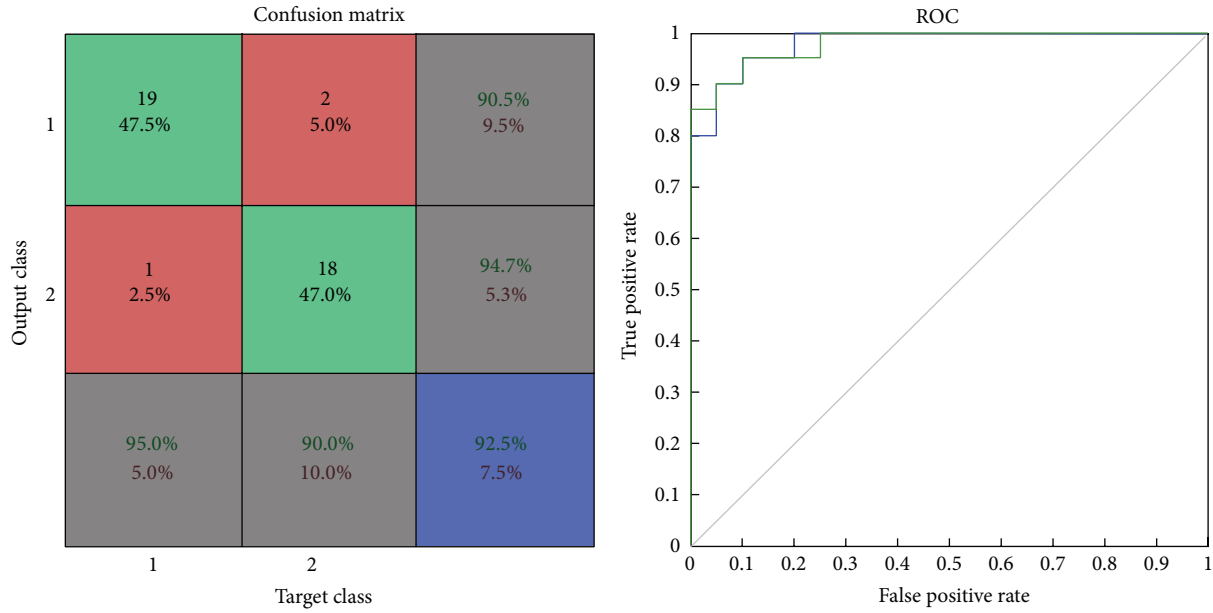
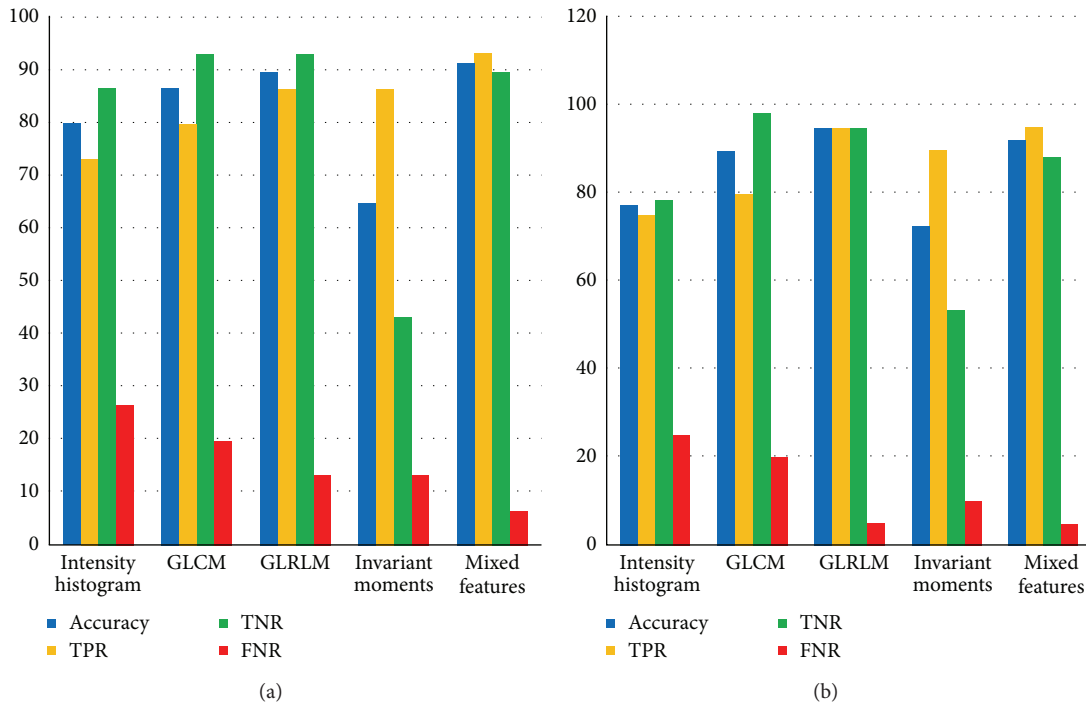Figure 8: Confusion matrix and ROC plot of mixed features testing data.

Figure 9: (a) Overall performance analysis of training dataset and (b) overall performance analysis of testing dataset.

dataset" as shown in Figure 9. Hence the selected features of the mixed feature set such as homogeneity (homom), sum of averages (savgh), difference variance (dvarh), information measure of correlation-1 (inf1h), information measure of correlation-2 (inf2h), inverse difference normalized (indnc), short run emphasis (SRE), short run high gray-level emphasis (SRHGE), and length (l) show excellent accuracy.

Along with the mixed features, the gray-level run length matrix features also showed an excellent accuracy. Training

set yielded an accuracy of 90%, but its testing dataset showed a better accuracy of 95%, that is, greater than both training and testing data accuracy of the mixed features. For testing the network against test data, GLRLM features gave better result. A low false negative rate decreases the possibility of misclassification. Both mixed and GLRLM features yielded a low false negative rate as shown in Table 6. The false positive rate of both mixed and GLRLM features is 5% and it is lower than the other networks.

TABLE 6: Performance analysis of all features.

FEATURE	Training (%)				Testing (%)			
	Accuracy	TPR	TNR	FNR	Accuracy	TPR	TNR	FNR
Intensity histogram	80	73.3	86.7	26.7	77.5	75	80	25
GLCM	86.7	80	93.3	20	90	80	100	20
GLRLM	90	86.7	93	13.3	95	95	95	5
Invariant moments	65	86.7	43.3	13.3	72.5	90	55	10
Mixed feature	91.7	93.3	90	6.7	92.5	95	90	5

8. Conclusion

In this study, five feature classifiers have been investigated for diagnosing the liver disease conditions. The accuracy of the classifier was based upon the feature set used, selected training samples, and the classifier's ability to learn from the training samples. From the above results, we have achieved our objective in finding the best classifier for liver disease diagnosis. Five sets of features such as GLCM, intensity histogram, GLRLM, invariant moments, and mixed features were extracted. These features were then selected and trained in neural network to determine the best set of features, which can determine the presence of disease conditions in the liver. A comparative approach revealed that both GLRLM and mixed feature set showed excellent accuracy in training as well as testing.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] A. Badawi and A. Youssef, "Tissue characterization of diffuse liver diseases using neural networks," in *Annual Meeting of Egyptian Society of Gastroenterology and Egyptian Society of Ultrasonography*, 1992.

[2] M. Singh, S. Singh, and S. Gupta, "A new quantitative metric for liver classification from ultrasound images," *International Journal of Applied Physics and Mathematics*, vol. 2, no. 6, 2012.

[3] P. Kumar, C. Prathap, and C. Dharshith, "An automatic approach for segmentation of ultrasound liver images," *Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 1, 2013.

[4] G. Li, Y. Luo, W. Deng, X. Xu, A. Liu, and E. Song, "Computer aided diagnosis of fatty liver ultrasonic images based on support vector machine," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '08)*, pp. 20–24, Vancouver, Canada, August 2008.

[5] C. Dietrich, C. Serra, and M. Jedrzejczyk, Ultrasound of the liver," EFSUMB—European Course Book.

[6] R. Suganya and S. Rajaram, "Content based image retrieval of ultrasound liver diseases based on hybrid approach," *American Journal of Applied Sciences*, vol. 9, no. 6, pp. 938–945, 2012.

[7] M. N. Hood and H. Scott, "Introduction to picture archive and communication systems," *Journal of Radiology Nursing*, vol. 25, no. 3, pp. 69–74, 2006.

[8] K. Kalyan, S. Jain, R. D. Lele, M. Joshi, and A. Chowdhary, "Application of artificial neural networks towards the determination of presence of disease conditions in ultrasound images of kidney," *International Journal of Computer Engineering & Technology*, vol. 4, no. 5, pp. 232–243, 2013.

[9] M. Seetha, I. Muralikrishna, B. L. Deekshatulu, B. L. Malleswari, and P. Hegde, "Artificial neural networks and other methods of image classification," *Journal of Theoretical and Applied Information Technology*, vol. 4, no. 11, p. 1039, 2008.

[10] J. Basu, D. Bhattacharyya, and T. Kim, "Use of artificial neural network in pattern recognition," *International Journal of Software Engineering and Its Applications*, vol. 4, no. 2, pp. 23–25, 2010.

[11] Y. M. Kadah, A. A. Farag, J. M. Zurada, A. M. Badawi, and A.-B. M. Youssef, "Classification algorithms for quantitative tissue

characterization of diffuse liver disease from ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 15, no. 4, pp. 466–478, 1996.

[12] J. H. Jeon, J. Y. Choi, S. Lee, and Y. M. Ro, "Multiple ROI selection based focal liver lesion classification in ultrasound images," *Expert Systems with Applications*, vol. 40, no. 2, pp. 450–457, 2013.

[13] W.-L. Lee, K.-S. Hsieh, Y.-C. Chen, and Y.-C. Chen, "A study of ultrasonic liver images classification with artificial neural networks based on fractal geometry and multiresolution analysis," *Biomedical Engineering—Applications, Basis and Communications*, vol. 16, no. 2, pp. 59–67, 2004.

[14] A. Plesea-Condratovici, C. Plesea-Condratovici, C. Neamtu, M. Banu, and M. Afteni, "Neural network application in Eco-ultrasound brightness prediction associated with hepatic diagnosis," in *The Annals of Dunarea De Jos University of Galati Fascicle V, Technologies in Machine Buiding*, pp. 15–18, 2012.

[15] M. Gletsos, S. G. Mougiakakou, G. K. Matsopoulos, K. S. Nikita, A. S. Nikita, and D. Kelekis, "A computer-aided diagnostic system to characterize CT focal liver lesions: design and optimization of a neural network classifier," *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 3, pp. 153–162, 2003.

[16] J. Brenard, *Digital Image Processing*, Springer, Berlin, Germany, 2005.

[17] N. Suematsu, Y. Ishida, A. Hayashi, and T. Kanbara, "Region-based image retrieval using wavelet transform," in *Proceedings of the 15th International Conference on Vision Interface*, pp. 9–16, Calgary, Canada, May 2002.

[18] M. Vasantha, V. Bharathi, and R. Dhamodharan, "Medical image feature, extraction, selection and classification," *International Journal of Engineering Science and Technology*, vol. 2, no. 6, pp. 2071–2076, 2010.

[19] S. Selvarajah and S. Kodituwakku, "Analysis and comparison of texture features for content based image retrieval," *International Journal of Latest Trends in Computing*, vol. 2, no. 1, 2011.

[20] M. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, pp. 179–187, 1962.

[21] N. Sharma, A. Bajpai, and R. Litoriya, "Comparison the various clustering algorithms of weka tools," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 5, pp. 73–80, 2012.

[22] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, no. 1, pp. 3–31, 2000.

[23] P. Andrey and T. Boudier, "Adaptive active contours," in *Proceedings of the ImageJ Conference*, 2006.

[24] A. Ahmadian, A. Mostafa, M. Abolhassani, N. Alam, and M. Gitti, *An Efficient Texture Feature Extraction Method for Classification of Liver Sonography Based on Gabor Wavelet*, Medicon, Tehran, Iran, 2004.

[25] M. Seetha, I. Muralikrishna, B. Deekshatulu, B. Malleswari, and P. Hegde, "Artificial neural networks and other methods of image classification," *Journal of Theoretical and Applied Information Technology*, vol. 4, no. 11, 2008.

[26] S. Poonguzhali and G. Ravindran, "Automatic classification of focal lesions in ultrasound liver images using combined texture features," *Information Technology Journal*, vol. 7, no. 1, pp. 205–209, 2008.

[27] M. Pietikäinen, T. Ojala, Z. Xu, and M. Pietikäinen, Rotation-Invariant Texture Classification Using Feature Distributions.

[28] J. Li, J. H. Cheng, J. Y. Shi, and F. Huang, "Brief introduction of back propagation (BP) neural network algorithm and its improvement," in *Advances in Computer Science and Information Engineering—Volume 2*, D. Jin and S. Lin, Eds., vol. 169 of *Advances in Intelligent and Soft Computing*, pp. 553–558, Springer, Berlin, Germany, 2012.

[29] M. Beale, M. Hagan, and H. Demuth, *Neural Network Toolbox User's Guide*, MATLAB, The MathWorks, 2013.