Abstract--In this paper, the self organization properties of genetic algorithms are employed to tackle the problem of feature selection and extraction in ultrasound images, which can facilitate early disease detection and diagnosis. Accurately identifying the aberrant features at a particular location of clinical ultrasound images is important to find the possibly damaged tissues. Unfortunately, it is difficult to exactly detect the regions of interest (ROIs) from relatively low quality of clinical ultrasound images. The presented evolutionary optimization algorithm presents a novel approach to building features for automatic liver cirrhosis diagnosis using a genetic algorithm. The extracted features provide several advantages over other feature extraction techniques which include: automatically construct feature set and tune their parameters, ability to integrate multiple feature sets to improve the diagnosis accuracy, and ability to find local ROIs and integrate their local features into effective global features. As compared with past approaches, we span a new way to unify the processing steps in a clinical application using the evolutionary optimization algorithms for ultrasound images. Experimental results show the effectiveness of the proposed method.

Index Terms—Feature construction, genetic algorithm, AdaBoost, region of interest, ultrasound image.

I. INTRODUCTION

The main steps in the automatic diagnosis process based on ultrasound images include data preprocessing, detection of regions of interest (ROI), feature extraction and selection, and classification. Data preprocessing aims at filtering speckle noise which is from the echo of non-uniform tissue in ultrasound imaging. The regions of interest (ROIs) objects in ultrasound images are often difficult to detect as they contain incomplete edges by interference of speckle noise. Therefore, it is important to select suitable image enhancement algorithm to improve the quality of the input ultrasound image for further process of ROIs detecting. The tasks of feature extraction and selection include measurements of size and shape of ROIs within the tissue, computation of texture features and elasticity indexes estimated from sequences of images [1]. Ultrasound image classification refers to annotating an image to a label among a predefined set of classes. The grade of liver cirrhosis is separated into 6 levels in the proposed system for liver disease diagnosis.

Many approaches including wavelet-based method [2], adaptive weighted median filter [3] and cubic-spline interpolation [4] have been proposed to filter speckle noise from ultrasound images. Based on the filtered images, segmentation of anatomical structures of interest from the input image can be performed manually or automatically using various approaches [5]. However, it is not a trivial work to segment the anatomical structures of interest from these images without priori tissue models. In contrast, many types of features including gray level distribution, intensity gradient, phase, fractal dimension, and texture can be estimated from ROIs of an ultrasound image [6-13]. In the case of liver disease identification, textures are the common features to characterize the input image. As an example, Basset et al. presented an ultrasound image interpretation system based on texture features to classify the severity of liver fibrosis [7]. Cao et al. proposed a liver fibrosis diagnosis system using the texture features characterized by fractal dimension and entropy [8]. Yeh et al. used the gray-level co-occurrence matrix (GLCM) extracted from the input ultrasound image to measure the grade of liver fibrosis [9]. The approach based on non separable wavelet transform (NSW) was also proposed to further improve the classification accuracy for liver fibrosis diagnosis [10]. However, the performance of these approaches would heavily depend on the locations of ROIs.

Clustering of estimated features into predefined classes may be performed using support vector machine (SVM), Fisher linear classifier, neural networks, Bayesian classifier, and so on. For example, Zaid et al. [11] used neural networks to classify the liver diseases in their clinical image application. Based on Gabor features extracted from ultrasound images, the AdaBoost classification scheme was used to improve the performance of liver cirrhosis diagnosis [12].
In this paper, we present an automatic liver cirrhosis diagnosis system using evolutionary feature construction. The presented evolutionary optimization algorithm presents a novel approach to building features using a genetic algorithm. In related work, the selection of effective ROIs for feature extraction is either through manual assessment or object segmentation. Automatic selection of effective ROIs without image segmentation is an important task for medical diagnosis as the shapes of anatomical structures of interest in captured ultrasound images are often vague. However, to our best knowledge, little work focused on this issue. In contrast, in this paper, the self-organization properties of genetic algorithms are employed to tackle the problem of ROIs selection. In the training phase, for each class, the parameters to generate effective ROIs are then learned and stored back to the database. In the classification phase, the ROIs and then their features are extracted from the input image based on these parameters. The extracted features provide several advantages over other feature extraction techniques which include: automatically construct feature set and tune their parameters, ability to integrate multiple feature sets to improve the diagnosis accuracy, and ability to find local ROIs and integrate their local features into effective global features. As compared with past approaches, we span a new way to unify the processing steps in a clinical application using the evolutionary optimization algorithms for ultrasound images. Experimental results show the effectiveness of the proposed method.

II. OPTIMIZING ROI SELECTION USING GENETIC ALGORITHM

![Diagram of the proposed ultrasound image interpretation system.](image)

The genetic algorithm (GA) [14] uses a meta-heuristic approach for solving hard combinatorial optimization problems. The idea of GAs came from an analog to biological evolution, in which provides an approach to learning. GAs represent hypotheses as chromosomes described by bit strings or symbolic expression. Rather than search from general-to-specific hypotheses, or from simple-to-complex, GAs generate an appropriate hypothesis by repeatedly mutating and recombining parts of the best currently known hypotheses. The search process begins with a population, or collection, of initial hypotheses. At each iteration, members of current population are evaluated relative to a given measure of fitness. The hypotheses of most fitness are selected probabilistically as seeds for producing the next generation. The process forms a generate-and-test beam-search of hypotheses, in which variants of the best current hypotheses are most likely to be considered next. A threshold defining an acceptable level of fitness for terminating the algorithm is used to stop the algorithm. Finally, GAs return the hypothesis in up-to-date population when the search process converges.

Figure 1 shows the block diagram of the system which involves the training phase and the classification phase. In the training phase, we automatically extract ROIs, optimized by GA, as speckle templates from ultrasound images in the patient database. The set of ROIs with the highest fitness is then used to train the SVM classifier for annotating the query image the grade of liver cirrhosis in the classification phase.

The first step, preprocessing, in the training procedure consists of three sub-steps: speckle noise filtering, scale-invariant feature transform (SIFT) [15], and feature extraction. The speckle noise filtering aims at eliminating the noise from the speckle template as much as possible without destroying the speckle pattern in the template. To achieve the goal, we choose the Gaussian filter to filter speckle noise. In computer vision, SIFT provides an approach to detect extreme feature points from an image to construct invariant descriptors with respect to position, scale, and rotation for object recognition. Instead of using the descriptors provided by SIFT, the feature extraction module detects the suitable features for medical diagnosis from the smoothed ultrasound images. Every feature point, detected by SIFT is indexed by the intensity gradient and defines a candidate 64x64 ROI with the feature point as the center. A feature vector to represent the speckle pattern of a ROI is then extracted from the ROI.

The texture features including the contrast, homogeneity, entropy, energy, and correlation can be extracted from the Grey Level Co-occurrence Matrix (GLCM) [13]. Given a specified displacement vector \((d_x, d_y)\) and a ROI \(R\), for each possible pair of grey levels \((g_1, g_2)\), the GLCM \(C_R(g_1, g_2)\) is defined as

\[
C_R(g_1, g_2) = \# \{(x, y), (x', y') \mid f_R(x, y) = g_1, f_R(x', y') = g_2, x = x' + d_x, y = y' + d_y\},
\]

where \(f_R(x, y)\) is the grey level of the pixel \((x, y)\) in \(R\) and \(#S\) returns the size of the set \(S\). Based on \(C_R\), the feature set \(F_R\) is used to characterize the content of the ROI \(R\). The information in \(F_R\) includes contrast, total energy, entropy, variance, correlation, local homogeneity, cluster shape,
clustering prominence, and fractal dimension. The effectiveness of $F_R$ for liver diseases diagnosis based on ultrasound images had been reported by the work of Mitrea et al. [13]. The details to compute these features from $C_R$ can refer to [13]. Instead of proposing new features for liver diseases diagnosis, in this work, we study the scheme to optimize the combination of ROIs when multiple ROIs are used in the interpretation of ultrasound images.

One disadvantage of using a single ROI for liver diseases diagnosis is that the system fails when the classification result using the low-level features extracted from the ROI does not agree with his high-level annotation. However, two ROIs are similar in terms of feature distance as their speckle patterns might be damaged by structural noises in relatively low quality of ultrasound images. To select an effective ROI is obviously not a trivial job as what structural noises included in the captured ultrasound images are usually unknown. In this work, we propose an evolutionary approach to optimize the selection of ROI using a genetic algorithm. Moreover, multiple ROIs are used to further enhance the performance of the liver diseases diagnosis.

Let $N$ ROIs be detected for an ultrasound image using the preprocessing step mentioned above. Notice that every ROI is associated with the intensity gradient of the center pixel and a feature vector $F$. The ROIs are sorted in the descending order according to the gradient values. The ROIs selection module uses an interval parameter to define the set $S$ of feature vectors for liver diseases diagnosis as follows

$$S = \{F_i | \nabla_{\min} \leq F_i \leq \nabla_{\max}, 0 \leq i < N\}, \tag{2}$$

where $F_i$ is the feature vector of the ROI indexed by $\nabla_i$. Obviously, the size of $S$ would depend on the interval $I$. This leads to the problem of un-balanced feature dimensions for different ROI combinations. To solve the problem, a simple dimension reduction scheme presented in our previous work [16] based on feature vector clustering is proposed to normalize the feature dimension for each combination of ROIs.

For each set $S$ of feature vectors, we compute its mean feature vector as follows

$$\mu = \frac{1}{|S|} \sum_{F \in S} F. \tag{3}$$

The distances between a feature vector $F_i$ and the mean vector $\mu$ in $S$ is defined as

$$d_i = D(F_i, \mu) \tag{4}$$

where $D(\cdot, \cdot)$ is the distance function between two feature vectors. We expect that the distance function $D()$ is non-negative, symmetric, and obeys the triangular inequality. Euclidean distance is used to meet the requirement above in this paper. The value of $d_i$ represents the scattering degree relative to the mean vector of the $i$th vector in $S$. The set $S$ is split if its distance variance is not small enough. To do that, the heart of the proposed method is to project the feature vectors in $S$ on a carefully selected line. We choose the vector $k$ with the largest value of $d_i$ (defined in Equation 4) in $S$ and the mean vector $\mu$, and consider the line $l$ that passes through them in the feature space. The projection value $p_i$ of the $i$th feature vector on the line is obtained by

$$p_i = d_i \frac{(F_i - \mu) \cdot (F_i - \mu)}{d_i} \tag{5}$$

Equation 5 maps feature vectors onto points on the line $l$, while preserving some of the distance information. For example, if the vector $i$ is reasonably close to the vector $k$, $p_i$ is small. Note that the value of $p_i$ is zero.

To separate the ROIs in $S$ into two clusters (Cluster 0 and Cluster 1), we use a one-bit scalar quantizer of the projection values. There are many existing methods for this purpose. The quantizer which uses the average projection value $p_{\text{avg}}$ as a threshold is one of the simplest. If the projection value is less than or equal to the value of $p_{\text{avg}}$ the feature vector is classified into Cluster 0, otherwise the vector is classified into Cluster 1. After classifying the feature vectors, two representative feature vectors $F_0$ and $F_1$ in a set are obtained by

$$(F_0, F_1) = (F_o = \frac{1}{N-q_{\text{notclassified}}} \sum_{F \in \text{Cluster}1} F, F_1 = \frac{1}{q_{\text{notclassified}}} \sum_{F \in \text{Cluster}0} F) \tag{6}$$

where $q$ is the number of vectors classified into Cluster 1 and $N$ is the size of $S$. Clustering of the feature vectors is repeated until the desired number of representative features are achieved. At least $K$-1 times of clustering are required if $K$ representative feature vectors are needed. After the clustering process is completed, all feature vectors in each cluster are regarded as identical and are represented by the mean vector of the cluster. Finally, all the representative feature vectors are cascaded to form the final feature vector for the corresponding ROI combination. Thus, the final feature vector to represent the input ultrasound image $U$ based on the interval parameter $I$ is defined as

$$F_I^U = F_0 \otimes F_1 \otimes \ldots \otimes F_K \tag{7}$$

where $\otimes$ is the cascade operator.

For each interval parameter $I$, we can construct a ROI combination from each database image and the representative feature vector for the ROI combination is obtained from (7). We divide the database images into two parts: the set of training images and the set of test images. The collection of representative feature vectors for the training images are used to learn a SVM classifier. To apply a genetic algorithm for the optimization of ROIs selection, the parameter $I = [\nabla_{\min}, \nabla_{\max}]$ is encoded as a chromosome which is the concatenation of $\nabla_{\min}$ and $\nabla_{\max}$. Next a fitness score is computed. The fitness score reflects how well the SVM classifies the test set using Equation 8 where $f_p$ is number of true positives, $f_n$ is number of false negatives, $t_p$ is the number of true negatives, and $f_n$ is the number of false positives. The fitness score is a real number in the range $[0,1]$. 

3
Equation 8 includes two terms to reduce the sensitivity to unbalanced numbers of negative and positive training examples. A fitness score based on classification accuracy would favor a SVM classifier to classify everything as negative, although it has no ability to discriminate positive examples from negative examples, when the negative examples are far more than positive examples in the training set. Whenever the fitness score of a chromosome is over a threshold it is added to a pool from which the selection module of the genetic algorithm update the chromosome population.

$$Fitness = \frac{t_p}{f_s + t_p} + \frac{t_n}{f_p + t_n}$$  (8)

III. THE CLASSIFICATION METHOD

Algorithm 1. Optimization of ROIs Selection

Initialize $P(1)$ of chromosomes.
Evaluate $P(1)$.

for $t = 1, 2, ..., T$
do
   for every chromosome do
      for every training image do
         Process image with feature construction
         Train chromosome’s SVM classifier
      end for
   end for
   for every test image do
      Process image with feature construction
      Use the output of SVM classifier to update fitness score
   end for
   Assign fitness score to the chromosome
end for

Recombine $P(t)$ to yield $C(t)$ using crossover and mutation.
Evaluate $C(t)$
Select $P(t+1)$ from $P(t)$ and $C(t)$.
end while

Let the maximal iteration to stop the genetic algorithm be $T$. The SVM classifiers with larger fitness scores in current population $P(T)$ are returned from the genetic algorithm to construct the classification bank for classifying the query image. Based on AdaBoost [17], algorithm 2 outlines how to select the SVM classifiers. $N$ represents the maximum number of SVM classifiers allowed in the final model, which is 5 in our case. The number of SVM classifiers in the model can be increased after training if desired. The resulting model is a list of SVM classifiers and coefficients that indicate how much to trust each classifier.

Algorithm 2. SVM Classifiers Selection.

Input: Sequence of $M$ examples $S = \{(x_i, y_i)\}, i = 1, \cdots, M$ with labels $y_i \in \{a_1, a_2, \cdots, a_K\}$; SVM classifiers with respect to current population $P(T)$;

Initialize: $D_i(i) = 1/M$, $i = 1, \ldots, M$.

for $t = 1, 2, \ldots, N$
do
   Remove the SVM classifier, $h_i$, with the largest fitness score from current population $P(T)$.
   Select a training data subset $S_t$ drawn from the distribution $D_t$.
   Calculate the error of $h_i$:
   $$e_i = \sum_{(x_t, y_t) \in S_t} D_t(i)$$
   if $e_i > 50\%$, break.
   Set $\beta_i = e_i/(1-e_i)$;
   Set $\alpha_i = \log(1/\beta_i)$;
   Update distribution:
   $$D_{t+1}(i) = \frac{D_t(i)}{\sum D_t(i)} \begin{cases} \beta_i, & h_i(x_t) = y_t \\ 1, & \text{otherwise} \end{cases}$$
end for
In the classification phase, shown in Fig. 1, the quality of a query image is first slightly improved by a Gaussian filter. For each SVM classifier, the stored parameter, gradient interval, is used to select the ROIs from the query image. Once ROIs of a query image is obtained, the features extracted from the ROIs are used to assess the degree of liver cirrhosis using the classifiers trained in the training phase. The normalization procedure mentioned in the previous section is also performed to construct the final feature vectors using (7). Multiple ROIs are extracted from a query image to judge what class the query image belongs to. More precisely, the part of a query image used to evaluate the specific degree of liver cirrhosis is adaptive according to the domain knowledge obtained from the training procedure.

As mentioned in previous section, we use $N$ SVM classifiers, selected by algorithm 2, to construct the final classification model. Thus, for each query image $q$, $N$ feature vectors are obtained for classification. Once a preset $N$ number of classifiers are generated, AdaBoost is ready for classifying the query image. Following the idea of AdaBoost, the weighted majority voting defined in Equation (9) is used to classify the query image. The weighted majority voting, corresponding to the total vote received by each class, is defined as

$$V_i = \sum_{t=0}^{N} \alpha_t, i = 1, \ldots, C$$

where $C$ is the number of classes. The effectiveness of the weighted majority voting is simple: those classifiers with good performance during training are rewarded with higher voting weights than the others. Recall that the voting weight $\alpha_t$ of classifier $h_t$ has a larger value when the corresponding normalized error $\beta_t$ obtained from Algorithm 2 is smaller. The value of $\alpha_t$ is therefore a measure of performance of $h_t$, and can be used to weight the classifiers. The logarithm of $1/\beta_t$ is usually used as the voting weight of $h_t$ to avoid the values of vote weights from asymptotically large caused by small training errors. At the end, the class that receives the highest total vote from all classifiers is the ensemble decision.

**IV. EXPERIMENTAL RESULTS**

The performance of the proposed system is evaluated with various test images, where each of them is manually labeled by a degree of liver cirrhosis. The patient database with 607 test samples is separated into two parts: the training database and the test database. These ultrasound images with 330x380 pixels and 256 grays were obtained using Toshiba SSA-700A Aplio50 B-mode ultrasound imaging system with different tissue harmonic 5.0 and 16 fps. Furthermore, ground truth of the patient database for performance evaluation is obtained through manually annotating the grade of liver cirrhosis for each ultrasound image by doctors from Department of Internal Medicine, Chang Gung Memorial Hospital, Keelung, Taiwan. The grade of liver cirrhosis for each ultrasound image is from 0 to 5. Level 0 indicates that the input ultrasound image is normal. Level 5 is the most serious in liver cirrhosis.

Figure 3 shows the GLCMs obtained from the representative ROIs for both clusters obtained by performing the proposed data reduction on an input image. These two matrices are obviously different according to the subjective evaluation. To further verify the effectiveness of the proposed method, the correct classification rate (CCR) defined as $C/N$, where $C$ and $N$ are the number of correct classification and the number of test samples, respective are used for performance comparison. The methods proposed by Cao et al. [8], YEH et al. [9], and Mitrea et al. [13] are also implemented for performance comparison. For each test image, the ROI for the compared methods is manually selected since the automatic ROI selection scheme is not provided in their approaches. In contrast, the proposed method automatically selects ROIs to provide an automatic mode for liver cirrhosis diagnosis. As shown in Tables 1 and 2, we test the performance for all methods based on the same set of test samples. Given a query image, in two (six) classes classification mode, the average values of $CCR$ to classify the implicit level of cirrhosis of the image are 56 (17), 81 (41), 60 (48) and 91 (67) for Cao’s, Yeh’s, Mitrea’s and proposed methods, respectively. Accordingly, the performance of the proposed method is much better than that of the compared methods. The classification accuracy for
some classes is not good since that few samples are provided in those classes.

Table 1. Performance comparison among Cao’s, Yeh’s, Mitrea’s, and proposed methods for two class classification in terms of CCR.

<table>
<thead>
<tr>
<th>Two classes</th>
<th>No cirrhosis</th>
<th>With cirrhosis</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cao’s method [8]</td>
<td>71</td>
<td>40</td>
<td>56</td>
</tr>
<tr>
<td>Yeh’s method [9]</td>
<td>86</td>
<td>76</td>
<td>81</td>
</tr>
<tr>
<td>Mitrea’s method [13]</td>
<td>36</td>
<td>84</td>
<td>60</td>
</tr>
<tr>
<td>Proposed method</td>
<td>Without GA</td>
<td>92</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>With GA</td>
<td>93</td>
<td>88</td>
</tr>
</tbody>
</table>

Table 2. Performance comparison among Cao’s, Yeh’s, Mitrea’s, and proposed methods for 6 classes classification in terms of CCR.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without GA</td>
<td>With GA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 0</td>
<td>64</td>
<td>86</td>
<td>36</td>
<td>71</td>
</tr>
<tr>
<td>Level 1</td>
<td>6</td>
<td>36</td>
<td>35</td>
<td>46</td>
</tr>
<tr>
<td>Level 2</td>
<td>0</td>
<td>46</td>
<td>69</td>
<td>77</td>
</tr>
<tr>
<td>Level 3</td>
<td>14</td>
<td>0</td>
<td>52</td>
<td>31</td>
</tr>
<tr>
<td>Level 4</td>
<td>20</td>
<td>10</td>
<td>20</td>
<td>70</td>
</tr>
<tr>
<td>Level 5</td>
<td>0</td>
<td>64</td>
<td>76</td>
<td>36</td>
</tr>
<tr>
<td>Average</td>
<td>17</td>
<td>41</td>
<td>48</td>
<td>55</td>
</tr>
</tbody>
</table>

**V. CONCLUSIONS**

In this paper, we have presented the method to construct an automatic system for liver disease prediction. The contributions of the paper include: (1) a preprocessing method is proposed to extract effective features from the selected ROIs of the input image; (2) using a genetic algorithm, a training procedure is proposed to optimize the selection of ROIs and to obtain the stable parameters for classification; (3) an AdaBoost based classification approach is presented to predict the possibility of liver cirrhosis for a patient based on the patient database and the input ultrasound image. As compared with past approaches, we span a new way to unify the processing steps in a clinical application using the evolutionary algorithms for interpreting ultrasound images.

Future work would be on increasing the size of patient database and generalized the proposed method to deal with ultrasound image sequences. Both of them would improve the prediction accuracy of the system.

**REFERENCES**


