Multi-Stage Optimization Support Vector machine by Analysis of Non-Linear Normalization Texture for Early Evaluation Staging Liver Fibrosis

NASROLLAH GOUDARZI[1], REZA FAGHIHI[1,2]
[1]Nuclear Engineering Department School of Mechanical Engineering, Shiraz University, Shiraz, Iran
[2]Radiation Research Center of Shiraz University, Shiraz, Iran
Correspondence to Nasrollah Goudarzi, Email : Goodarzi.n5@gmail.com

ABSTRACT

Background: A reliable non-invasive approach is still in development to determine precisely the stage at which fibrosis is located. In this study, we used a Computer-aided diagnosis (CAD) to analyze magnetic resonance imaging (MR) images in a human's body to determine the stage of liver fibrosis with relative precision in CAD. The histopathological patterns of liver fibrosis are one of the most important biomarkers for the diagnosis and classification of chronic liver diseases, from the first stage to the end, on radiography in magnetic resonance imaging (MR) images.
Aim: To investigate various types of data obtained from MR images for the selection of optimal parameters and characteristics along the optimized stratification of fibrosis.
Method: A total of 250 patients were scanned with 1.5T and 3T superconducting scanners for abdominal examinations. All cases were examined by needle biopsy, as a golden standard in our test, and range from 0 (without fibrosis) to 5 (cirrhosis). For each case, at least four sequential fuzzy images are obtained by the MRI scanner: before contrast, arterial phase, gateway phase and balance. In both imaging methods, 18 texture features have been calculated from a gray level coincidence matrix (GLCM), which is extracted with a ROI in the liver as a set of input vectors. Each combination of these subset assays is evaluated using a two-stage optimized support vector machine (SVM) and a left-side method to diagnose fibrosis in non-cirrhosis or cirrhosis.
Result: Findings : Regardless of which method is used, it can be said that according to the accuracy (AR) calculated from each compound, the desired number of texture features for classification the degree of liver fibrosis is in the range of 4 to 7 . The overall performance calculated by the total sum of the mean maximum AR value belonging to all 18 characteristics in MRI images is 68.14% and 71.98%, respectively. Moderate bruising and entropy, among the 18 features of texture, are among the most common features in all three sets of visual data. The correlation property has the lowest AR value and is excluded from all data sets as an effective feature.
Conclusions: Conclusion: Evaluation of the balance of MR images in CAD can determine the effective stage in which the human liver fibrosis is located. This article describes the progress made in developing a non-invasive approach based on the design of advanced classification for determining the stage of liver fibrosis. Classification accuracy was avoided by implementing a two-stage SVM optimization algorithm in MRI imaging. Comparison of AR performance with 18 selected texture features revealed that MRI 3T images for liver fibrosis classification were better than MRI 1.5 T is. Finally, it can be said that texture analysis in the balance phase is more effective than any other fuzzy image.
Keywords: Liver fibrosis, non-linear normalization texture, MR images, MRI

INTRODUCTION

People who suffer from chronic liver disease are at risk for horrific complications, such as hepatocellular carcinoma (HCC) and hepatic impairment[1]. In diagnosis and evaluation of hepatic cirrhosis and chronic hepatitis, it is necessary to determine the degree of liver fibrosis as one of the most important cirrhosis indices and also one of the most important factors for predicting HCC incidence[2]. HCC is one of the most common malignancies in people affected by these diseases[3]. Liver fibrosis is a texture repair process that responds to liver damage. The fibrosis is characterized by collagen extracellular accumulation and proteins in the colon. This extracellular aggregation fills up empty space due to damage to the original texture. If the damage sustained in the long term is stimulated due to different causes, the extracellular matrix cannot be degraded or reduced. Fibrosis reversibility depends on determining its stage. Studies have shown that garlic liver fibrosis can be reversed in the early stages (F1-2) by anti-fibrotic

treatment. Fibrous, moderate (F3), is only partially reversible, and fibroblasts (F4), advanced fibroids, fibrosis and cirrhosis (F5) are irreversible[4,5]. Therefore, early diagnosis and precise determination of the stage in which liver fibrosis is located is important. The evaluation of liver histology is a reliable approach to determining the stage of fibrosis, but at the same time it should be noted that liver biopsy is required in this method that is invasive and cannot be routinely used in clinics[6]. Non-invasive methods include magnetic resonance techniques (MRs) that are used to determine liver fibrosis[7]. Among the MR elastographygraphy (MRE) techniques, the MR is a novel technique for quantitatively assessing the mechanical features of textures in the laboratory environment. The MR elastographygraphy is performed using a vibration source to generate low-frequency mechanical waves in the texture. The imaging of the waves created using the MRI contrast technique is performed, and then the wave information is processed to produce quantitative images that show mechanical features, such as texture hardness. Studies

show that MR elastographygraphy is an accurate method for detecting and determining the stage of liver fibrosis[8,10]. However, MR elastographygraphy are subject to limitations in assessing liver fibrosis in patients with moderate to severe iron overload that causes hemoglochromatosis or hemosiderosis. Because in this case, the liver signals are so low that the waves cannot be adequately conceived. Also, MR elastography has some shortcomings, including high costs, limited access to a MR unit, and a longer review period[9,10,11]. It should also be noted that MR imaging has been refined by hepatocyte-specific contrast and resonant imaging has been used to determine the stage in which liver fibrosis is present[12,13]. However, these MR methods still do not have the precision to meet the clinical needs of liver fibrosis.

Recently, MRIs have been paired with computer-assisted diagnostic technology to diagnose liver fibrosis. Several studies have examined the accuracy of the determination of the stage of liver fibrosis by the approach that combines MRI imaging with CAD. Cato and colleagues have come to the conclusion that CAD-based texture analysis in images recovered from gadolinium by the balance phase is a strong application for the diagnosis of liver fibrosis at a time when different methods have been used[14]. They also found that CAD analysis for liver texture features that were reflected in MR images enhanced by contrast for hepatocytes can be used to distinguish F3-4 from F1-2, yet these analyzes cannot have Difference between F0, F1 and F2[15]. House and colleagues stated that the specific features of the extracted texture from unburdened T2 images that are lacking in regeneration can separate the normal texture of the liver from the fibrotic texture of the liver, with the difference that there is no clear separation between F3 and F1-2[16]. The method presented by Sela et al was developed automatically by using a backward-vector multi-storey binary vector classification model that evaluates fibrosis in f-MRI images of rat liver. Slowly and achieves precision 9.96% and 72.7% for F0 versus F1-5 and F1-3 versus F4-5 [17]. Zheng and colleagues performed the classification accuracy of 85.0%, 66.0% and 70.0% for F0, when they performed SVM-based classification analyzes for texture features in portrait imagery improved by gadolinium. F0-2 versus F5, F0 versus F1-2 and F1-2 versus F3-4[18].

Computational intelligence methods have been used to explore predictive models for detecting hepatocellular carcinoma in patients with liver cancer. Evolutionary algorithms are used as efficient and robust methods for the evolution of the best-fitted parameter values that optimize the selection of features. Different computational methods have been developed to explore a more robust set of molecular features for liver fibrosis. These methods include those that incorporate other nature-inspired algorithms, including particle swarm optimization, into categories such as support vector machine (SVM). The aim of this study was to improve the accuracy of the determination of the stage of liver fibrosis in MR images that have been optimally optimized in conjunction with CAD technology. Previous studies have shown that the fibrotic feature can show differences in the images[14,18,19]. Therefore, in this study, we use a CAD system to analyze the balance-phase images from human beings at various stages of liver

fibrosis. The optimized support vector machine algorithm at the two-level was used as a nonlinear normalization for compression of images and features of the feature embedded in the Gray level co-occurrence matrix (GLCM)[20]. Because pixel structure analysis uses a series of mathematical models that extracts parameters from gray levels of pixels in digital images that the radiologist's eyes cannot detect, this method has the potential to help detect the radiologist unnecessarily. There are other costly methods of imaging or biopsy. The aim of the present study was to investigate the diagnostic value of computerized characterization of the structure of the feature for classification the stages of liver fibrosis. However, a very important step in this work is the method of choosing a feature property that is often overlooked. Current methods use any of the uncontrollable or monitored monitoring options. When a linear property selection technique is used, the features selected linearly with the class labels (maximum association), while correlating with each other (at least redundancy), are related to the input property space. However, this set of selected features is not linearly the best option for a non-linear classification, which maps the input property space to a higher 3D space (shown as kernel space). Accordingly, for nonlinear classification, it is best to choose the characteristics that lead to the best performance in the nonlinear kernel space. On the other hand, the selection of the mapped function for the input space is higher in the three dimensional space, which depends on the application and its related data types. These nonlinear predictions are often modeled using 8 kernel functions. In this paper, we propose a new strategy for selecting features that can create a classification in the core space. Thus, if we use a nonlinear kernel, the features selected in the nonlinear kernel space show the best performance, or if a linear kernel is used, the characteristics for the linear classification will be the best. To use two types of kernels, we define a new kernel that separates both linear and non-linear kernels, so they are used from both types of models. Accordingly, in this article we will consider the following: (1) The parameters (1) of our proposed method will select the characteristics that best classify data in core space, through the learning of multiple cores; (2) optimization of parameters The kernel functions in the SVM classification are accomplished by two steps. (3) We design the kernel functions in a way that proposes a nonlinear mixing strategy for determining the classification of liver fibrosis.

Tools and Materials:

Magnetic Resonance Imaging: Since February 2015 to March 2017, a total of 160 patients were hospitalized in Abu Ali Sina Hospital in Shiraz using a MR 3T scanner in a 6-channel spatula arrangement under ventricular examination. The imaging protocols used include TR/TE, 2.3/82 milliseconds in the plane scanning plane :7mm slice thickness and 1 mm inter-slice gap, the internal imaging matrix, 512 × 512, the slice thickness of the 4 mm and the inter-slice gap was 2 mm. In the phase of dynamic contrast control: 7 mm slice thickness and inter-slice gap1 mm ; contrast agent injection; Gd-DTPA; dose equal to 0.2 mmol/kg; and injection rate of 2 to 4 ml/s Was made. After 16 seconds, 50 to 60 seconds and 2 to 4 minutes after the contrast agent injection, scan the arterial phases, gateway

and balance. Among the data sets, there are 18 patients without a history of liver disease and liver dysfunction, and in other patients, the fibrosis stage is confirmed by the result of liver biopsy. A total of 50 other hospitals were located in the Dena hospital, located in Shiraz province. These items were obtained using a MR, 1.5T MR scanner and included in our tests. The weighted T1 image sequences obtained by Gadolinium (1.615.16, 512 × 512 matrix, 90 degree flip angle, unit acquisition, 1, 18 for 26 seconds of acquisition time), with a slice thickness of 8 mm and inter-slice gap of 2 mm. This study was approved in both hospitals and received written consent from all patients.

EXPERIMENTAL MATERIALS

To evaluate the contribution of texture features after a MRI scan, a total of 250 patients who suffered or lacked liver fibrosis were analyzed in a hospital, the stage where liver fibrosis was located. And in accordance with the reference guidelines established by the program for the prevention and treatment of viral hepatitis (VHPTA). In these studies, 250 items were scanned by the MR.

Table 1. the patients' statistics with regard to VHPTA system, in five stages of fibrosis.

Stage of fibrosis	Score	Description of hepatic fibrosis Degree	Number of MRI cases Abu Ali Sina's hospital	Number of MRI cases Dena hospital	Groups
Normal	0	No fibrosis	41	4	Non-cirrhosis (Group 0)
Mild fibrosis	1	Fibrous expansion of some portal areas, with or without short fibrous septae	28	2	
	2	Fibrous expansion of most areas and bridging fibrosis appeared	20	12	
Severe fibrosis	3	Most fibrous septum and lobular structure disorder	14	10	Cirrhosis (Group 1)
	4	Early cirrhosis, diffuse fiber Hyperplasia	18	19	
Cirrhosis	5	Cirrhosis	39	3	

ROI selection: We opened DICOM files on a server for radiologists, and it was made possible for them to choose ROIs by moving and clicking on the mouse and automatically calculating the texture features along with the ROI and Save them as an Excel file for further analysis. In this experiment, ROIs were selected manually and sporadically from each consecutive image. Radiologist with over 20 years of history of radiological imaging has been responsible for this work. ROIs are shown as red squares in Figure 1. ROIs are depending on the instructions shown. Due to the distribution of liver dispersed, due to the distribution of liver fiber fusion, large blood vessels and hepatocytes in the liver are eliminated. Three sizes of ROI (32 × 32, 20 × 20, 16 × 16 pixels) are generated individually and according to the starting point of the square, and all ROIs are preprocessed in the presence or absence of Sobel filters. Due to technical problems and the subject matter in this paper, only one of the slice thicknesses is used in MRI.

The CAD system: Figure 1 shows the process of determining the stage of fibrosis by the CAD system. First, ROI is manually selected. Next, raw data is normalized. Then, feature property vectors are extracted. Finally, support vector machine is optimized. Two classes are used separately to classify liver fibrosis.

Fig. 1: Flow diagram of living liver fibrosis with CAD analysis of MRI.
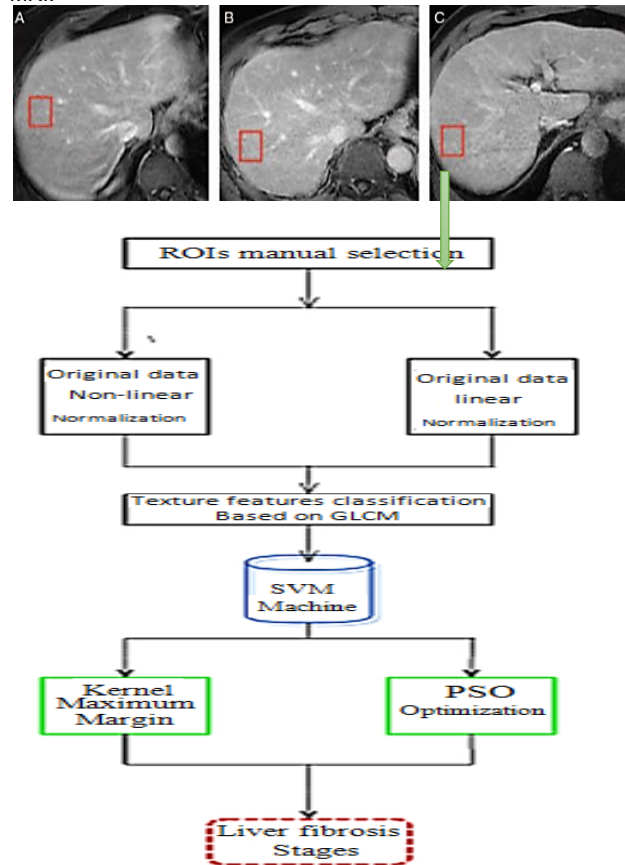
3.3 Gray level co-occurrence matrix and texture features

GLCM [21,22] is a matrix of relative frequencies Pij (d, θ), in which two adjacent pixels on the image are separated by d and θ. One of them has a gray level i and another with a gray level j. In this paper, a total of 18 characteristics were investigated. These features include: auto correlation, contrast, correlation, cluster superiority, cluster differences, heterogeneity, energy, entropy, sexuality, maximum probability, mean total, variance, total variance, difference variance, total entropy, entropy difference , Measurements of correlation information 1, correlation information measurement 2, normalized inverse correlation and normalized inverse difference torque. Each feature property was extracted from four directions (θ = 0, 45, 90, 135,180), respectively. Then 80 texture features were extracted for each ROI. The texture features are defined by the following equations:

Auto-correlate:
Contrast :

$$f_1 = \sum_i \sum_j (ij) p(i, j) \qquad (1)$$

Correlation:

$$f_2 = \sum_{n=0}^{G-1} n^2 \left\{ \sum_{i=1}^{G} \sum_{j=1}^{G} p(i, j) \| i - j \| = n \right\} \qquad (2)$$

Cluster superiority

$$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \qquad (3)$$

Cluster differentiation

$$f_4 = \sum_i \sum_j \left(i + j - \mu_x - \mu_y \right)^4 p(i, j) \qquad (4)$$

Dissimilarity:

$$f_5 = \sum_i \sum_j \left(i + j - \mu_x - \mu_y \right)^3 p(i, j) \qquad (5)$$

Energy:

$$f_6 = \sum_i \sum_j |i - j| \, p(i, j) \qquad (6)$$

Antropy:

$$f_7 = \sum_i \sum_j p(i, j)^2 \qquad (7)$$

Homogeneity:

$$f_8 = \sum_i \sum_j p(i, j) \log(p(i, j)) \qquad (8)$$

Maximum likelihood:

$$f_9 = \sum_i \sum_j \frac{p(i, j)}{1 + (i - j)^2} \qquad (9)$$

Variance:

$$f_{10} = \underset{i,j}{\mathrm{MAX}} \, (p(i, j)) \qquad (10)$$

Sum of the average:

$$f_{11} = \sum_i \sum_j \left(i - \mu_m \right)^2 p(i, j) \qquad (11)$$

μ_m is the average of $p(i, j)$

Sum of the antropy

$$f_{12} = \sum_{i=2}^{2G} i p_{x+y}(i) \qquad (12)$$

$$P_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{G} \sum_{j=1}^{G} p(i, j), \quad k = 2, 3, \cdots, 2G$$

Sum of the variance

$$f_{13} = -\sum_{i=2}^{2G} P_{x+y}(i) \log \left\{ P_{x+y}(i) \right\} \qquad (13)$$

Variance difference:

$$f_{14} = \sum_{i=2}^{2G} \left(i - f_{13} \right)^2 P_{x+y}(i) \qquad (14)$$

:
Antropy difference:
:

$$f_{15} = \text{variance of } p_{x-y} \qquad (15)$$

$$P_{x-y}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{G} \sum_{j=1}^{G} p(i, j), \quad k = 0, 1, \cdots, G - 1$$

The data measurement with respect to alliance 1

$$f_{16} = -\sum_{i=0}^{G-1} P_{x-y}(i) \log \left\{ P_{x-y}(i) \right\} \qquad (16)$$

:
The data measurement with respect to alliance 2:

$$f_{17} = \frac{HXY - HXY1}{\max\{HX, HY\}} \qquad (17)$$

:
The reversed normalized difference

$$f_{18} = \sqrt{1 - \exp[-2.0(HXY2 - HXY)]} \quad (18)$$

$$HXY = -\sum_i \sum_j p(i, j)\log(p(i, j)),$$

$$p_x(i) = \sum_{j=1}^{G} p(i, j), \quad p_y(j) = \sum_{i=1}^{G} p(i, j)$$

$$HXY1 = -\sum_i \sum_j p(i, j)\log\left\{ p_x(i)p_y(j) \right\},$$

$$HXY2 = -\sum_i \sum_j p_x(i)p_y(j)\log\left\{ p_x(i)p_y(j) \right\}$$

Normalized inverse difference torque:

$$f_{19} = \sum_{i=1}^{G}\sum_{j=1}^{G} \frac{p(i, j)}{1 + |i - j|/G^2} \quad (19)$$

$$f_{20} = \sum_{i=1}^{G}\sum_{j=1}^{G} \frac{p(i, j)}{1 + (i - j)^2/G^2} \quad \mathbf{20}$$

Classification of SVM Model: Selecting a feature or variable as the process of selecting a sub-set of distinctive features is used to build the best model, for example, for classification or regression. As previously supervised, feature selection approaches can be classified in two main categories: unsupervised and supervised. Previously, the features were chosen without regard to the marks of that class, while later the choice of features was based on the highest association with the signs and features of that particular class and the least amount of the selected features. Among the methods for selecting the monitored feature, in recent years the choice of scattered and weighted features [23] has attracted the most attention, because this is a simple procedure and has a better performance.

Suppose that there are N experimental samples, each of which has characteristic d space. We can sort the property vectors in a $X \in \mathbb{R}^{d \times N}$; matrix and arrange their particular marks $y \in \mathbb{R}^{1 \times N}$. Selecting a dispersed feature to minimize goals is then:

$$\min_{\mathbf{w}} \|\mathbf{y} - \mathbf{w}^{\mathrm{T}}\mathbf{X}\|_2^2 + \lambda\|\mathbf{w}\|_1, \quad (21)$$

In order to better represent the samples, we use the weight vector $\mathbf{w} \in \mathbb{R}^d$. This weighing vector is considered as a ℓ_1 norm to have a compact and dense set of features. Clearly, the set of features selected under these settings would be appropriate if we could design a linear classification model (such as linear SVM), which is because these features are designed to minimize and maximize Relationship with the marks of that class is selected in the main feature space. However, for a nonlinear classification, we must select features that multiply minimization and maximize the same relationship in the kernel space.

Nonlinear classification is usually obtained by first plotting the main property vectors in a large-scale space using a nonlinear mapping function (referred to as $\phi(\cdot)$) and classification is then done in the kernel space. Although, instead of explicitly mapping a function, a kernel function is often used that defines the similarity of the specimen (or feature vectors)[24]. This is because many learning machine algorithms (such as SVMs) can be expressed entirely in terms of dot products, and under certain conditions the kernel functions can be defined as dot products (infinite dimensions) in one Feature space is expressed.

Here, in order to select features and classify the data, we adopt a similar formula to the kernel-based SVM with a specific kernel to account for the feature selection. We suggest the use of the kernel function on each single feature and the definition of the aggregated or dense kernel through the total weight of all these kernels. In this way, we can select the features that have the greatest role in making a better classification model by adjusting l1 on the kernel weight vector.

Geometry modeling kernel-based feature selection : in the core kernel-based SVM formulation there is a Max-Margin decision range in kernel space that uses the loss function L with a hyper parameter C, exchange:

$$\min_{f \in \mathcal{H}} \frac{1}{2}\|f\|_{\mathcal{H}}^2 + C\sum_{j=1}^{N} L(y_j, f(\mathbf{x}_j)), \quad (22)$$

f is related to the production space of the kernel Hilbert $\mathcal{H}$, and $\mathbf{x}_j$ and y_j correspond to the jth sample and the corresponding sign. In terms of the nonlinear non-linear function, f is defined as $f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}) + b$. As. As we discussed earlier, we do not directly use the mapping function φ (0). For this purpose, Theorem 35 integrates, according to the case, the answer for w can always be presented as a linear combination of experimental data. So :

Fig. 2. (a) the conventional Kernel functions structure .(b) the implied kernels for every single characteristic, through the correlated weights with respect to every accumulated characteristics. $k_m(.,.)$ Shows the kernel function applied for the characteristics.
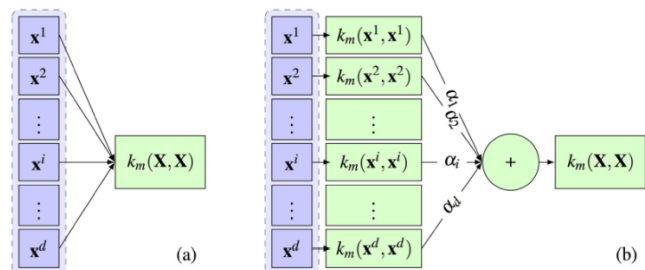
$$f(\mathbf{x}) = \sum_{j=1}^{N} \beta_j y_j \phi(\mathbf{x}_j)^T \phi(\mathbf{x}) + b = \sum_{j=1}^{N} \beta_j y_j k(\mathbf{x}_j, \mathbf{x}) + b.$$
(23)

So the optimization problem for SVM is reduced to solve $\beta_j, \forall\, j$. Here, we tend to choose the features that can make the best nonlinear SVM model. Accordingly, we create a kernel for each feature, and then we collect all of these kernels by weighted average. Therefore, the kernel is calculated between the two samples $\mathbf{x}_j$ and $\mathbf{x}$.

$$k(\mathbf{x}_j, \mathbf{x}) = \sum_{i=1}^{d} \alpha_i k(x_j^i, x^i).$$
(24)

This relationship is similar to the [25,26] learning framework of the multiple kernel, while making each kernel only on a single feature. By using the new kernel k on all experimental samples, we will have $\mathbf{K} = k(\mathbf{X}, \mathbf{X}).$

Fig. 3 illustrates the process of making the kernel matrix for all samples (2b) compared to the custom kernel build (2a). Now we can reinforce the choice of the most distinctive features in the kernel space by setting the **α** weigh-vector, so the target function will be:

$$\min_{f \in \mathcal{H}, \alpha} \frac{1}{2}\|f\|_{\mathcal{H}}^2 + C\sum_{j=1}^{N} L(y_j, f(\mathbf{x}_j)) + \lambda\|\alpha\|_1,$$
$$s.\,t.\ \ \alpha \geq 0.$$
(25)

Although we know that usually different features and kernels form different data kernels. Therefore, we recommend using different kernel types for all of these features and let our method choose which kernel type and which feature is the best classification model. Accordingly, if we define k different type of kernel, then the combination of $m \in \{1, \ldots, \kappa\}$ is a different kernel for *d* features, as:

$$k(\mathbf{x}_j, \mathbf{x}) = \sum_{i=1}^{d}\sum_{m=1}^{\kappa} \alpha_{i,m} k_m(x_j^i, x^i).$$
(26)

in this way, $\alpha \in \mathbb{R}^b$ is defined as the entire bond $\{\alpha_{i,m}\}_{i=1}^{d}\ \forall\, m \in \{1, \ldots, \kappa\}$. Notice that $b = d \times \kappa.$ Because there is a ℓ_1 regularization on this vector, we can conclude the different types of kernel and determine which kernel has the most role based on the features selected feature(s) in the data classification.

Stage one : optimaization kernel based and max-margin classification
To solve the first step optimization problem in (22), we only need to make sure that the loss function is recognizable or differentiable. Then we can easily divide the problem into two convex parts and solve it using the gradient descent method. The first part of the problem involves solving *f*, in which we assume that **α** is constant. By this method, a simplified SVM optimization with a constant kernel is specified:

$$\min_{f \in \mathcal{H}} \frac{1}{2}\|f\|_{\mathcal{H}}^2 + C\sum_{j=1}^{N} L(y_j, f(\mathbf{x}_j)).$$
(27)

Then, with holding the constant *f*, we can obtain the best **α** values by solving the second part of the problem:

$$\min_{\alpha} \frac{1}{2}\|f\|_{\mathcal{H}}^2 + C\sum_{j=1}^{N} L(y_j, f(\mathbf{x}_j)) + \lambda\|\alpha\|_1,$$
$$s.\,t.\ \ \alpha \geq 0.$$
(28)

The only problematic phrase has the form is the last word. Because we want to apply $\alpha \geq 0,$ **α** limit, we use the image operator $\mathcal{P}(\alpha) = \max(0, \alpha)$ for each repetition, which ensures that the last term, according to It can be cleaned up to **α**, so this goal can be easily optimized using descending gradient [27]. To ensure convergence of the algorithm, based on [26,28], we choose the gradient step size according to the

Armijo rule [29]. the optimization process should be intermittently divided between two parts of the problem to achieve convergence. The second part of the kernel problem It learns through weight optimization **α** in $f(\mathbf{x}_j),$ while the first part of the problem uses each SVM solver to learn the SVM function.

This algorithm contains stop criteria. It is important to note that the main objective function (25) is a convex function, and therefore has a generalized optimism. It is shown, under such a configuration with a convex goal function, a descending gradient designed to converge to a global optimum[27].

Note that our proposed framework integrates the regularization of ℓ_1 into the max-margin classification (such as SVM) to select the feature and classification. As described above, since the optimization process consists of two phases of optimization, the same idea can be used with other kernel methods, using its own kernel and integrating the ℓ_1 configuration into their formulas. Choosing the kernel function is very important, and it's the application and type of data associated with the desired application. The kernel functions are based on the scale and the degree of similarity between the samples. One of the most famous kernels is the radial base function (RBF) or the Gaucin kernel, whose formula is as follows:

$$k_{\mathrm{RBF}}(\mathbf{x}_j, \mathbf{x}) = \exp\left(\frac{-D(\mathbf{x}_j, \mathbf{x})}{2\sigma^2}\right),$$
(29)

Where $D(.,.)$ is the distance between two samples and σ's is hyper-parameter of kernel . usually, a squared Euclidean norm will be regarded as the displace size between two samples :

$$D_{\|\cdot\|_2^2}(\mathbf{x}_j, \mathbf{x}) = \left\|\mathbf{x}_j - \mathbf{x}\right\|_2^2 = \sum_{i=1}^{d}(x_j^i - x^i)^2.$$
(30)

Although this metric location must be based on the data specification. As explained earlier, in this work, we use the volume of ROIs associated with MRI images, and therefore all non-negative features. Hence, our feature vectors, if

normalized, are similar to the histogram or probability distribution functions (PDF). Many of the previous work typically normalize the data using the z-score, which modifies the values of the characteristic with zero and standard deviation of 1 to a common scale. Although this essentially changes the features and physical features of the feature, here, we normalize the feature vectors in the range (0.1) and compare them to PDFs or pseudo-histogram features that are widely used in computer applications. There are a number of other metrics that are used for irrational or quasi-histogram characteristics[30]. One of these criteria is χ^2.

$$D_{\chi^2}(\mathbf{x}_j, \mathbf{x}) = \frac{1}{2} \sum_{i=1}^{d} \left(\frac{(\mathbf{x}_j^i - \mathbf{x}^i)^2}{\mathbf{x}_j^i + \mathbf{x}^i} \right). \tag{31}$$

Another well-known place criterion in computer applications is the location of EMD (Earth Mover's Distance) [31], which evaluates the lack of similarity between two multidimensional distributions in a feature space. Directly, with two distributions given, we can call one as a crime Consider the wide sphere in space and the other as the set of cavities in the same space. EMD is the minimum amount of movement required to fill the cavities. For a one-dimensional case, like our case, EMD has a simple answer to the closed form:

$$D_{\text{EMD}}(\mathbf{x}_j, \mathbf{x}) = \sum_{i=1}^{d} \left| \text{cdf}(\mathbf{x}_j)^i - \text{cdf}(\mathbf{x})^i \right|, \tag{32}$$

Here cdf (0) is a dummy distribution function and $| 0 |$ The definite value of the function. One of the famous kernels used for the intersection kernel pseudo histogram features (the intersection kernel of the hysterogram (HIK)) that characterizes the similarities of the pseudo-histogram characteristics as defined below:

$$k_\cap(\mathbf{x}_j, \mathbf{x}) = \sum_{i=1}^{d} \min(\mathbf{x}_j^i, \mathbf{x}^i). \tag{33}$$

Since, based on the distribution of data, some data sets may use better kernels (for example, in a local feature space), we also define the linear kernel function as follows:

$$k_{\text{LIN}}(\mathbf{x}_j, \mathbf{x}) = \mathbf{x}_j^\text{T} \cdot \mathbf{x}. \tag{34}$$

To find the best kernel that can be different for each data set, use all of these definitions for the kernel and define its kernel formula by equation (26) by aggregating all types of kernels $(\forall\, m, m \in \{1, \ldots, \kappa\})$, for each feature $(\forall\, i, i \in \{1, \ldots, d\})$. Therefore, the set of all kernels can be defined as the linear kernel, the RBF kernel with different spatial metrics, and the intersection kernel as follows:

$$\{k_{\text{LIN}}, k_{\text{RBF}} + \| \cdot \|_2^2, k_{\text{RBF}} + \text{EMD}, k_{\text{RBF}} + \chi^2, k_\cap\}, \tag{35}$$

With $\kappa = 5$ different types of kernel. Our optimization framework, through the optimization of α, chooses which feature or features, and the type or types of kernels with the data set for the problem, are the best and most suitable.

Stage two: Particle swarm optimization algorithm

The weight factor u was inherently constant in most studies of the mixed kernel function, which makes the results of the research largely random. In this paper, in order to obtain a better classification effect, at this stage we select a particle swarm optimization algorithm to optimize the support vector machine parameters, which can increase the ability to learn and generalize the classification. At this stage, the model of particle swarm optimization algorithm is considered possible for each optimization problem as a particle, the position of each particle is described with a group of spatial vectors and velocity vectors. Particles find their optimal answer and optimal group response by changing their speed and location[34] and then update their speed and location through the speed and location equations. The speed update formula is:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1 rand()\left(pbest_{ij}(t) - x_{ij}(t)\right)$$
$$+ c_2 * rand()\left(gbest_j(t) - x_{ij}(t)\right) \tag{36}$$

The update formula is as follows:

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1) \tag{37}$$

The PSO's basic algorithm only needs to measure quantitative parameters, so it's easy to use. But the main problem with this algorithm is that the particle is easily trapped in the local minima, so the search precision is not high enough. Some researchers suggest advanced algorithms based on compatibility feedback, [35] since particles move toward an individual and group optimal position, so this advanced algorithm reduces their oscillation in a certain range and also increases their tendency Improved compatibility position. Although this index only takes into account the choice of population, it is difficult to come out of diversity, after the particles are trapped in the local minima, this defect is significant when solving problems with more than one local extreme.

In 1984, Shannon[36] introduced the concept of entropy and used it as a scale of information. On the one hand, for a particular system, if the randomness is too large and disorderly, the entropy of the information of this system is great. On the other hand, if a system is defined by certain rules, the entropy of the information system is very small.

For entrainment of particles, the entropy of the information reflects the distribution of particles in the response space, the larger the entropy of the information is, the particle distribution is more dispersed; the smaller the entropy of the information, the denser the particle distribution. In order to effectively prevent the loss of the algorithm to the optimal local point and to improve the convergence and accuracy of the algorithm, we combined integration of entropy information and compatibility feedback to improve the diversity of the particle population.

The entropy of particle information can effectively measure the diversity of particle populations. A particle jump with a certain probability when reducing diversity can limit the premature convergence of particles and allow particles to jump out of the optimal local point to find better results; instead, particles can be replicated from the initial stages to End, ability to discover new areas. The

computational formula for entropy particle information [37] is as follows:

$$NH_\delta^{xi}(B) = -\log_2 \frac{\delta_B(x_i)}{n} \tag{38}$$

$$NH_\delta(B) = -\frac{1}{|N|} \sum_{i=1}^{n} \log_2 \frac{|\delta_B(x_i)|}{|N|} \tag{39}$$

So that the entropy of the i-th $NH_\delta(B)$ particle information $NH_\delta^{xi}(B)$ and entropy is the entire group's information. The proposed location changes in this article are as follows:

$$NH_\delta(B) - NH_\delta^{xi}(B) > 0 \tag{40}$$

$$|F_g - F_m| < -\frac{F_m}{N} \tag{41}$$

$$rand > \frac{1}{1 + \log_2(j)} \tag{42}$$

As j is the number of the current stage, i is the particle sign, N is the number of particles. Formula (40) means that the entropy of i-th particle information is less than entropy of the total information of particle swarm, so that the optimization process can be turned into a bad state at this moment, and the population diversity gradually decreases. In the formula (41), F_g is the optimum overall compatibility, F_m is the average adaptation of all particles in the current repeats; if the definite value of the difference between these two is less than - (F_m) / N, the tendency to move to the maximum optimal position should be at this point Increase. Formula (42) is the probability of a particle mutation that gradually decreases with increasing number of repetitions. Formula (40) is an estimation of particle information entropy, Formula (41) is a Particle Adjustment Evaluation, Formula (42) is a mutation probability, particles that have these three conditions are mutated. The mutation method, the random throw of 20 new particles into the overall response space, is the formulation of the particles as follows:

$$pop(j, :) = (pop_{max} - pop_{min}) * rand + pop_{min} \tag{43}$$

As pop (j) shows the location of the j-th particle, pop_{max} and pop_{min}, it shows the boundaries of the particle. Calculate compatibility based on the location of new particles, if this value is less compatible than the previous one, update the position of this particle. This process not only increases the range that the particle searches for the overall space of the answers, but also increases the diversity of the population of the particle, which can effectively prevent local convergence. This advanced algorithm is particularly suitable for optimization problems with a higher dimension and more than one local extermum.

The old PSO-RBF method of the SVM kernel needs to optimize two parameters (width g of the RBF kernel function, the penalty parameter c. The mixed PSO of the SVM kernel needs to optimize three parameters: width g of the RBF kernel function, penalty parameter c, and mixed weight coefficient u. The basic process of classification of the SVM mixed kernel function is based on the entropy of the particle swarm optimization algorithm as follows.

Step One: Generate random velocity and initial particle position

Step Two: initialization of the PSO parameter

Step Three: Updated location and particle velocity according to formulas[36,37].

Step Four: Calculating the amount of compatibility of each particle, P_{best} is a particle of a set of current optimal positions; G_{best} is a collection of general optimal positions in the primary population.

Step Five: If a particle is allowed to mutate based on formulas[38,39,40,41,42] mutate it; if the compatibility is better than P_{best}, update the P_{best} with the new magnitude position.

Step Six: If the end criteria are available, go back to step 7 and otherwise go back to step 3.

Step Seven: In order to create the values the parameters c, g, u and the velocity of classification accuracy of the test set; the flowchart of the algorithm is shown in Fig. 3.

Fig. 3: Advanced PSO-Mixed Flowchart SVM Kernel
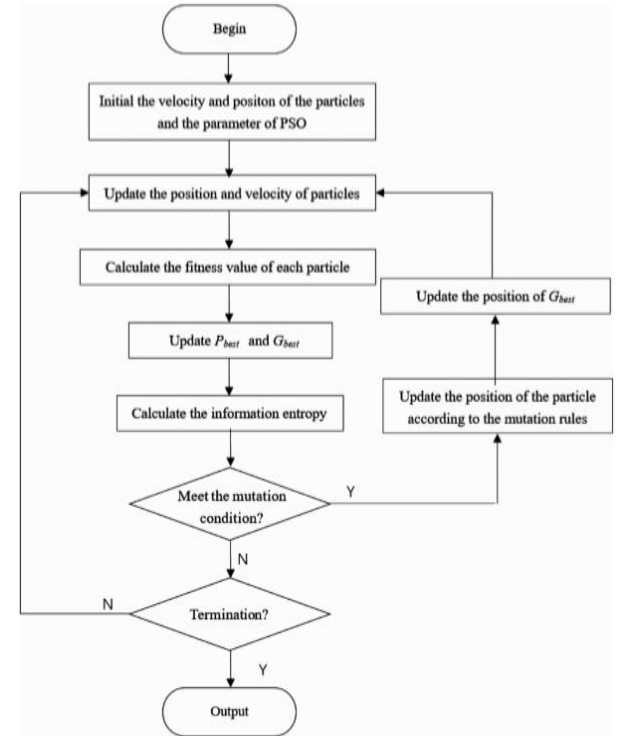
Leave-one-out cross-validation (LOOCV) method: The purpose of using SVM's optimized model design is to determine which of the subsets of the 18 texture features should be used to generate the best prediction model. In most modeling methods, if we compare the features of the subsets using the error rate in the sample, the best performance will occur when all 18 features are used. With this in mutual validation, the best-fit model typically includes a subset of features that are really information.

The method of assigning a sample (LOO)[38], as its name implies, is that it considers a single set of natural data sets that contain sample M and validation data. And it

should be said that the M-1 samples will remain as training data for building a SVM classification model. After confirming this model, the sample is passed to the training dataset and another sample is selected. The approach is repeated as each observation in the sample is used once and as validation data. After a loop a set of M training times and test samples of M, all items in a dataset are approved using the SVM model. In this study, the number of M samples that are selected according to the set, MR 3T and 1.5T are respectively 50 and 18, respectively.

Mutual validation based on the assignment of an item is usually used in the analysis of a very small data set. Although the LOOCV methodology is often expensive because such training should be repeated over and over again, it should be noted that this method is still a robust and accurate method for assessing the predictive model. In some studies, several modified methods, such as minimum squares, nuclear regression, and cross-validation can be significantly improved by pre-calculating certain values that are frequently required in training or through the use of rules. However, careful attention is required in order to maintain the "total amazement" of the validation set of the teaching method; otherwise, opposition may arise. In order to develop the modeling approaches in a precise manner, in our experiment, we choose the traditional LOOCV method.

Calculation of texture features with different conditions Texture: Text features display different values if using different preprocessing methods, different image quality and ROI sizes. Therefore, we need to optimize each of these conditions before parsing and optimizing the subset of the property.

First, the specified accuracy w/o for a soubel filter is tested in the dataset which is arranged in order to determine the best SVM input pattern, and the half of the data set is eliminated by the result of this test. [39-40]. Secondly, an optimal size of the three different sizes of ROI has been selected[41].

Since the number of selected ROIs is too high, the texture features of 18 × 10 in the order arranged are reduced by averaging the value of the features from the 10 selected ROIs and the set of new data will be converted to only 18 textures. After all the optimal processing steps are completed, the total ROI in one case is reduced from 1600 to 60. This step can significantly reduce the number of input samples to ensure the efficient training of SVM.

3.7 Optimizing the features Besides the above optimizing problem

In addition to the high optimization problem, there is a challenge to optimize 18 texture features, for example, which number and which of the best features are to show fibrosis? In this paper, data samples in the SVM classification model are divided into two categories: mild liver fibroses S0, S1, S2 classified as negative samples, and cases of severe liver fibrosis and cases of cirhosis , S3, S4, S5, which are classified as positive samples. Of the data sets arranged in order, each possible combination of 18 features is performed using the SVM model in the LOOCV method for accuracy comparison. This method is as follows:

Step 1: Non-cirrhosis and cirrhosis groups are seen as Groups 0 and Group 1 containing M samples in each group, respectively. 18 features are derived from GLCM and averaged over 10 selected ROIs. (2M × 18 features are introduced as input candidates).

Step 2: The number of input items (n) is selected from a combination of 18 features. (250-1 obtained a different composition, in which $n \in ^{1,15}$.

Step 3: Depending on the selected items n Under the set of step 2, the input characteristics of 22M × n are eliminated from group 0 and group 1, respectively. A single instance with n feature is selected as validation data, and the remaining 2M-1 samples are used as training data.

Step 4: The SVM classifier is modeled using 2M-1 training samples and confirms the remaining sample of the k test. If k belongs to group 0 and is correctly classified, then TN = TN + 1; so if group 1 is sorted correctly then TP = TP + 1.

• Step 5: If unknown samples exist in 2M, then step 3 is executed; otherwise, the accuracy rate is calculated (AR): AR0 = TN / M, AR1 = TP / M, AR = (AR0 + AR1) / 2.

Step 6: If the ergodic process does not end, Step 2 runs, otherwise, Step 7 will be executed.

Step 7: End. $2^{15}-1 = 38,865$ AR is produced for further processing.

Evaluation of the performance of texture features: Each of the data sets arranged respectively have a value of 38,865 of AR values, which can be divided into 18 groups according to the number of items selected in a subset [AR1, AR2,, AR18]. Also, each AR^i has a number of sub-sets of a combination of items in the 18 features, in which the maximum AR^i value is defined as Max AR^i. Since the MR data set contains 4 fuzzy images, the maximum value of AR^i in a dataset arranged as j is defined as Max AR^{ji}. Additionally, the sum of the mean values of the maximum values in the complete data set obtained by the K method can be calculated as follows:

$$AR_k = \left[\frac{1}{m}\sum_{j=1}^{m} MaxAR_j^1, \frac{1}{m}\sum_{j=1}^{m} MaxAR_j^2, \ldots, \frac{1}{m}\sum_{j=1}^{m} MaxAR_j^i, \ldots, \frac{1}{m}\sum_{j=1}^{m} MaxAR_j^n \right] \quad (44)$$

Where, K=1 : MRI 3T : MRI 1.5T .

RESULTS AND DISCUSSION

Accuracy of staging fibrosis: The main polynomial function of the SVM is considered as the dominant position in the classification, which can indicate that the data set is relatively scattered after being mapped to a high-dimensional space. The distribution of data points in a dataset is unknown after mapping to a high-dimensional space, which indicates that the previous method of assigning the value of u is not empirically reliable. The efficiency of this classification can be evaluated with accuracy in the test set, classification accuracy rate and computational time line graphs in the PSO-RBF core optimized, the PSO hybrid corea, and PSO hybrid core support vector machine, as it has been shown flowchart Figure 3 it has been shown.

As shown in Fig. 4, by comparing the PSO hybrid core kernel with the PSO-RBF core, several classification accuracy rates are improved, suggesting that the linear combination of the polynomial cores and the RBF core is

better than the single RBF core in generalization capability. By comparing the improved PSO hybrid core with the PSO core, the rate of classification accuracy of the data set is changing, but the rate of classification accuracy of the collection has improved to a certain extent due to the increase in the collection.

The features and examples of these three datasets are relatively large, the distribution of data points in a high-dimensional space can be complicated after using the combination core function. And the improved PSO accuracy in the optimization process is better under these circumstances, so the accuracy of the final classification is improved. Figure 5 shows that the PSO kernel computing time and the PSO core are longer than the PSO-RBF core due to the high optimization of the parameters and computation of the often internal product. This paper focuses primarily on the impact test of the classification algorithm, and we do not process the feature selection. Therefore, we must believe that the calculation time will be greatly reduced after a series of feature selection steps that improve the accuracy of classification.

The results showed that the accuracy of the classification of the algorithm in this paper is greater than the traditional PSO-SVM algorithm. This is due to the fact that the SVM hybrid core has both global and local characteristics, which can efficiently extract the dataset character mapping and adapt to complex data. In addition, improved PSO based on entropy of information can optimize the SVM parameters, especially the weight factor, which can greatly play the role of the combined core functions. The result of optimizing the various data sets is different weights, which indicates that the empirical allocation method is not reliable for the combined core weight. From a public data collection website, we found a classification application of the three datasets used in this article. However, in contrast to the above method, the accuracy of our method in these multiple data sets is higher. In general, the training time of the classification model is due to the optimization of the additional parameters and the internal product calculations longer than the normal core SVM, but it is acceptable when the time of calculation of the test data after the modeling is made. This paper focuses on verifying the effectiveness of our classification algorithm, and we do not process feature selection. Therefore, we believed that after a series of selection steps, the computing time would be significantly reduced, which also improves the classification accuracy. We also analyze the composition of several common core functions, the next research can work on the use of other linear combinations of functions or even the investigation of complex nonlinear complexes, we can also process different methods of selecting features before the classification, or Improve the PSO algorithm to increase parameter accuracy and improve classification efficiency.

The accuracy of different imaging conditions and different combinations of the number of texture features is different. In the studies we have been done earlier [18], we focused on the performance of MRI-based fibrosis in the Sobel filter. The results obtained from these studies show that the best accuracy of the variable is 0.76 without the pre-processing of the Sobel with 0.64 with the filter. Compared with the highest scores obtained in direct use of row data, we can say that the Sobel filter does not have significant performance in the pre-processing of MR images. The reason for this may be the fact that the Sobel filter not only strengthens the fibrosis pattern, but also improves background sound. We also concluded that the different sizes of ROI affect the classification, and the size 20×20 is the optimal size between the three sizes of ROI [26]. Therefore, this test was performed using raw ROI 20 × 20 raw data

Choosing the optimal number of features: In order to determine the subset of the features that are really information, LOOCV runs on each of the arranged image data sets that are arranged in a regular fashion. In this modeling technique, each combination of 18 features is designed to provide the best AR from 1 to 15 feature subtypes in Table 1. As shown in Fig. 3, the optimal number of texture features for classification the degree of fibrosis is 4 to 7, regardless of which method is used. The best overall average of the maximum AR value is 82.42% and 87.60%, respectively, obtained by 6 features of MRI 3T images in hospital , and also in the case of 5 features of MR 1.5T Available at Dena Hospital, this amount will be equal to 85.12%. The overall performance is computed by the sum of the mean of a maximum of 18 features. Also, the results show that MR 3T has a better quality in the fibrosis classification than 1.5T, in addition, the lowest total sum of the mean maximum AR value shown in Table 1 occurs when all of the 15 The feature is used together to build the SVM prediction model, which indicates the importance of selecting the features of the notification.

Distribute features containing useful information: Based on the sum total results of the mean of the maximum AR values, we conclude that the number of subsets of the features, which range from 5 to 7, is better. Another interesting point to look at is what these 5 to 7 features are in reality. A threshold value of AR is set to 0.7 to select the subset features that have the best performance. If the AR value is greater than 0.7, the number of each of these items in this AR will be plus 1. According to these statistics, a discriminating force histogram is produced for each feature present in the classification. These histograms are produced in Tables 2a and 2b and shown in Figure 7a and 7b after normalizing the scale in a range of [0-1].

Performance of different fuzzy images: As shown in Fig. 8, AR values are proportional to the injection of contrast units, and both CT and MR images reach the highest level of performance in the equilibrium phase. The MR can effectively display fibrotic texture and the balance-of-phase image is presented as a major tool for the interpretation of cirhosis.

The complete processing of credit with the SVM method in LOO is very time consuming, because the number of combinations that are derived from 18 texture features is very high. In our experiment, a 40-member group (40 + 1), which consists of a data set of group 1, uses a rigorous learning and evaluation rounds in an array of data, for example, doing it on average from 6 to 9 The hours takes place at the workstation (4.3G processor with 4G RAM, Windows 7, 64-bit version). However, when the modeling is completed, we can use this prediction model to detect a new dataset within a second. Although some

studies have introduced their effective methods for educational prototypes with improved algorithms and have stated that the methods on the part of them can greatly reduce the processing time of validation, but it should be noted that braininess and precision is a matter that can still be considered bytes. In addition, it is difficult to determine a golden standard for examining the results of the studies. Therefore, in this study, we have chosen the most reliable method of validation, which is the LOOCV method, in order to guarantee the objectivity of our results. That includes a

similar approach to human observation in experiments for conducting investigations, which would be a big challenge if it is integrated with CAD. One of the studies we have already done [42] shows some preliminary results in comparing human observation with computerized diagnostics. Although, only with regard to the fibrosis pattern, the function of radiologists is significantly lower than computer algorithms, but the trend of accuracy change on both sides is the same.

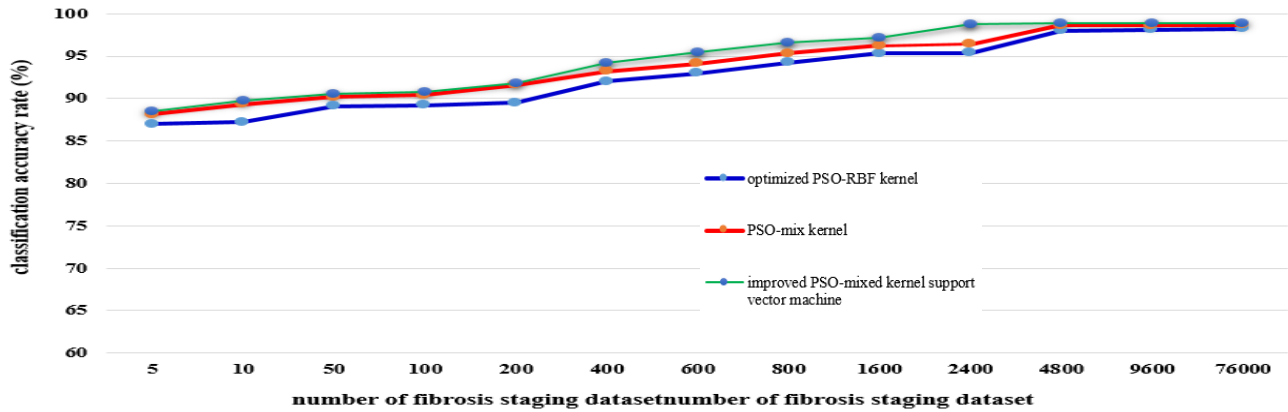Fig. 4. The classification exact rate in three algorithms.

Fig 5: computational sacrifice of optimized PSO-RBF kernel, PSO-mix kernel, and improved PSO-mixed kernel support vector machine
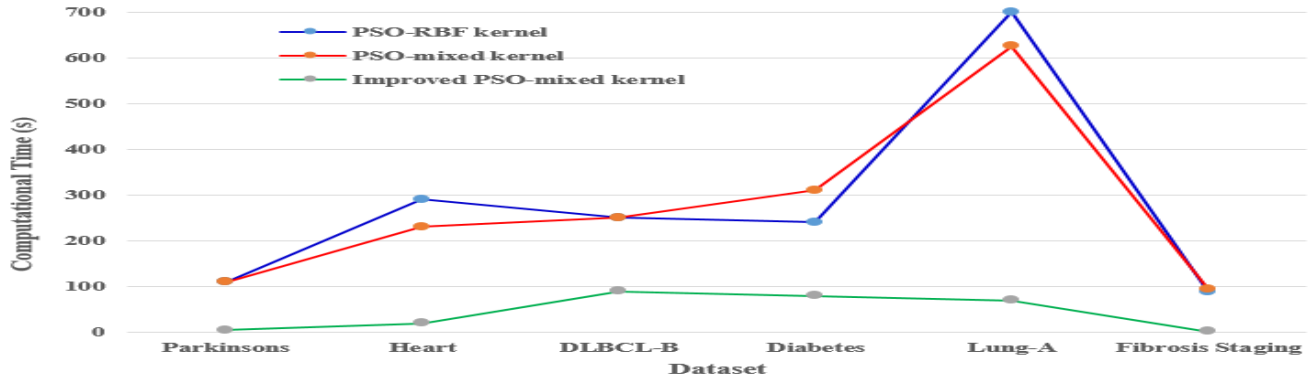

Fig. 6: the accuracy curve, by which selection number of texture features , shows three different functions in 2 modes.
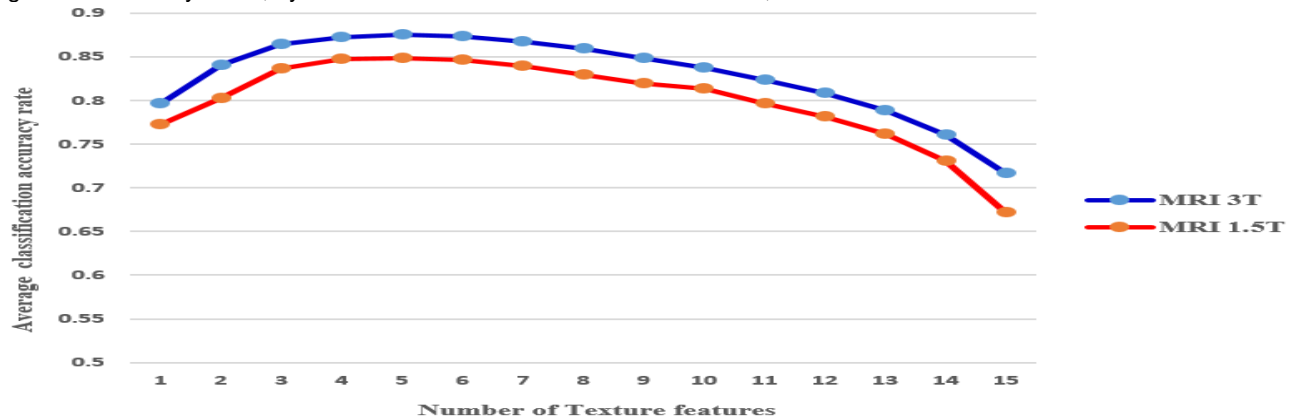
Fig. 7a. by histograms there is a distinguishing force for all 18 exploited features of the 3 various data series, capability of every characteristic in discriminating fibrosis texture by means of counting increased numbers in AR values , is investigated.
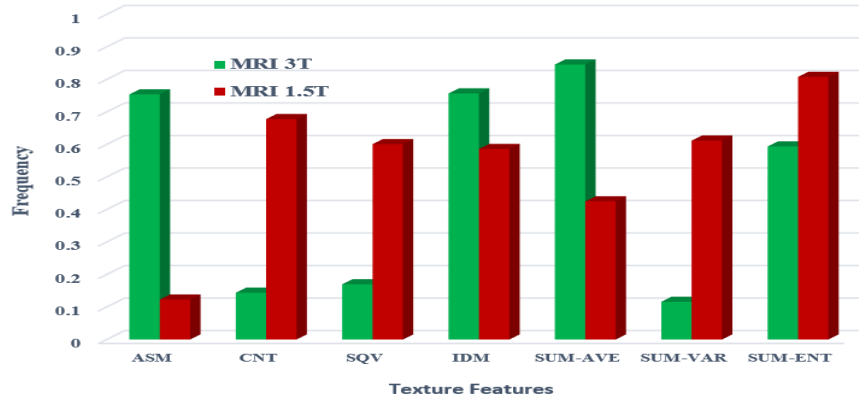

Fig. 7b. by histograms there is a distinguishing force for all 18 exploited features of the 3 various data series, capability of every characteristic in discriminating fibrosis texture by means of counting increased numbers in AR values , is investigated.
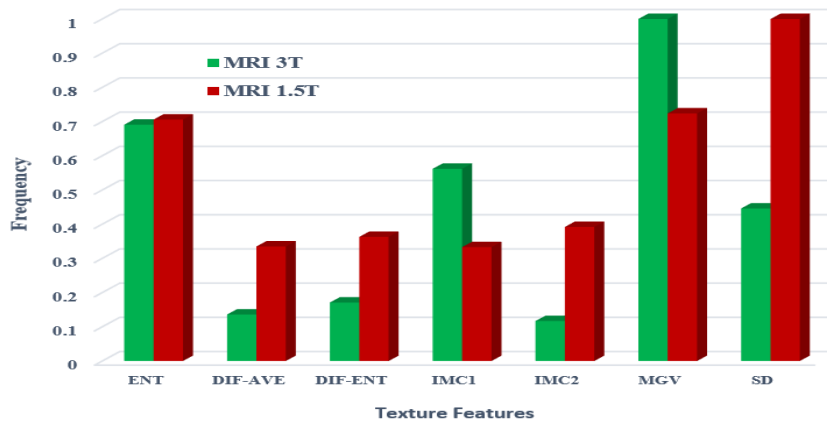

Fig 8: the AR value in 4 depicted phases . the texture analysis in equilibrium phase with respect to depicted phases is more efficient.
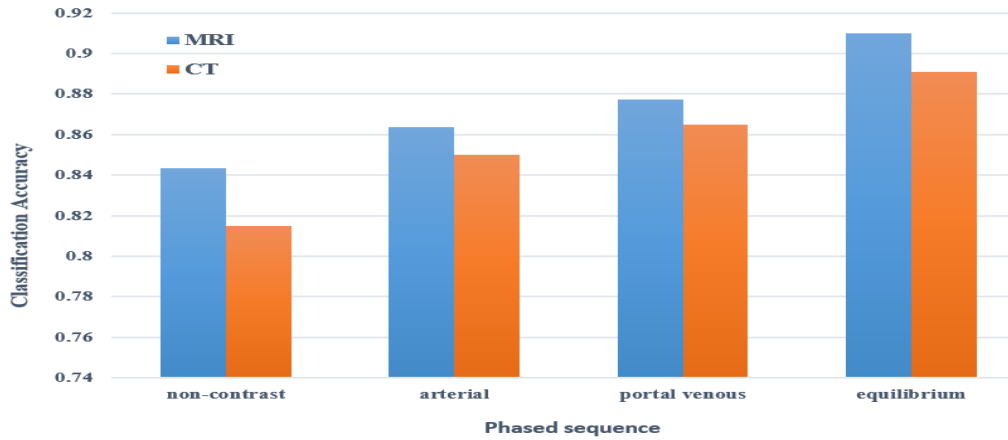

Table 1: The AR values for MRI to use in different number of features as the SVM input vectors.

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
MRI3T	0.797	0.841	0.865	0.873	0.876	0.874	0.868	0.860	0.849	0.838	0.824	0.809	0.789	0.761	**0.717**
MRI1.5T	0.773	0.803	0.837	0.848	0.849	0.847	0.840	0.830	0.820	0.814	0.797	0.782	0.762	0.731	**0.672**

Table 2a . the importance magnitude value in MRI provisions.

Number	ASM	CNT	COR	SQV	IDM	SUM_{AVE}	SUM_{VAR}	SUM_{ENT}
MRI3T	0.754	0.144	0.000	0.170	0.757	0.846	0.116	**0.594**
MRI1.5T	0.123	0.678	0.000	0.601	0.586	0.425	0.612	**0.0.808**

Table 4b. the importance magnitude value in MRI provisions.

Feature	ENT	DIF$_{AVE}$	DIF$_{ENT}$	IMC1	IMC2	MGV	SD
MRI3T	0.691	0.136	0.171	0.562	0.117	1.110	**0.446**
MRI1.5T	0.706	0.335	0.363	0.333	0.392	0.724	**1.110**

CONCLUSION

In chronic liver disease, liver fibrosis is considered as one of the most important prognostic / diagnostic prognosis. Advanced stages of liver fibrosis are diagnosed with clinical signs of liver dysfunction and liver disease. In summary we can say that we have evaluated the accuracy of the determination of the fibrosis stage using MR images optimized in the equilibrium phase in the liver fibrosis model. The evaluation of MR images in relation to CAD system that differs between non-linear normalized texture features in different hepatocytes with different fibrotic texture sizes. We were able to effectively distinguish the differences in the liver textures that were present in the stages of liver fibrosis. We found that the nonlinear normalization of images, the key to achieving, is better to detect. The proposed CAD analysis of routine MR images shows the hope in improving accuracy in determining the phase of fibrosis with a non-invasive approach. In this study, a new nonlinear combination of the classification process using SVM has been investigated. An algorithm for the classification of the SVM core function is optimized, which consists of two steps. On the one hand, the combined core function used in this paper is the nonlinear combination of the polynomial function and the RBF core in the backup vector machine, and its ability to learn and generalize it from the RBF core is more advanced. On the other hand, the PSO algorithm based on the entropy of information for The search for relative kernel parameters is used, information entropy feedback can effectively improve the diversity of particles and improve the accuracy of parameters optimization, thus increasing the accuracy of classification. Compared to the PSO-RBF core and The PSO core of the compound, the SVM of the PSO core, can classify precision in the evaluation of the staging of Kirobriz DI, and this also suggests that this algorithm has a good practical value for evaluating the hepatic fibrosis staging.

Limitations and future offers: There are limits to our study, and we need to improve them in the future. First, the data sets are divided into two groups. This is due to the unbalanced distribution of each phase of fibrosis, as well as the goal of reducing the complexity of validation. In the next step, our studies will examine the quantitative fibrosis in 5 steps with incremental datasets. Second, this study does not combine images from multiple phases due to computational complexity. The prominent features of each phase can be combined to achieve 80 features (18 × 4) per item, and the execution of such an action in the future will show us the importance of a variable of all phases. Thirdly, this study was conducted as a preliminary study, with a great deal of time and effort to manually access ROIs and record validation values. We are now beginning to develop an automated algorithm for choosing ROI in the liver region, and we have integrated the GLCM, LOOCV, and SVM in one program. Finally, all ARs that are classified in the texture analysis are not so prominent (less than 0.8) that can be used as a separate model for determining the stage of fibrosis in clinical methods. This may be due to the fact that the fibrosis pattern is clearly not different at different stages, and even experienced radiologists cannot diagnose fibrosis just because of the texture in the pictures. Different stages of fibrosis can be distinguished using texture features, along with other effective features, such as shape, volume, stretching, etc. In addition, it is recommended to extract morphological features combined with statistical characteristics in order to achieve superior results. It is also suggested in future studies that we will examine the selection of morphological features along with statistical characteristics in the MRI and SPECT / CT hybrid imaging for better results.

REFERENCE

1. El-Serag HB, Mason AC. Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med1999;340(March):745–50.
2. Yatsuhashi H, Yano M. Natural history of chronic hepatitis C. J Gastroenterol Hepatol 2000;15(May (Suppl.s2)):111–6.
3. Wynn TA. Cellular and molecular mechanisms of fibrosis. J Pathol 2008;214(Jan(2)):199–210.
4. 2001 Viral Hepatitis Prevention and Treatment Plan: Chinese Society of Hepatology
5. Fallow J and Hayes P 2011 Pathogenesis and treatment of hepatic fibrosis: is cirrhosis reversible? Clinical Medicine 11179–83.
6. West J and Card T R 2010 Reduced mortality rates following elective percutaneous liver biopsies Gastroenterology 139 1230–7.
7. Bonekamp D, Bonekamp S, OuHY, TorbensonMS, Corona-Villalobos CP, Mezey E and Kamel I R 2013 Assessing liver fibrosis: comparison of arterial enhancement fraction and diffusion-weighted imaging J. of Magnetic Resonance Imaging 00 1–10.
8. Bohte AE et al 2014 Non-invasive evaluation of liver fibrosis: a comparison of ultrasound-based transient elastography and MRelastography in patients with viral hepatitis B andC European Radiology 24 638–48.
9. Kim BH, Lee JM, Lee Y J, LeeKB, SuhKS, Han JKand Choi B I 2011MRelastography for noninvasive assessment of hepatic fibrosis: experience from a tertiary center in Asia J. of Magnetic Resonance Imaging 34 1110–6.
10. Venkatesh SK, Wang G, Lim SGand Wee A 2014 Magnetic resonance elastography for the detection and staging of liver fibrosis in chronic hepatitis B European Radiology 24 70–8.
11. Venkatesh SK, YinMand Ehman R L 2013 Magnetic resonance elastography of liver: technique, analysis, and clinical applications J. of Magnetic Resonance Imaging 37 544–55.
12. Feier D, Balassy C, Bastati N, Stift J, Badea R and Ba-Ssalamah A 2013 Liver fibrosis: histopathologic and biochemical influences on diagnostic efficacy of hepatobiliary contrast-enhancedMRimaging in staging Radiology 269 460–8.
13. Bülow R, Mensel B, Meffert P, Hernando D, EvertMand Kühn J P 2013 Diffusion-weighted magnetic resonance imaging for staging liver fibrosis is less reliable in the presence of fat and iron European Radiology 23 1281–7.
14. Kato H, Kanematsu M, Zhang XJ, Saio M, Kondo H, Goshima S and FujitaH2007 Computer-aided diagnosis of hepatic fibrosis: preliminary evaluation of MRI texture analysis using

the finite difference method and an artificial neural network American Journal of Roentgenology 189 117–22.

15. Goshima S, Kanematsu M, Kobayashi T, Furukawa T,Zhang XJ, Fujita H, Watanabe H, Kondo H, MoriyamaNand BaeKT 2012 Staging hepatic fibrosis: Computer-aided analysis of hepatic contours on gadolinium ethoxybenzyl diethylenetriaminepentacetic acid-enhanced hepatocytephase magnetic resonance imaging Hepatology 55 328–9.

16. HouseMJ, Bangma S J, Thomas M, GanME, AyonrindeOT, Adams L A, Olynyk JKand St Pierre TG2013 Texture-based classification of liver fibrosis using MRI J. of Magnetic Resonance Imaging 41(2) 322–8.

17. Sela Y, Freiman M, Dery E, Edrei Y, Safadi R, Pappo O,Joskowicz L and Abramovitch R 2011 fMRI-based hierarchical SVMmodel for the classification and grading of liver fibrosis IEEE Trans. Biomed. Eng. 58 2574–81.

18. Semelka RC, Chung J J, Hussain SM, MarcosHB and Woosley J T 2001 Chronic hepatitis: correlation of early patchy and late linear enhancement patterns on gadoliniumenhancedMRimages with histopathology initial experience J. of Magnetic Resonance Imaging 13 385–91.

19. Zhao Y P, GuoDM, Liu L, LiuWH,MuWY and Zhang P2015 Apparent diffusion coefficient measurements and Gd-DTPA enhanced-imaging in staging hepatic fibrosis in rats International Journal of Clinical and Experimental Medicine 8 2197–204.

20. Haralick RM, Shanmugam Kand Dinstein I 1973 Textural features for image classification IEEE Transaction on Systems,Man, and Cybernetics 3 610–21.

21. Virmani J, Kumar V, KalraNand KhandelwalN2011 Prediction of Cirrhosis based on singular value decomposition of gray level co-occurence marix and a neural network classifier IEEE Developments in E-systems Engineering 56 146–51.

22. Azmi R and NoroziN2011 A new Markov random field segmentation method for breast lesion segmentation inMR images J. of Medical Signal and Sensors 1 156–64.

23. Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S. & Ma, Y. Robust face recognition via sparse representation. IEEE TPAMI 31, 210–227(2009).

24. Hofmann, T., Schölkopf, B. & Smola, A. J. Kernel methods in machine learning. Ann. Stat. 1171–1220 (2008).

25. Varma, M. & Babu, B. R. More generality in efficient multiple kernel learning. In ICML, 1065–1072 (2009).

26. Saghafi, B., Rajan, D. & Li, W. Efficient 2d viewpoint combination for human action recognition. Pattern Analysis and Applications 19, 563–577 (2016).

27. Calamai, P. H. & Moré, J. J. Projected gradient methods for linearly constrained problems. Mathematical programming 39, 93–116 (1987).

28. Chapelle, O., Vapnik, V., Bousquet, O. & Mukherjee, S. Choosing multiple parameters for support vector machines. Machine learning 46, 131–159 (2002).

29. Liu, P.-L. & Der Kiureghian, A. Optimization algorithms for structural reliability. Structural safety 9, 161–177 (1991).

30. Cabral, R. S., De la Torre, F., Costeira, J. P. & Bernardino, A. Matrix completion for weakly-supervised multi-label image classification. IEEE TPAMI (2015).

31. Ling, H. & Okada, K. An efficient earth mover's distance algorithm for robust histogram comparison. IEEE TPAMI 29, 840–853 (2007).

32. Kumar A, Ghosh SK, Dadhwal VK. Study of mixed kernel effect on classification accuracy using density estimation. Proceedings of the ISPRS Commission VII Symposium 36 (Part 7); 2006.

33. Hovestadt T, Binzenh€ofer B, Nowicki P, et al. Do all inter-patch movements represent dispersal? A mixed kernel study of butterfly mobility in fragmented landscapes.J Anim Ecol. 2011;80:1070–1077.

34. Kennedy J. Particle swarm optimization. In:Encyclopedia of machine learning. US: Springer. 2010.p. 760–766.

35. Tang Y, Wang Z, Fang J. Feedback learning particle swarm optimization. Appl Soft Comput. 2011;11:4713–4725.

36. Shannon CE. A mathematical theory of communication. ACM SIGMOBILE Mobile Comput Commun Rev.2001;5:3–55.

37. Hu Q, Pan W, An S, et al. An efficient gene selection technique for cancer recognition based on neighborhood mutual information. Int J Mach Learn Cybern.2010;1:63–74.

38. Schapire R. Leave one out error stability, and generalization of voting combi-nations of classifiers. Mach Learn 2004;55:71–97.

39. Li W, Zhang X, Kanematsu M, Hara T, Zhou X, Fujita H, et al. Development ofan automated method for differentiation of cirrhotic liver in abdominal MRimages. Med Img Inf Sci (MII) 2004;21(2):137–43.

40. Wu C, Chen Y. Texture features for classification of ultrasonic liver images. IEEETrans Med Imag 1992;11(2):141–52.

41. Kato Hiroki, Zhang Xuejun, Kondo Hiroshi, et al. Computer-aided diagnosis ofchronic liver disease using 3 Tesla MRI artificial intelligence: optimization ofthe region of interest in texture analysis. In: Proceeding of 38th Japan magneticresonance medicine conference. 2010 (in Japanese).

42. Kato Hiroki, Kanematsu Masayuki, Zhang Xuejun, et al. Computer-aided diag-nosis of hepatic fibrosis: preliminary evaluation of MRI texture analysis usingthe finite difference method and an artificial neural network. Am Roentgen RaySoc (AJR) 2007;189:117–22.