

Research Article

Computational Classification and Diagnosis of Alcoholic Liver Diseases Using General Regression Neural Network

Naiping Li, Yongfang Jiang, Jin Ma, Bo He, Wei Tang, Mei Li, Qing Huang, and Ting Yuan

Liver Diseases Research Center, The Second Xiangya Hospital, Central South University, Changsha, Hunan 410011, China

Correspondence should be addressed to Yongfang Jiang; jiangyongfang@hotmail.com

Received 2 January 2014; Revised 20 March 2014; Accepted 21 March 2014; Published 22 May 2014

Academic Editor: Gongnan Xie

Copyright © 2014 Naiping Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alcoholic liver diseases cause high incidence of death worldwide. However, computational diagnosis and classification of alcoholic hepatitis have not yet been established. In this study, we used general regression neural network (GRNN) model with a high-performance classification ability to diagnose and classify alcohol hepatitis. We used tenfold cross-validation to demonstrate the error rate of networks. The results show an accuracy of 80.91% of the back diagnosis in 110 patients and the accuracy of 81.82% of predicting-diagnosis in 11 patients referring to the clinical diagnosis made by a group of experts. This study suggested that using the liver function tests as the input layer variables of GRNN model could accurately diagnose and classify alcoholic liver diseases.

1. Introduction

There are many types of hepatitis diseases including diseases caused by viruses, toxic chemicals, alcohol consumption, and so forth. Long-term excessive drinking can cause extensive necrosis within the liver or lead to even liver failure. Fatty liver is an early sign of alcoholic liver damage and then it develops into alcoholic hepatitis, alcoholic liver fibrosis, and finally alcoholic liver cirrhosis. Statistics have consistently shown that cirrhosis and its sequelae are responsible for about 2% of all causes of death in the USA, among them 80% to 90% of cirrhosis was alcohol-related [1]. In China, over 20% of alcoholics have alcoholic liver disease (ALD) [2]. An appropriate treatment can save a significant number of lives if patients were promptly diagnosed in the early stages. Unfortunately, accurate diagnosis of ALD in the early stages has never been an easy task. Importantly, different pathological stages of ALD require different therapeutic strategies [3]. Therefore, accurate diagnosis and classification of ALD are crucial for clinical management. However, many factors make diagnosis and classification of ALD difficult, often delaying the correct diagnosis. For example, ALD is currently classified by clinic symptoms, biochemical manifestations, and pathologic manifestations [2, 3]. However, clinical symptoms and biochemical alterations are not disease specific.

Pathological tests are usually not acceptable for patients. Therefore, it is important to develop a reliable and powerful medical diagnostic system to classify disease stages, reduce diagnosis time, and improve diagnosis accuracy. However, medical diagnosis by nature is a complex cognitive process. Currently, application of computational methods, such as neural networks, has shown great potential in the diagnosis of liver diseases. However, application of neural networks to diagnose and classify ALD has not yet been established.

Different neural networks have been applied to liver disease diagnosis except for ALD. For instance, Nakano et al. investigated the ability of neural networks to discriminate between mild and severe chronic active hepatitis and demonstrated that neural networks are capable of predicting the diagnosis for 78% of the cross-validation group [4]. Kim et al. adopted a prediction model including logistic regression, a decision tree, and a neural network to analyze the risk factors of liver disease and obtained 72.09% of accuracy [5]. Bascil and Temurtas adopted Levenberg-Marquardt algorithm to study the hepatitis disease diagnosis and obtained a classification accuracy of 91.87% [6]. Lin used an intelligent model of integrated classification and regression (CART) and case-based reasoning (CBR) techniques for the diagnosis of liver diseases and the CBR diagnostic accuracy rate is 90.00% [7]. Lin and Chuang established an intelligent liver diagnosis

model which integrates artificial neural networks, analytic hierarchy process, and case-based reasoning methods to examine if patients suffer from liver disease and to determine the types of the liver disease [8]. One of the numerous successful examples goes to Chou et al. [9], which adopted a prediction model including multivariate adaptive regression splines (MARS) and ANN (artificial neural network) to handle breast cancer diagnostic problems. The accuracy of the model was as high as 98%. However, we have not found any study that considers a system to diagnose and classify the alcoholic liver diseases.

GRNN is one of the most popular neural networks and was adopted because of its strengths [6, 10, 11]: (1) parallel structure and better generalization make GRNN super than almost any other neural network; (2) GRNN is a noniterative neural network structure and its data flows from input layer to the output layer. Therefore, there is no iterative learning present such as in the case of multilayer perceptrons (MLP) making them faster to some extent; (3) GRNN works quite accurately with light datasets and performs well on noisy data compared to backpropagation neural networks (BPNN); (4) GRNN is also very unswerving and as the size of the dataset increases the error approaches towards zero. These are the reasons why GRNN is being used in medical classification, predictive and diagnostics problems. In this study, we used general regression neural network (GRNN) pattern recognition to diagnose and classify ALD.

2. Material Studied, Methods, and Techniques

2.1. GRNN Network Structure

2.1.1. GRNN Theory. Generalized regression neural network theory is based on nonlinear regression analysis, in which the greatest probability value (y) is actually calculated through regression analysis of dependent variable Y on independent variable x . Assume that $f(x, y)$ represents the known joint continuous probability density function of a vector random variable, x , and a scalar random variable, y . The regression of y on x is given as follows:

$$\hat{Y} = E\left(\frac{y}{x}\right) = \frac{\int_{-\infty}^{\infty} y f(\mathbf{X}, y) dy}{\int_{-\infty}^{\infty} f(\mathbf{X}, y) dy}. \quad (1)$$

\hat{Y} is the predicted value under input X [11]. The density $f(X, Y)$ can be estimated from a training set by using the consistent estimator proposed by Parzen [12]:

$$\begin{aligned} \hat{f}(X, y) &= \frac{1}{n(2\pi)^{(p+1)/2} \sigma^{p+1}} \\ &\times \sum_{i=1}^n \exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \\ &\times \exp\left(-\frac{(X - Y_i)^2}{2\sigma^2}\right). \end{aligned} \quad (2)$$

Here, the probability estimator $\hat{f}(X, y)$ is based upon sample values X_i and Y_i of random variables x and y , where n

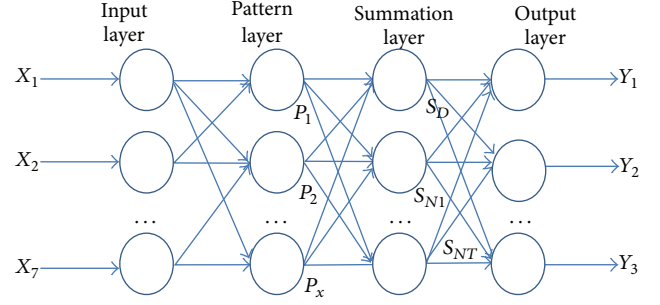


FIGURE 1: Generalized regression neural network model diagram for diagnosis of alcoholic liver disease.

is the number of sample observations, p is the dimension of the vector variable x , and σ is the Gaussian width factor (here called smoothing factor) as shown in (2). In particular, using $\hat{Y}(X)$ instead of $f(x, y)$ into (1) and interchanging the order of integration and summation yield the desired conditional mean (3), designed:

$$\begin{aligned} \hat{Y}(X) &= \left(\sum_{i=1}^n \exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \right. \\ &\quad \times \left. \int_{-\infty}^{\infty} y \exp\left(-\frac{(Y - Y_i)^2}{2\sigma^2}\right) dy \right) \\ &\quad \times \left(\sum_{i=1}^n \exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \right. \\ &\quad \times \left. \int_{-\infty}^{\infty} \exp\left(-\frac{(Y - Y_i)^2}{2\sigma^2}\right) dy \right)^{-1}. \end{aligned} \quad (3)$$

Due to $\int_{-\infty}^{\infty} z e^{-z^2} dz = 0$, the available network output $\hat{Y}(X)$ is

$$\hat{Y}(X) = \frac{\sum_{i=1}^n Y_i \exp\left[-(X - X_i)^T (X - X_i) / 2\sigma^2\right]}{\sum_{i=1}^n \exp\left[-(X - X_i)^T (X - X_i) / 2\sigma^2\right]}. \quad (4)$$

The estimated $\hat{Y}(X)$ can be visualized as a weighted average of all the observed value, Y_i , where each observed value is weighted exponentially according to its Euclidean distance from X . When the smoothing parameter σ is made large, the estimated density is forced to be smooth. As σ becomes very large, $\hat{Y}(X)$ assumes the value of Y_i associated with the observation closest to x . For intermediate value of σ , all values of Y_i are taken into account, but those corresponding to points closer to X are given heavier weight.

2.1.2. GRNN Network Structure. A GRNN is a four-layer feedforward neural network based on nonlinear regression theory: input layer, pattern layer, summation layer, and output layer [13] as shown in Figure 1. The input layer merely transports the input variables (data attributes) to the pattern layer. The number of input variables in input layer depends on the total number of the observation parameters. In this

study, the diagnosis and classification of alcoholic liver disease were established by generalized regression neural network model, in which the input layer contains seven variables. The number of neurons in the pattern layer is equal to the number of training samples, n , while each neuron corresponds to different training samples [10, 11]. The summation layer includes two units: the first unit sums all the outputs of the pattern layer and evaluates the numerator of (4), and the second unit evaluates the denominator of (4). Each unit in the pattern layer is connected to each of the two units in the summation layer. The output unit merely divides the two outputs of the summation layer to get the predicted value. In this study the output layer contains three parameters: alcoholic cirrhosis with decompensated liver function, alcoholic cirrhosis with compensated liver function, and alcohol hepatitis.

2.2. Case Study

2.2.1. Patient Data. The data for this study was collected from 200 patients who showed symptoms of alcoholic liver diseases. All patients were male, aged 20 to 55 years with average of 43 years. These patients had a disease duration of 5 to 30 years. The dataset contained clinic symptoms, biochemical manifestations, and pathologic manifestations of the 121 patients. Based on several different diagnostic criteria for alcoholic liver disease, several experts diagnosed the 121 patients. 34 cases were clinically diagnosed for alcoholic hepatitis, 30 cases were clinically diagnosed for alcoholic liver cirrhosis with compensated liver function, and 57 cases were clinically diagnosed for alcoholic liver cirrhosis with decompensated liver function. Seven liver function tests, such as gamma glutamyl transpeptidase (GGT), total bile acid (TBA), alkaline phosphatase (ALP), total bilirubin (TBIL), alanine aminotransferase (ALT), aspartate aminotransferase (AST), and aspartate/alanine (AST/ALT), were used for computational diagnosis and classification using the GRNN model.

2.2.2. Method. In this study, the alcoholic liver diseases were diagnosed and classified with GRNN model. The input layer contained seven variables: GGT, TBA, ALP, TBIL, ALT, AST, and AST/ALT, which correspond to X_1 , X_2 , X_3 , X_4 , X_5 , X_6 , and X_7 in the GRNN, respectively. The output layer had three parameters: alcoholic cirrhosis with decompensated liver function, alcoholic liver cirrhosis with compensated liver function, and alcohol hepatitis, represented using (1, 0, 0) (0, 1, 0), and (0, 0, 1), respectively. Among the 121 samples, 110 samples were selected as training samples and 11 samples were diagnosing samples to test the rationality of the model (Table 1).

MATLAB R2008a is used to implement the GRNN methods. There are five procedures to establish the GRNN model to diagnose and classify alcohol hepatitis. Firstly, before training, the data is normalized. The inputs and the targets always fall within a specified range as following:

$$X_n = \frac{2(x - \min p)}{\max p - \min p} - 1, \quad (5)$$

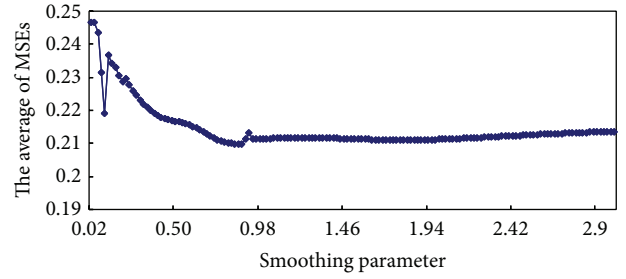


FIGURE 2: The relationship between σ and average of MSEs for 10 separate tests.

where $\min p$ and $\max p$ are the minimum and maximum values in a pattern.

Secondly, k -fold cross-validation and receiver operating characteristic curves are by far the most popular error-estimation approaches in machine learning [14, 15]. In this paper, we used tenfold cross-validation to demonstrate the error rate of networks.

Thirdly, the smoothing parameter σ is selected according to the minimum of the average of MSEs (mean square error) for k separate tests (k -fold cross-validation).

Fourthly, the GRNN model to diagnose and classify alcohol hepatitis is established according to the selected smoothing parameter σ . Finally, the established GRNN model is used to diagnose and classify alcohol hepatitis.

3. Results

A K -folding scheme with $K = 10$ is applied (using K -folding means dividing the dataset to k sets and use $k - 1$ of them for training and one for testing. This is repeated k times and then the average of the result is taken). The training procedure for each dataset is repeated 10 times, each time with 90% of the patterns as training and 10% for testing. All the reported results are obtained by averaging the outcomes of the 10 separate tests. The mean square error (MSE) is used to evaluate the GRNN model performance.

In this study, the smoothing parameter σ is selected according to the minimum of the average of MSEs for 10 separate tests (A K -folding scheme with $K = 10$). Figure 2 shows the relationship between σ and average of MSEs for 10 separate tests. We find that the average of MSEs for 10 separate tests is 0.209789 at the lowest points when the smoothing parameter σ is 0.86. Therefore, the smoothing parameter 0.86 is used to establish GRNN model. The training procedure for each dataset is repeated 10 times when each time with 90% from the patterns as training samples and 10% from the patterns used as testing samples. We find that the GRNN models have a good performance (MSE in the range (0.16992, 0.27269)).

Using the established GRNN model, the 110 training samples were back diagnosed and given an accuracy rate of 80.91% (Table 2). 11 diagnosing samples were applied for diagnosis and given an accuracy rate of 81.82% (Table 3). Furthermore, Fisher's linear discriminant is also used to diagnose and classify alcohol hepatitis. The coefficients of

TABLE I: The seven indicators of liver function tests and clinical diagnosis.

Case number	Liver function tests							Diagnosis		
	GGT X_1	ALT X_2	AST X_3	AST/ALT X_4	TBA X_5	ALP X_6	TBIL X_7	Y_1	Y_2	Y_3
1	1081.1	44.6	86.3	1.94	52.1	81.3	71.4	1	0	0
2	605.3	36.4	115.5	3.17	101.8	120.8	81.3	0	1	0
3	76	20.1	30.3	1.51	32.2	103.5	21.2	0	1	0
4	111.9	33.8	92	2.72	126.5	128.5	377	1	0	0
5	60.4	22.5	41.5	1.84	105	116.4	89.8	1	0	0
6	95.7	51.1	68.5	1.34	31.1	77.2	46.3	1	0	0
7	262.4	61.5	187.2	3.04	151.8	456	1234.2	1	0	0
8	1016	52.9	79.3	1.5	206.7	298.6	53.5	0	1	0
9	890	48.5	131.7	2.72	24.6	640	90	0	1	0
10	175.6	88.3	96.9	28.9	132.9	171.8	202.2	0	1	0
11	121.9	42.3	48	1.13	14.8	114.4	114.4	0	1	0
12	166.6	15.3	31.2	2.04	3.4	144.7	57.5	0	1	0
13	157.5	14.5	30	2.07	1.3	133.3	43	0	1	0
14	57.9	39.3	20.8	0.53	1	68.3	8.7	0	0	1
15	457.4	26.7	251	9.4	52.5	205.8	20.4	1	0	0
16	349.9	17.9	123.8	6.92	17.6	178.6	16.2	1	0	0
17	322.8	14.2	71.3	5.02	33.1	236.2	23.7	1	0	0
18	248.6	8.5	46.8	5.51	69.9	311.8	13.2	1	0	0
19	297.1	49.3	73.4	1.49	33.8	137.3	56	1	0	0
20	57.3	39.4	59.7	1.52	3	79.3	48.9	1	0	0
21	70	113.4	140	1.23	36.6	118.8	40	1	0	0
22	653.6	28.7	58	2.02	34.5	160.7	33.4	1	0	0
23	451.6	31.7	57.8	1.82	9	191.2	34.4	1	0	0
24	151.1	37.3	83.3	2.23	1498	75.6	103.8	0	0	1
25	140.4	47.7	52.5	1.1	1.6	74.2	40.1	0	1	0
26	79.4	43.7	28.4	0.65	0.8	110.7	12.4	0	1	0
27	28.4	28.2	27.8	0.99	1.2	45.6	9.2	0	1	0
28	104.2	25.7	44	1.71	33.4	110.3	9.7	1	0	0
29	314.9	226.3	858.1	3.79	77.9	64.6	445	1	0	0
30	268	3.8	46.3	12.18	113.1	206.1	743.5	1	0	0
31	72.8	69.8	0.96	5.1	3.8	187.8	176.4	0	0	1
32	39.6	50.5	1.28	28.1	45.1	273	101.8	1	0	0
33	90.9	24.5	49	2	77.2	199.5	55.5	1	0	0
34	78.9	23.1	47	2.03	89.7	111	76.2	1	0	0
35	69.2	20.6	42.1	2.04	104	159.5	57.4	1	0	0
36	20.6	28.9	48.6	1.68	5.8	78.5	70.4	1	0	0
37	26.7	36.5	60	1.64	38.7	86.9	57.4	1	0	0
38	214.3	33.9	102.9	3.04	106.3	261.6	728.1	1	0	0
39	200.5	21.1	64.3	3.05	121.9	217.9	273.7	1	0	0
40	127.4	20.7	66.6	3.22	164.3	232.6	128.7	1	0	0
41	335.6	188.6	206.8	1.1	242.1	166.1	527.9	0	0	1
42	323.6	29.8	61.9	2.08	166.8	341.2	108.9	0	0	1
43	257.1	1224.6	574.8	0.47	269.5	28.3	118.1	0	0	1
44	2406.6	145.3	303.2	2.09	15.1	157.3	25.9	0	0	1
45	2319	192.8	342.3	1.78	42.3	210.3	33.8	0	0	1
46	1831.2	192	215.7	1.12	76.9	181.8	20.9	0	0	1
47	1804.2	202.7	204	1.01	59.9	204.1	16.3	0	0	1
48	901.9	72.4	69.6	0.96	16.5	171.7	11	0	0	1
49	50.7	13.2	15.8	1.2	9.6	41.7	18	0	0	1

TABLE 1: Continued.

Case number	Liver function tests							Diagnosis		
	GGT X_1	ALT X_2	AST X_3	AST/ALT X_4	TBA X_5	ALP X_6	TBIL X_7	Y_1	Y_2	Y_3
50	32.9	11.7	15.8	1.35	5.7	35.8	13.9	0	0	1
51	35	41.2	48.5	1.18	2.6	31.2	22.1	0	0	1
52	41.4	26.6	50	1.88	51.4	122.6	19.9	1	0	0
53	15.8	22	38.6	1.75	3.4	124	16.8	0	0	1
54	44.1	30	57.3	1.74	163.9	185.8	52.3	1	0	0
55	88.2	21.3	17.7	0.83	3.9	81.1	20.5	1	0	0
56	102.3	22.6	21.4	0.95	5.1	121.8	18.5	1	0	0
57	1419.2	22.3	56.9	2.53	59.6	232.4	45.4	0	0	1
58	731.7	23.7	53.5	2.26	54.6	123.9	26.6	0	0	1
59	1038.6	52.9	100.2	1.89	59.6	179.3	41	1	0	0
60	196	90	69.6	0.77	1	57.6	16.2	0	0	1
61	146.6	45	34.8	0.77	2.1	52.6	13.5	0	0	1
62	40.5	13.7	14.9	1.09	3.5	65.6	11.7	0	0	1
63	56.4	48.2	29.2	0.61	23	84.4	5.3	0	0	1
64	100.6	19.3	21.1	1.09	35.5	111.4	84.1	0	0	1
65	98.6	20	92.9	4.65	139.5	93.7	331	0	1	0
66	213.6	21	102	4.65	123	162	187	0	1	0
67	767.2	24.7	75.8	3.07	16.6	89.7	15.4	1	0	0
68	641.1	32.5	84.2	2.59	38.4	117.5	13.1	1	0	0
69	230.1	29.2	32.8	1.12	40.7	101	14.3	1	0	0
70	224.6	17.4	45.9	2.64	23.8	104.7	28.1	1	0	0
71	526.9	43.9	23.5	0.54	3.4	5.9	19.4	0	0	1
72	782.6	41.1	143.3	3.49	49.8	179.5	42.3	0	0	1
73	336	75	95	1.27	76.7	128	49.74	1	0	0
74	34	19	41	2.16	25.7	86	21.4	1	0	0
75	31	40	75	1.88	28.3	121	36.09	1	0	0
76	56	18	39	2.17	27.5	39	24.31	1	0	0
77	88	46	79	1.72	16.9	65	37.47	1	0	0
78	923	28	156	5.57	60.2	286	67.95	1	0	0
79	36	18	28	1.56	9.2	72	27.24	1	0	0
80	46	32	32	2.9	11.9	92	15.71	1	0	0
81	331	36	40	0.91	26.4	165	26.11	1	0	0
82	35	45	76	1.98	26.3	123	31.09	1	0	0
83	115	62	114	1.84	29.9	61	33.37	1	0	0
84	43	25	30	1.2	14.2	69	2.1	1	0	0
85	27	7	20	2.86	12.4	33	19.9	1	0	0
86	42	26	31	1.19	14.8	83	15.16	1	0	0
87	140	16	23	1.44	35.1	64	32.22	1	0	0
88	47	31	29	0.94	12.9	93	14.71	1	0	0
89	335	74	94	1.57	74.7	118	44.74	0	0	1
90	24	28	29	2.72	13.5	91	23.36	0	0	1
91	58	25	37	2.57	26.5	99	23.31	0	0	1
92	46	58	31	0.53	14.2	117	33.6	0	0	1
93	59	48	40	0.83	13.9	106	27.6	0	0	1
94	63	8	22	2.75	7.7	95	25.53	0	0	1
95	38	37	79	1.13	18.4	109	69.67	0	1	0
96	46	36	31	0.84	12.5	123	15.71	0	1	0
97	39	22	27	1.59	11.2	91	17.24	0	1	0
98	36	38	81	2.13	16.4	108	67.67	0	1	0

TABLE I: Continued.

Case number	Liver function tests							Diagnosis		
	GGT X_1	ALT X_2	AST X_3	AST/ALT X_4	TBA X_5	ALP X_6	TBIL X_7	Y_1	Y_2	Y_3
99	12	11	1	0.09	16.3	58	14.2	0	1	0
100	682	24	72	3	65.4	135	43.49	0	1	0
101	14	23	28	1.22	13	99	19.26	0	1	0
102	417	26	81	3.1	55.6	127	39.26	0	1	0
103	85	28	27	0.96	27.83	19.4	41.73	0	1	0
104	392	24	75	3.13	54.6	137	40.26	0	1	0
105	432	62	119	1.92	28.8	83	47.58	1	0	0
106	22	28	54	1.93	44.3	187	35.58	1	0	0
107	62	56	29	0.52	8.4	99	31.11	0	0	1
108	301	37	41	1.11	27.4	175	25.11	0	0	1
109	391	58	60	1.03	18.1	93	67.46	0	1	0
110	662	21	74	3.01	66.4	134	33.49	0	1	0
111*	97.3	42.5	112.4	2.64	78.6	99	64.1	1	0	0
112*	328.6	40.3	163.3	4.03	67.1	103.8	146.9	0	1	0
113*	42.6	38.1	49.2	1.29	18.7	84.2	79.8	1	0	0
114*	16.9	28	1.66	12.3	37.6	146.5	68	1	0	0
115*	70.1	130.3	1.86	29.6	8.3	487.5	165	1	0	0
116*	12.5	24.1	1.93	35.5	1.9	140.6	127.6	0	1	0
117*	11.1	21.6	1.95	26.7	5.7	127.6	131.2	0	1	0
118*	14.3	25.4	1.78	33.1	7.5	155.5	164.9	0	1	0
119*	14.3	25.4	1.78	33.1	7.5	155.5	164.9	0	1	0
120*	1972	81	173	2.14	34.3	226	40.4	0	0	1
121*	514	342	550	1.61	37.5	368	57.46	0	0	1

Note: * samples need diagnosis.

Fisher's linear discriminant functions were listed in Table 4. And the classify results were listed in Tables 2 and 3. It shows the 110 training samples were back diagnosed and given an accuracy rate of 54.55% (Table 2). 11 diagnosing samples were applied for diagnosis and given an accuracy rate of 72.72% (Table 3).

The GRNN diagnosis showed high consistency with clinical diagnosis and classification made by experts. These results showed the excellent modeling abilities of our GRNN model in alcohol hepatitis diagnosis and classification.

4. Discussions

It is important to classify alcoholic liver diseases because different disease stages, such as active hepatitis and cirrhosis, require significantly different therapeutic strategies [3]. This paper presents a study on diagnosis and classification of alcoholic hepatitis by general regression neural network (GRNN) pattern recognition. Results obtained from GRNN were compared to clinical diagnosis made by a group of experts. The accuracy of back diagnosis of the 110 training samples was 80.91% and the accuracy of the predicting-diagnosis of the 11 samples was 81.82% in accordance with clinical diagnosis made by experts. It was seen that GRNN could be successfully used to help diagnose and classify alcoholic hepatitis.

The architecture of the neural network used in this study is the multilayered feedforward network architecture with

7 input variables and 3 output indicators. The number of input variables was determined by 7 liver function tests and the number of output indicators is represented as the classification of the disease development stages. Using this GRNN model, alcoholic liver disease could be accurately staged. Notably, the changes of these liver function tests are not disease specific. They can only reflect the degree of liver injury and substitute disease development stages. This is because any factors, such as viruses and toxicities, could induce liver damages resulting in similar changes in liver function tests. Therefore, these variables could not be used to distinguish hepatitis caused by different etiologies. In this study we demonstrated that using these variables for the input layer of GRNN model successfully classified the alcoholic liver disease. This is the first report using GRNN model with liver function tests for the classification and diagnosis of alcoholic liver disease. Due to the similarity in liver function damage, we proposed that using the liver function tests and the GRNN model can also classify disease stages for other kinds of hepatitis.

5. Summary

We developed a general regression neural network model which accurately diagnosed and classified alcoholic hepatitis. Seven liver function tests were used for input variables while three disease development stages were used for output

TABLE 2: Back diagnosis results of alcoholic liver disease by generalized regression neural network and Fisher's discriminant method.

Case number	GRNN			Diagnosis	
	Y_1	Y_2	Y_3	Fisher's discriminant method	Actual class
1	0.7333	0.0722	0.1945	3	1
2	0.2441	0.6648	0.0911	1	2
3	0.4268	0.2788	0.2944	1	2
4	0.923	0.077	0	1	1
5	0.6669	0.1676	0.1655	1	1
6	0.5903	0.1971	0.2126	1	1
7	1	0	0	1	1
8	0.0005	0.9995	0	3	2
9	0	1	0	1	2
10	0	1	0	2	2
11	0.3795	0.3947	0.2258	1	2
12	0.3564	0.4832	0.1605	1	2
13	0.3843	0.4087	0.207	1	2
14	0.3364	0.2546	0.4091	1	3
15	1	0	0	1	1
16	0.9994	0.0006	0	1	1
17	0.9802	0.007	0.0128	1	1
18	0.9998	0	0.0002	1	1
19	0.5994	0.2525	0.1481	1	1
20	0.557	0.2059	0.2371	1	1
21	0.926	0.0587	0.0153	1	1
22	0.4967	0.4358	0.0675	1	1
23	0.4623	0.0536	0.4841	1	1
24	0	0	1	3	3
25	0.4885	0.2275	0.284	1	2
26	0.332	0.342	0.3259	1	2
27	0.3165	0.2665	0.4169	1	2
28	0.4672	0.2703	0.2624	1	1
29	1	0	0	1	1
30	1	0	0	1	1
31	0.0011	0.0004	0.9986	1	3
32	1	0	0	2	1
33	0.9744	0.02	0.0056	1	1
34	0.6255	0.1865	0.188	1	1
35	0.8381	0.0996	0.0623	1	1
36	0.5616	0.1982	0.2402	1	1
37	0.581	0.2021	0.217	1	1
38	1	0	0	1	1
39	0.9932	0.0064	0.0004	1	1
40	0.9983	0.0015	0.0003	1	1
41	0	0	1	1	3
42	0.0009	0	0.9991	1	3
43	0	0	1	3	3
44	0	0	1	3	3
45	0	0	1	3	3
46	0.0053	0	0.9947	3	3
47	0.0298	0	0.9702	3	3
48	0.711	0.1472	0.1418	3	3

TABLE 2: Continued.

Case number	GRNN			Diagnosis	
	Y_1	Y_2	Y_3	Fisher's discriminant method	Actual class
49	0.3041	0.2591	0.4368	1	3
50	0.2854	0.2651	0.4495	1	3
51	0.3327	0.2919	0.3753	1	3
52	0.5193	0.2386	0.2421	1	1
53	0.4097	0.2898	0.3006	1	3
54	0.9936	0.0042	0.0022	1	1
55	0.3821	0.2598	0.3581	1	1
56	0.3659	0.3531	0.281	1	1
57	0.9549	0.0244	0.0207	3	3
58	0.4492	0.4788	0.072	3	3
59	0.1617	0.1325	0.7058	3	1
60	0.3793	0.1808	0.4399	1	3
61	0.3573	0.2243	0.4184	1	3
62	0.3651	0.2475	0.3874	1	3
63	0.3601	0.2679	0.372	1	3
64	0.2964	0.3042	0.3993	1	3
65	0.0702	0.9298	0	1	2
66	0.0148	0.9844	0.0008	1	2
67	0.6305	0.3154	0.0541	3	1
68	0.4137	0.5228	0.0635	1	1
69	0.5176	0.2629	0.2195	1	1
70	0.5594	0.2479	0.1927	1	1
71	0.0018	0.0072	0.991	3	3
72	0.1715	0.4762	0.3523	1	3
73	0.6397	0.0706	0.2898	1	1
74	0.4973	0.2033	0.2993	1	1
75	0.5845	0.2201	0.1953	1	1
76	0.4607	0.1992	0.3401	1	1
77	0.6562	0.1594	0.1845	1	1
78	0.9975	0	0.0025	1	1
79	0.4482	0.2218	0.33	1	1
80	0.4704	0.1748	0.3548	1	1
81	0.4815	0.1543	0.3643	1	1
82	0.5944	0.2115	0.1941	1	1
83	0.8249	0.0871	0.088	1	1
84	0.4062	0.2358	0.358	1	1
85	0.5017	0.1501	0.3482	1	1
86	0.4166	0.2545	0.3289	1	1
87	0.4653	0.2036	0.3311	1	1
88	0.3803	0.287	0.3327	1	1
89	0.6071	0.0856	0.3074	1	3
90	0.4651	0.185	0.35	1	3
91	0.3184	0.2007	0.4809	1	3
92	0.3226	0.3469	0.3304	1	3
93	0.3806	0.307	0.3124	1	3
94	0.4373	0.1971	0.3656	1	3
95	0.5349	0.2874	0.1777	1	2
96	0.3463	0.3551	0.2986	1	2
97	0.4186	0.2587	0.3227	1	2

TABLE 2: Continued.

Case number	Diagnosis				Actual class
	Y_1	GRNN Y_2	Y_3	Fisher's discriminant method	
98	0.6125	0.2184	0.1692	1	2
99	0.2418	0.4192	0.339	1	2
100	0.2815	0.6907	0.0278	1	2
101	0.3848	0.296	0.3192	1	2
102	0.3386	0.4991	0.1623	1	2
103	0.2116	0.3976	0.3909	1	2
104	0.4748	0.3862	0.139	1	2
105	0.6053	0.1463	0.2483	1	1
106	0.9314	0.0413	0.0273	1	1
107	0.3397	0.3032	0.3571	1	3
108	0.4228	0.1859	0.3913	1	3
109	0.4611	0.36	0.179	1	2
110	0.2598	0.7091	0.0311	1	2

Note: Bold indicates diagnosis; 1, 2 and 3 represent alcoholic cirrhosis with decompensated liver function, alcoholic liver cirrhosis with compensated liver function, and alcohol hepatitis, respectively; similarly hereinafter.

TABLE 3: Diagnostic results of the tested samples.

Case No.	Diagnosis				Actual class
	Y_1	GRNN Y_2	Y_3	Fisher's discriminant method	
1	0.7835	0.1375	0.079	1	1
2	0.1992	0.7026	0.0982	1	2
3	0.54	0.2244	0.2356	1	1
4	0.0105	0.0001	0.9893	1	1
5	1	0	0	2	1
6	0.0021	0.9979	0	2	2
7	0.0314	0.9686	0	2	2
8	0.013	0.987	0	2	2
9	0.013	0.987	0	2	2
10	0	0	1	3	3
11	0.5	0.5	0.5	1	3

TABLE 4: Coefficients of Fisher's linear discriminant functions.

Variables	Coefficients		
	1st	2nd	3rd
X_1	-0.001	-5.03E - 005	0.002
X_2	-0.001	0.000	0.007
X_3	0.006	0.004	0.001
X_4	0.028	0.112	0.000
X_5	0.002	0.002	0.004
X_6	0.017	0.014	0.013
X_7	-0.001	-0.001	-0.002
(Constant)	-2.154	-2.817	-3.017

Note: 1st, 2nd, and 3rd express the first, second, and third functions of Fisher's linear discriminant method, respectively.

indicators. The clinical diagnosis was confirmed by a group of experts. The accuracy of the back diagnosis in the 110 training samples and the accuracy of predicting-diagnosis in the 11 samples were evaluated. Using GRNN, ALD could be

accurately diagnosed and classified and the accuracy could be around 80.91%–81.82%. GRNN also reduces diagnosis time and improves the diagnosis accuracy which will benefit young physicians. However, it should be noted that the current international classification of alcoholic liver disease still follows no single standard, and this paper is limited because of a small size of cases. Also, we only consider three categories for the classification and more categories need further study.

Conflict of Interests

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

Acknowledgment

The authors would like to acknowledge three anonymous reviewers for their valuable comments, which led to a significant improvement over an early version of the paper.

References

- [1] R. S. O'Shea and A. J. McCullough, "Treatment of alcoholic hepatitis," *Clinics in Liver Disease*, vol. 9, no. 1, pp. 103–134, 2005.
- [2] Y. M. Li, J. G. Fan, B. Y. Wang et al., "Guidelines for the diagnosis and management of alcoholic liver disease," *Journal of Digestive Diseases*, vol. 12, no. 1, pp. 45–50, 2011.
- [3] M. D. Zeng, Y. M. Li, C. W. Chen et al., "Guidelines for the diagnosis and treatment of alcoholic liver disease," *Journal of Digestive Diseases*, vol. 9, pp. 113–116, 2008.
- [4] H. Nakano, Y. Okamoto, H. Nakabayashi, S. Takamatsu, H. Tsujii, and H. Matsuoka, "Application of neural network to the interpretation of laboratory data for the diagnosis of two forms of chronic active hepatitis," *International Hepatology Communications*, vol. 5, no. 3, pp. 160–165, 1996.
- [5] Y. S. Kim, S. Y. Sohn, and C. N. Yoon, "Screening test data analysis for liver disease prediction model using growth curve," *Bio-medicine and Pharmacotherapy*, vol. 57, no. 10, pp. 482–488, 2003.
- [6] M. S. Bascil and F. Temurtas, "A study on hepatitis disease diagnosis using multilayer neural network with Levenberg Marquardt training algorithm," *Journal of Medical Systems*, vol. 35, no. 3, pp. 433–436, 2011.
- [7] R. H. Lin, "An intelligent model for liver disease diagnosis," *Artificial Intelligence in Medicine*, vol. 47, pp. 53–62, 2009.
- [8] R.-H. Lin and C.-L. Chuang, "A hybrid diagnosis model for determining the types of the liver disease," *Computers in Biology and Medicine*, vol. 40, no. 7, pp. 665–670, 2010.
- [9] S.-M. Chou, T.-S. Lee, Y. E. Shao, and I.-F. Chen, "Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines," *Expert Systems with Applications*, vol. 27, no. 1, pp. 133–142, 2004.
- [10] L. Dong and X. Li, "An application of grey-general regression neural network for predicting landslide deformation of dahu mine in China," *Advanced Science Letters*, vol. 6, pp. 577–581, 2012.
- [11] D. F. Specht, "A general regression neural network," *IEEE Transactions on Neural Networks*, vol. 2, no. 6, pp. 568–576, 1991.
- [12] E. Parzen, "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, 1962.
- [13] S. A. Hannan, R. P. Manza, and R. J. Ramteke, "Generalized regression neural network and radial basis function for heart disease diagnosis," *International Journal of Computer Applications*, vol. 7, pp. 7–13, 2010.
- [14] N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*, Cambridge University Press, 2011.
- [15] L. Dong and X. Li, "Comprehensive models for evaluating rockmass stability based on statistical comparisons of multiple classifiers," *Mathematical Problems in Engineering*, vol. 2013, Article ID 395096, 9 pages, 2013.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

