A New Application of Fuzzy Neural Network to ECG Recognition based on Rough Set Theory

Xianming Huang, Yanhong Zhang

Electronic Information and Control Engineering College, Beijing University of Technology
100022 Beijing, China
huangxm@email.com

Abstract-- **When recognizing and analyzing wave signals in electrocardiogram (ECG) automatically, present methods have some limitation and do not meet physicians' expectations. A new application of fuzzy neural network to ECG recognition based on rough set theory is presented. Firstly the recognition rules for characteristic points in ECG are reduced by using rough set theory. Then the reduced rules are used to construct the fuzzy neural network structure whose membership functions are optimized by supervisal study method. Finally, the FNN is employed in recognizing characteristic points in ECG. There are some discussions about correlative arithmetic such as sizer method, difference method and how to choose difference parameters. Also in this paper, an example of this new application is discussed. We adopt MIT-BIH data to verify R wave recognition and the detection rate is proved higher than the one of routine recognition method.**

***Index terms*—ECG recognition, rough set, FNN**

Ⅰ. INTRODUCTION

Electrocardiogram (ECG) represents the recording of the electrical potential of the heart. It is a standard tool used to diagnose heart disease. The Holter ECG device is used most frequently for recording the ECG. Physicians apply the device to a patient when they need to monitor his/her ECG to find the few abnormal cycles in the ECG throughout the day. A typical cycle of an ECG is shown in Fig.1. Physicians first locate such characteristic points as Q points, R points and S points in the ECG from which they locate P complexes, QRS waves, T complexes and U waves in the ECG. These waves and complexes are defined in Fig.1. Physicians then calculate the parameters of those waves and complexes to determine whether the ECG shows signs of cardiac disease or not. The parameters are the amplitude and the interval of each wave, such as RR interval, PP interval, QT interval and ST segment. But recognition of the characteristic points and calculations of the parameters is a tedious routine for the physician; nearly 100000 cardiac cycles per patient are recorded in a day with the Holter device. The physician has to analyze this large amount of ECG data to search for only a few abnormal cardiac cycles in the ECG. So there is an urgent need for an automatic ECG analyzing system to lighten the burden of physicians. Some automatic ECG analyzing systems have been developed. But an ECG analyzing system that is good enough to meet physicians' expectations has not yet been developed because it is difficult to locate the characteristic points with a computer. Also, the shape of the ECG varies with each patient. The ECG recorded from the same patient changes as time passes.

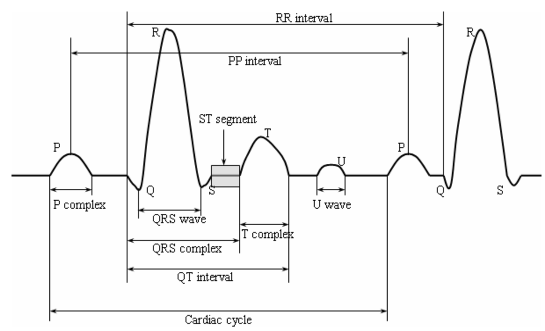

Fig. 1. A cardiac cycle of an ECG

Since the characteristic points are the reference points used to recognize other waves and complexes, the first and key step is to recognize the characteristic points for implementing an automatic ECG analyzing system. The prevailing methods of recognizing the characteristic points include filter method, the syntactic method and soft compute method.

Filter method is a simple and high-speed arithmetic and can be embedded into much hardware such as MCU easily. So it is suitable for clinical monitors. The ECG signals are driven through a special filter designed for such a special cycle as QRS wave. When a QRS wave arrives, the filter will export a pulse output whose amplitude reflects the compositive characteristic information including the amplitude and frequency of the QRS wave. This method is sensitive to noise signals and has a high miscalculation rate. When the ECG signals have serious movement fake track or baseline excursion, the distinguishing performance of filter method will be affected seriously.

Another method of recognizing the characteristic point is the syntactic method. In this method, the first step is the determination of a set of primitives. These primitives should provide a compact but adequate description of the ECG in terms of the specified structural relations. In the next step we determine parameters to represent the primitives for pattern recognition. This method is based on the assumption that the ECG is composed of peaks and segments which are primitives to constitute the ECG. However, primitive selection is both problem- and pattern-dependent and there is no general solution to this problem. The primitives constituting the ECG will vary slightly with each patient and the patient's condition. The primitives determined before ECG processing may be inadequate to describe the ECG for different patients. In such case, the syntactic methods may not work well, because the ECG patterns that are assumed to emerge by the syntactic method may change with time.

Soft compute method generally indicates the wavelet transform method. The wavelet transform method has particular advantage of analyzing changed signals. In this method the signal is decomposed into different-scale linear transforms. The characteristic points and waves can be determined by analyzing power spectrum density of ECG signal and the relation between the wavelet transform scale and the signal's frequency. But this method has too large compute quantity and has not been used to clinical application.

Studies on neural networks have been actively pursued. There have been a number of industrial applications of neural networks. There are some neural networks which have learning and self-organizing abilities. Neural networks have the ability to learn patterns in response to newly input patterns. Those learning and self-organizing abilities are appropriate for ECG recognition, because the ECG signals will change their shapes according to the patient's physical condition. The neural network trained well can classify all kinds of ECG signals automatically. In this paper, a new application of fuzzy neural network to ECG recognition based on rough set theory is presented. Firstly the recognition rules for characteristic points in ECG are reduced by using rough set theory. Then the reduced rules are used to construct the fuzzy neural network structure whose membership functions are optimized by supervisal study method. Finally, the FNN is employed in recognizing characteristic points in ECG and an example of this new application is discussed. We adopt MIT-BIH data to verify R-wave recognition and the detection rate is proved higher than the one of routine recognition method.

II. THE SUMMARIZATION OF ROUGH SET THEORY

Rough set theory, introduced by Zdzislaw Pawlak in the early 1980s, is a new mathematical theoretic method to deal with knowledge's and data's representation, learning and induction. It develops an important research branch in the field of intelligent information processing. It doesn't rely on accessional information out of data set and analyzes, reasons, discovers reliant relation among data just from the point of view of data's discriminable attribute. It provides convenience for knowledge reduction. The basic idea of rough set theory is exporting decision-making or classified rule of some concept by knowledge reduction on the premise that classifying ability of information system hold the line.

In rough set theory, a basic tool to represent and dispose knowledge is decision-making table. Research object of rough set is an object set which is often described by attribute-value couple. That set is denoted as U. In a general way, a decision-making table knowledge representation system can be denoted as $S=<U,R,V,f>$. Here U is a set of objects, which is also called domain. $R=C\bigcup D$ is attribute set, and subset C and D is called condition attribute set and decision-making attribute set respectively, $D\neq\Phi$. $V=\bigcup V_r\,(r\in R)$ is the set of attribute value. V_r is value domain of attribute r. $f:U\times R\rightarrow V$ is an information function which designates the attribute value of each object x in U. For each attribute subset $B\subseteq R$, an undiscriminable dualistic relation (indiscernibility relation) $IND(B)$ can be defined as

$$IND(B)=\{x,y\,|\,(x,y)\in U^2\forall b\in B(b(x)=b(y))\}.$$

That how to acquire harmonious rules furthest from decision-making table have to be researched when acquiring knowledge based on rough set. By reducing original decision-making table, a record from new reduction decision-making table represents a class of samples which have same rule characteristic and reliant relation between decision-making attribute and condition attribute. Some definitions are introduced.

Definition 1: Given knowledge representation system $S=<U,R,V,f>$, for each subset $X\subseteq U$ and indiscernibility relation B, lower/upper approximate set of X can be defined as

$$B^-(X)=\bigcup\{Y_i\,|\,(Y_i\in U\,|\,IND(B)\wedge Y_i\bigcap X\neq\Phi)\}$$

$$B_-(X)=\bigcup\{Y_i\,|\,(Y_i\in U\,|\,IND(B)\wedge Y_i\subseteq X)\}$$, and

$$U\,|\,IND(B)=\{X\,|\,(X\subseteq U\wedge\forall x\forall y\forall b(b(x)=b(y)))\}$$

is U's partition by indiscernibility relation B.

Definition 2: Dualistic group $(B_-(X),B^-(X))$ is called rough set based on B.

Definition 3: $POS_B(X)=B_-(X)$ is called X's B positive domain.

We adopt general arithmetic to reduce decision-making table. The arithmetic to be used is as follows:

Let condition attribute set of original decision-making be $P=\{a_i\,|\,i=1,\cdots,n\}$, decision-making attribute set be $D=\{d\}$.

General arithmetic--

For each condition attribute a_i carry out the process until condition attribute set does not change.

{ When a_i is deleted ,

if $POS_{(P\setminus\{a_i\})}(D)=POS_P(D)$

then delete column a_i lied in and unite repeated rows

else hold a_i }

Upper and lower approach of decision-making attribute class in allusion to condition attribute class can be calculated in each rule. The believe degree of each rule can be calculated by following formula.

$$m(x_k) = \min \frac{\left|[x_k]_{R_i} \cap D_k\right|}{\left|[x_k]_{R_i}\right|}$$

Thereinto x_k represents the k th rule (k =1,2‥,**n**). D_k represents the decision-making attribute class of the k th rule. R_i represents the class in allusion to the i th condition attribute of the rule (i =1,2‥,**n**).

III. THE ARITHMETIC OF RECOGNIZING CHARACTERISTIC POINTS

A. The Principle of Difference Method

We apply difference method to recognizing the characteristic points in ECG for its high-speed and high-efficiency trait. Let x_i ($i = 1,2,\cdots,N$) be original data of the ECG. Since sampling data is not less than 2 seconds generally, and the sampling rate of MIT-BIH database which will be used in this paper is 360Hz, we select $N \geq 360 \times 2 = 720$. So we define k points first order difference as $y_i = x_i - x_{i-k}$ ($k \geq 1,\ k+1 \leq i \leq N$), k points second order difference as $z_i = y_{i+k} - y_i = x_{i+k} + x_{i-k} - 2x_i$, ($k \geq 1$, $k+1 \leq i \leq N-k$). The minimal amplitude and time limit of the ECG waves are determined by international institution as 20 mv voltage deflexion and 6ms time limit. This criterion accords with eyeballing resolving power of human being. So k can be calculated by the formula:

$$(k-1)/f \leq 6/1000$$

f is the sampling rate of the ECG. We get

$$1 \leq k \leq 1 + 360 \times 6/1000 = 3.16$$

In this paper we use k =3 to calculate the attributes of the characteristic points. The first order difference reflects the slope of a point. The second order difference reflects the change rate of the slope.

B. The Reduction of Recognition Rules for Characteristic Points based on Rough Set Theory

We apply decision-making table reduction arithmetic in rough set theory to obtaining recognition rules of characteristic points in ECG. Firstly the decision-making table of characteristic points is established. For each point 12 condition attributes are calculated:

c1. 1 point first order difference;
c2. 1 point second order difference;
c3. k points first order difference;
c4. k points second order difference;
c5. k points first order difference of the left n th point;
c6. k points second order difference of the left n th point;
c7. 1 point first order difference of the right n th point ;
c8. 1 point second order difference of the right n th point;
c9. k points first order difference of the right n th point;
c10. k points second order difference of the right n th point;
c11. 10 points first order difference;
c12. 10 points second order difference.

We select k =3, n =3 according to the forenamed representation. There is a decision-making attribute d in the decision-making table which represents the type of the characteristic point. 4 decision-making attribute values are chosen:

- The start point of P or T wave denoted as 1;
- The end point of P or T wave denoted as 2;
- The R wave crest denoted as 3;
- The P or T wave crest denoted as 4.

P and T wave are classified into the same class because they are similar from the point of view of the forenamed condition attributes. When practicing the recognition, they can be distinguished by their locations relative to the location of R wave. We adopt the data from the data file 101.dat in MIT-BIH database whose time segment is from 12 second to 14 second (just a demo, other data can also be adopted). Since the condition attributes are all continuous, discrete disposal must be implemented firstly. By calculating the difference value of each point, we found the difference values of most points fall into the region [-44, 80]. So we have the discrete disposal method denoted as follows:

Tab. 1. The Method of Discrete Disposal for the Difference

No.	Attribute	Values						
		0	1	2	3	4	5	6
1	first order difference	<-44	[-44, -10)	[-10, 0)	[0, 10)	[10, 30)	[30, 80)	>=80
2	second order difference	<-37.6	[-37.6, 0)	[0, 41.2)	[41.2, 80)	>=80		

17 sampling points are adopted and the original decision-making table for characteristic points in ECG is established. From Tab.2 we can see that 4 start points of P or T wave, 4 end points of P or T wave, 5 wave crest points of R wave and 4 wave crest points of P or T wave are adopted. We use forenamed decision-making table reduction arithmetic to reduce Tab.2 and get final decision-making table (Tab.3) for simplified rule. From Tab.3 we can obtain the reduced rules for recognizing 4 kinds of characteristic points. For example, the minimum decision-making rules for recognizing R wave crest are:

Rr1. c11=6 $\longrightarrow$ d=3 Rr2. c2=0 $\longrightarrow$ d=3
Rr3. c1=4 $\longrightarrow$ d=3 Rr4. c1=5 $\longrightarrow$ d=3

The reduction result is related to some factors. If these factors change, the reduction result may be adjusted accordingly. These factors are: the selection of condition attribute in decision-making table, the selection of condition attribute after the reduction and the selection of sampling points.

Tab. 2. Original Decision-making Table of Characteristic Points

No.	Condition Attribute												d
	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10	c11	c12	
1	3	2	2	2	2	2	3	1	3	2	2	1	1
2	3	1	3	2	3	1	3	1	4	1	3	2	1
3	2	2	2	2	2	2	3	1	4	1	2	1	1
4	3	2	4	1	2	2	3	2	3	2	4	3	1
5	2	2	1	2	2	1	2	2	3	2	1	1	2
6	2	1	2	2	1	2	2	2	2	1	1	1	2
7	3	2	2	2	1	2	2	1	3	2	1	1	2
8	3	1	2	2	1	2	3	1	2	2	1	1	2
9	4	0	6	0	3	4	1	2	0	4	6	5	3
10	4	0	6	0	4	4	1	2	0	4	6	6	3
11	5	0	6	0	2	4	1	2	0	4	6	5	3
12	4	0	6	0	4	4	1	2	0	4	6	6	3
13	5	0	6	0	2	4	1	2	0	4	6	5	3
14	3	1	4	0	4	2	2	2	1	2	5	4	4
15	3	1	4	1	3	2	2	1	2	1	4	3	4
16	3	1	4	1	3	2	2	2	2	1	5	4	4
17	3	1	3	1	4	1	2	2	2	1	4	3	4

Tab. 3. Final Decision-making Table for Simplified Rule

No.	Condition Attribute					d
	c1	c2	c5	c8	c11	
1	X	X	X	X	2	1
2	X	X	2	1	X	1
3	3	X	2	X	X	1
4	X	X	X	X	3	1
5	2	X	X	1	X	1
6	X	X	2	X	4	1
7	X	2	X	X	4	1
8	3	2	X	2	X	1
9	X	X	X	X	1	2
10	2	X	X	2	X	2
11	X	X	1	X	X	2
12	2	1	X	X	X	2
13	X	X	X	X	6	3
14	X	0	X	X	X	3
15	4	X	X	X	X	3
16	5	X	X	X	X	3
17	X	X	X	X	5	4
18	X	1	4	X	X	4
19	3	X	4	X	X	4
20	3	1	X	2	X	4
21	X	X	X	1	4	4
22	X	X	3	X	4	4
23	X	1	X	X	4	4
24	X	1	3	2	X	4
25	3	X	3	2	X	4
26	X	X	4	X	4	4

C. To Construct Fuzzy Neural Network Structure

The model of fuzzy neural network is：Let the network's input be x_1, x_2 (just a demo, more inputs can also be adopted)，output be u，then fuzzy relation between output and input can be described as:

$$R^i : if \;\; x_1 \; is \; A_1^i \;\; and \;\; x_2 \; is \; A_2^i \;\; then \;\; u \; is \; B^i .$$

FNN is a feedforward neural network which has 3 concealed layers (Fig.2).

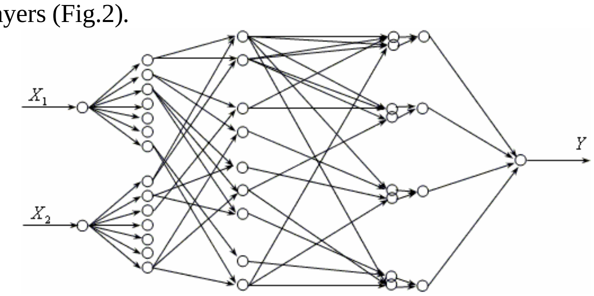

Fig. 2. The Structure of FNN

This is a kind of fuzzy decision-making system based on join mechanism. There are five layers in the system. The first layer is input layer and in this layer language variable $\{x_j\}$ is input. x_1 is the forenamed c2 and x_2 is the forenamed c11. The second layer is fuzzy layer and in this layer the node is morpheme node which can be used to represent membership function $m_{A_{ij}}(x_j)$.

$$m_{A_{ij}}(x_j) = e^{-[(\frac{w_s x_j - w_c}{1/w_d})^2]} \quad (i = 1,2,\cdots,7; j = 1,2,)$$

w_s is quantity factor. w_c is center element of membership function. w_d is scale factor of membership function. In the second layer the language nodes are connected with the corresponding morpheme nodes completely. The nodes of the third layer are rule nodes and have the function of fuzzy "and" operation. All nodes in the third layer form fuzzy rule base. The number of nodes in this layer is the number of initial rules derived from simplest decision-making table. The line connecting the third layer with the fourth layer has the rational function of join mechanism lest the rule matching process should happen. The line in the third layer prescribes premise condition of rule nodes and the line in the fourth layer prescribe conclusion of rule nodes. The fourth layer is rule decision-making layer and in this layer each node is rough nerve cell. Rough nerve cell can be looked upon as the combination of upper approach r^- and lower approach nerve cell r_-. The one connecting with r^- is upper approach rule condition of the decision-making in the condition of initial rule's condition class. The one connecting with r_- is lower approach rule condition of the decision-making. The link weight between two nerve cells is rough membership degree of rules (believe degree of rules). The output of r^- and r_- are:

$$outputr^- = \max(f(inputr^-), f(inputr_-))$$
$$outputr_- = \min(f(inputr^-), f(inputr_-))$$

The joint section of r^- and r_- represents information exchange. The input of r^- and r_- are transferred to next normal nerve c and the output of c is:

$$outputc = \frac{(outputr^- - outputr_-)}{(outputr^- + outputr_-)}$$

The link in the fourth layer carry out fuzzy "or "operation and compose weighted rules. The fifth layer is output layer and outputs network decision-making signal.

The relation of input and output in controller network are as follows:

Let the input of the j th node in the L th layer be I_j^L and output be O_j^L.

The first layer: $I_j^1 = w_{s\,j} \bullet x_j$， $O_j^1 = I_j^1$， $j = 1,2$

The second layer: $I_j^2 = O_1^1 - w_{c\,j}$， $j = 1,2,\cdots,7$ ；

$$I_j^2 = O_2^1 - w_{c\,j} \text{，} \; j = 8,9,\cdots,14 \text{ ；}$$

$$O_j^2 = w_{d\,j} \bullet e^{-(I_j^2)^2} \text{，} \; j = 1,2,\cdots,14$$

$j = 1,2,\cdots,7$ represent 7 fuzzy sets of temperature error which correspond to membership function $m_{1j}(x_1)$. $j = 8,9,\cdots,14$ represent 7 fuzzy sets of temperature error rate which correspond to membership function $m_{2(j-7)}(x_2)$.

The third layer: $I_k^3 = O_i^2 \bullet O_j^2$, $i = 1,2,\cdots,7$;

$j = 8,9,\cdots,14$; $O_k^3 = I_k^3, k = 7 \bullet (i-1) + (j-7)$

The fourth layer： $I_k^4 = 1 \Big/ (\sum_{k=1}^{49} O_k^3) \bullet O_k^3$， $O_k^4 = I_k^4$

The fifth layer: $I_1^5 = \sum_{k=1}^{49} (O_k^4 \bullet w_{b\,k})$, $O_1^5 = I_1^5$

According to these formulas, the output of rough-fuzzy neural network is:

$$u = \sum_{k=1}^{49} P_k(x_1, x_2) \bullet w_{b\,k} ,$$

$$P_k(x_1, x_2) = \frac{\prod_{j=1}^{2} m_{A_{jk}}(x_j)}{\sum_{k=1}^{49}\left(\prod_{j=1}^{2} m_{A_{jk}}(x_j)\right)}.$$

D. The Adjustment of Neural Network

The FNN adopts supervisal study method to optimize membership function. For FNN the error function is defined as

$$E = \frac{1}{2}\sum^{m}(T_0(t) - T(t))^2$$

m is the number of study samples derived from data samples by rough processing. $T_0(t)$ is anticipant output value of the system and $T(t)$ is actual output value. BP error back promulgate method adjusts weighted values of network so that the error function E has the minimal value. The parameters of membership functions or fuzzy rules are optimized. The membership function's center $w_{c\,i}$ and width $w_{d\,i}$ are used as adjusted parameter. Let s_j^l represent error back promulgate signal of the j th node in the L th layer, then the signal of each layer are:

The output layer:

$$w_{c\,i}(t+1) = w_{c\,i}(t) + h[T_0(t) - T(t)]\frac{w_{d\,i}u_i}{\sum w_{d\,i}u_i}$$

$$w_{d\,i}(t+1) = w_{d\,i}(t) +$$

$$h[T_0(t) - T(t)]\frac{w_{ci}u_i(\sum w_{d\,i}u_i) - (\sum w_{c\,i}w_{d\,i}u_i)u_i}{(\sum w_{d\,i}u_i)^2}$$

$$s_i^5 = T_0 - T$$

The fourth layer: $s_i^4(t) =$

$$[T_0(t) - T(t)]\frac{w_{c\,i}u_i(\sum w_{d\,i}u_i) - (\sum w_{ci}w_{d\,i}u_i)u_i}{(\sum w_{d\,i}u_i)^2}$$

The third layer: $s_i^3 = \sum_k s_k^4$

The second layer:

$$w_{d\,ij}(t+1) = w_{d\,ij}(t) - h\sum_k s_k^3 u_i \frac{2(u_i - w_{cij})^2}{w_{d\,ij}^3}$$

It has been proved that this method's convergence speed is faster than general method's by simulation experiment.

IV. A TEST OF THE ARITHMETIC

An example is presented which applies forenamed theories to recognizing R wave crest. We adopt the data from the data file 103.dat in MIT-BIH database whose time segment is from 12 second to 14 second. The forenamed reduced decision-making rules Rr1,Rr2 are used as restriction rules. By the forenamed arithmetic we can obtain the search set {290,291,292,619,620,621}. After the average processing the characteristic points of R wave crest are obtained as {291, 620}.

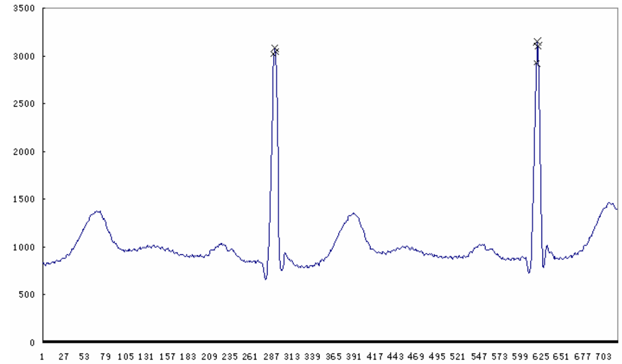

Fig. 3. The Search Set for R Wave Crest in 103.dat

Sometimes the results do not contain the characteristic point we wanted, maybe contain the nearby points. Under the circumstances sizer method is adopted. Quasi characteristic points of R wave crest are filtered according to two sizer principles: the amplitude of R wave crest is large; the interval between two neighboring R waves is usually not more than 0.25 second. The validity of this sizer method is proved by the following example.

Next we adopt the forenamed recognition arithmetic of R wave crest to recognize R wave. The sizer principles are also used. Three types of ECG data (Normal, PVC, BBB signals) of MIT-BIH database is adopted to test the detection rate of R wave by using the arithmetic of recognizing characteristic points based on rough set. The test results are as follows;

Tab. 4. The Detection Rate of R Wave in Normal ECG Data

database file	actual R wave	detected R wave	detection rate(%)
101.dat	1865	1850	99.2
103.dat	2084	2070	99.33
112.dat	2539	2480	97.68
115.dat	1953	1924	98.52
117.dat	1535	1504	97.98
121.dat	1863	1797	96.46
122.dat	2476	2455	99.15
230.dat	2256	2240	99.29

Tab. 5. The Detection Rate of R Wave in PVC ECG Data

database file	actual R wave	detected R wave	detection rate(%)
208.dat	2956	2550	86.27
221.dat	2427	2379	98.02
233.dat	3079	2945	95.65

Tab. 6. The Detection Rate of R Wave in BBB ECG Data

database file	actual R wave	detected R wave	detection rate(%)
109.dat	2532	2469	97.51
111.dat	2124	1872	88.14
118.dat	2278	2192	96.22
212.dat	2748	2712	98.69
231.dat	1571	1538	97.9

The average detection rate of R wave in Normal ECG signal can be calculated to be 98.49% from tab.4. The average detection rate of R wave in PVC ECG signal can be calculated to be 93.05% from tab.5. The average detection rate of R wave in BBB ECG signal can be calculated to be 95.82% from tab.6. So the total average detection rate of R wave is calculated to be 96.39%. From the concerned literature we find out that the average detection rate of R wave by routine recognition method is 95.65%. So the feasibility, practicability and validity of the new ECG recognition arithmetic based on rough set theory are proven.

V. CONCLUSION

We have demonstrated a new application of fuzzy neural network to ECG recognition based on rough set theory in this paper. Firstly the recognition rules for characteristic points in ECG are reduced by using rough set theory. Then the reduced rules are used to construct the fuzzy neural network structure whose membership functions are optimized by supervisal study method. Finally, the FNN is employed in recognizing characteristic points in ECG. As a result of experiments, the arithmetic was proven to be applicable to practical R wave recognition in an ECG. The R wave is a reference wave to recognize the successive significant points to implement an automatic ECG analyzing system.

REFERENCES

[1] G. Y. Wang, *Rough set theory and knowledge discovery*, China XJTU Press, 2001.12

[2] S. M. Szilagyi, "Event Recognition, Separation and Classification from ECG Recordings", *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 20, No. 1, 1998

[3] Y. Zhang, J. S. Wang, C. J. Wang, "Design of Intelligent Monitor for Newborns and Identification Arithmetic of QRS Wave from ECG Signals", *Journal of Anshan Institute of I. &S. Technology*, Vol. 24, No. 5, 2001.10

[4] L. F. Wen, Z. H. Meng, "New Developments of QRS – complex Detection Methods", *Biology Medicine Engineering*, Vol. 24, No. 5, 2001

[5] G. Y. Huang, M. C. Zou, "Application of Fuzzy Pattern Recognition Algorithm on Dynamic Cardiogram QRS Pattern Statistics", *Journal of Yanshan University*, Vol. 25, No. 3, 2001.7

[6] G. Li, W. Y. Ye, " An artificial-intelligence approach to ECG analysis", IEEE Engineering in Medicine and Biology Society, Vol. 22, No. 1, 2000.4

[7] L.M.Fu, L.C.Fu. "Mapping Rule-based Systems into Neural Architecture", *Knowledge Based Systems*, Vol.3, No.1, 1990

[8] L.M.Fu. "Knowledge-based Connectionism for Revising Domain Theories", *IEEE Trans.on System,Man,and Cybernetics*,Vol.23, No.1, 1993

[9] L.M.Fu. "Rule Generation from Neural Networks", *IEEE Trans.on Systems,Man,& Cybernetics*,Vol.24,No.8, 1994

[10]Pawlak Z. "Rough set", *J. of comput &inf science*, 1982

[11]Yi Jikai. *Intelligent Control Technology*, China BJPU press. 1999