Prediction of Acute Hypotensive Episodes Using Random Forest Based on Genetic Programming

Zhun Fan, *Senior Member, IEEE*, Youxiang Zuo, Dazhi Jiang and Xinye Cai, *Member, IEEE*

***Abstract*—At Intensive Care Unit (ICU), acute hypotensive episode (AHE) can cause serious consequences. It can make the organs broken, or even the patient dead. Generally AHE is predicted by the doctor clinically. In order to forecast the AHE automatically, this paper proposes an algorithm based on the genetic programming (GP) and random forest (RF). The algorithm obtains features of the signal through the Intrinsic Mode Function (IMF) signal produced by applying empirical mode decomposition (EMD) to the arterial blood pressure (MAP) signal. Then the feature sets and the data sets are grouped to evolve decision functions via GP. Finally, a random forest is formed and the classification result is obtained by voting. The achieved accuracy of the proposed method is 77.55%, the sensitivity is 80.55% and specificity is 75.14% after the five-fold cross-validation.**

***Index Terms*—acute hypotensive episode, empirical mode decomposition, genetic programming, random forest**

I. Introduction

HYPOTENSE is characterized by abnormal low blood pressure values clinically, which is also one of the recurrent situations at ICU. AHE is defined as any period of 30 minutes or more during which at least 90% of the mean arterial pressure (MAP) measurements are at or below 60 mmHg. The occurrence of AHE is a critical event in ICU. If not treated promptly and properly in time, it may cause an irreversible organ damage, and eventually death [1]. Thus, time is precious for the professionals, because if there is a method to predict the AHE, the doctors would have enough time to select an effective treatment to intervene patients with AHE.

Actually, for patients in the ICU, massive amounts of clinical information are collected in real-time, which include arterial blood pressure (ABP), heart rate (HR), and oxygen saturation (SO2), etc [1]. In this work we attempt to train a classifier to forecast AHE using the ABP data.

In 2001, J. Bassale [2] described the parametric and nonparametric methods to analyze and characterize ABP prior to hypotension episodes. He concluded that ABP variability and shape have the potential to predict hypotension. In 2002, C. Crespo et al. [3] hypothesized that changes in the ABP signal morphology could be detected prior to the onset of hypotension. They found that the variance of the ABP signal and the variance of the percussion wave slope were the significant factors when predicting the AHE. Also, in 2007, M. Saeed [4] introduced a temporal similarity metric based on transforming time series data into an intuitive symbolic representation. They used wavelet decomposition to characterize time series dynamics at multiple time scales. Their algorithm was employed to identify similar physiologic patterns in hemodynamic time series from ICU patients, with a potential to be used in the detection of imminent hemodynamic deterioration [1]. In 2008, L. Lehman [5] present a similarity-based searching and pattern matching algorithm that identifies time series data with similar temporal dynamics in large-scale, multi-parameter databases. In 2009, X. Chen et al. [6] explored six basic indices derived from ABP data near the forecast window including mean ABP and diastolic ABP. M. A. Mneimneh [7] proposed a rule-based approach which used the mean arterial blood pressure signal as an indicator for predicting AHE. PA Fournier [8] used information divergence between two distributions to identify the most discriminative features, and then used these features to train the classification model based on a nearest neighbor's algorithm. In 2011, T. Rocha [1] proposed the application of neural network multi-models to predict the AHE under two phases. First phase mainly trained the models based on the analysis between the current blood pressure time signal and a collection of historical blood pressure templates, and then in the second phase, the multi-model structure was employed to detect the occurrence of AHE on the basis of predicting the future evolution of current blood pressure signal. In 2013, V. Awandekar [9] proposed an algorithm consisting of probability distributions of MAP and information divergence methods to calculate the statistical distance between two probability distributions. The Bhattacharyya Distance is found to be a very accurate method to calculate such statistical distance.

A method based on random forest was proposed in this paper, because random forest classifier is one of the most

This work was supported by the National Natural Science Foundation of China (No.: 61175073), in part by Natural Science Foundation of Guangdong Province (No.: S2013010013974), in part by the Shantou University National Foundation Cultivation Project (No.: NFC13003).

Zhun Fan, is with the Guangdong Provincial Key Laboratory of Digital Signal and Image Processing and the Department of Electronic Engineering, Shantou University, Shantou, 515063 CHN (e-mail: zfan@stu.edu.cn).

Youxiang Zuo, is with the Department of Electronic Engineering, Shantou University, Shantou, 515063 CHN (e-mail: 13yxzuo@stu.edu.cn).

Dazhi Jiang is with the Department of Computer Science, Shantou University, Shantou, 515063 CHN (e-mail: dzjiang@stu.edu.cn).

Xinye Cai is with the College of Computer Science and Technology, Nanjing University of Aeronautics, Nanjing, Jiangsu, 210016 CHN (e-mail: xinye@nuaa.edu.cn).

successful ensemble learning algorithm which have been proven to be very popular and powerful techniques in the pattern recognition and machine learning community for high-dimensional classification problems[10][17]. The traditional method of growing the decision tree in the random forest is CART methodology [10]. Because of the powerful global searching ability of GP, in this paper, the decision trees in the random forest are replaced by decision functions which are obtained through GP. And as the result was shown in the experiment result part, this method has a better accuracy than the traditional method. The algorithm, namely the random forest based on genetic programming (GP-RF) mainly includes two parts - the first part is evolving the decision functions automatically using GP, and the second part is voting by the decision functions to get the final result. Features was obtained in this paper through the Intrinsic Mode Function (IMF) signal which produced by applying empirical mode decomposition (EMD) to the MAP signal. EMD is a method of processing non-stationary signal, which can obtain the local features via stepwise decomposition of the original signal. After the features are obtained, the data set and the feature set are grouped, and then a decision function will be evolved by each group via GP. A random forest is formed by the evolved decision functions from all groups, and the classification result is obtained by voting finally.

II. DATA PROCESSING

Before the algorithm is described, the dataset used in this paper, and the method on how to get the feature set are first explained as follows.

A. Dataset

Data used in this paper was collected from PhysioNet by the MIMIC-II project (Multi-parameter Intelligent Monitoring for Intensive Care),a Bioengineering Research Partnership project funded by the US National Institutes of Health and the National Institute of Biomedical Imaging and Bioengineering, with additional support from Philips Medical Systems. The MIMIC-II project has collected data from about 30,000 ICU patients up to now [1] [11]. An integrated user-friendly relational database was developed for browsing patients' clinical information (e.g. lab results, fluid balance, medications, nurses' progress notes, etc). Based upon its unprecedented size and scope, MIMIC-II has proven to be an important resource for the research of intelligent patient monitoring, supporting efforts in medical data mining and knowledge-discovery [12]. In this paper, we used 1599 pieces of data records which include 799 AHE records and 800 No_AHE records from the MIMIC-II database.

The signal we used from the database is ABP signal, which is sampled at 125Hz. We focus on the mean arterial pressure (MAP) that is actually a combination of systolic arterial blood pressure (SABP) and diastolic arterial blood pressure (DABP), which is most often calculated as follows:

$$MAP = DABP + \frac{SABP - DABP}{3}$$

In this paper, we predict occurrence of AHE within half an hour only based on the MAP signal.

B. Empirical Mode Decomposition (EMD)

The features used in this paper are obtained from the IMF produced by EMD. EMD is an adaptive method for decomposing a signal into AM-FM modulated components [13]. The method was introduced by N. E. Huang in 1998 as a nonlinear and non-stationary signal processing tool and it has been used in a large variety of applications [14].

EMD decomposes a complicated signal into several components called IMFs. Each IMF signal has some specifications [13]:

1) The number of extrema (maxima and minima) is equal to the number of zero crossings or differs only by one.

2) They are locally symmetric and the mean of top and bottom envelope of each IMF is zero.

As discussed in [14], the decomposition of original signal into IMFs is an iterative algorithm which stops until some conditions are satisfied. The conditions are as follow [13].

1) The residual signal becomes smaller than a predefined threshold.

2) The residual signal is a monotonic function that cannot be further decomposed into more IMFs.

The process of EMD can be summarized into the following steps [13]:

1) Assign i=1

2) Find all extrema (maxima and minima) of signal$x(n)$

3) Get the envelops of minima ($e_{min}(n)$) and maxima ($e_{max}(n)$) of the signal $x(n)$.

4) Compute the mean of minima and maxima envelopes:

$$m(n) = \frac{e_{min}(n) + e_{max}(n)}{2}$$

5) Compute the difference of the main signal and the mean signal: $h(n) = x(n) - m(n)$

6) Continue the steps 1-4 with $h(n)$ as a new signal or stop the iteration according to the stop conditions. Assign$c_i(n) = h(n)$, then continue the process with the residual signal $x(n) - h(n)$ as a new $x(n)$ and increase i by one.

After the process of EMD, the original signal can be written as the sum of its IMFs as follow:

$$x(n) = \sum_{i=1}^{N} c_i(n)$$

For example, Figure 1 is the MAP signal and its IMFs, the MAP signal is shown in the first sub-figure of Figure 1.

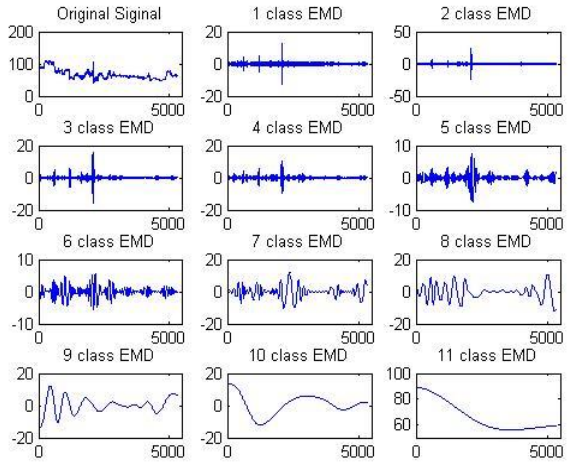

Figure 1. The original MAP signal and the decomposed IMFs

C. Feature Extraction

After the IMFs of each record are obtained by EMD, some statistical features from the original signal and the IMFs are extracted, including minimum, mean, max, median, and variance of the original signal and the IMFs.

In addition, some other features are calculated, including:

1) The maximum instantaneous frequency of each IMFs. The maximum instantaneous frequency is calculated as follows:

$$\max\left(\frac{d}{dt}(\angle hilbert(x(t)))\right)$$

This feature means the most instantaneous variations of the signal's IMFs, in which *hilbert* is a well know mathematical transformation that results in a complex signal, and the $\angle$ means the phases of this complex signal [13].

2) High frequency energy to low frequency energy ratio. After the MAP signal was decomposed into IMFs, the MAP signal can be written as:

$$MAP = \sum_{i=1}^{N} c_i(n)$$

Where $c_i(n)$ is the i^{th} IMF of the MAP. Now we can define the energy of high frequency E_1 and the energy of low frequency E_2 as follow,

$$E_1 = \sum_{i=1}^{\lfloor N/2 \rfloor} \varepsilon\{c_i(n)\}$$

$$E_2 = \sum_{i=\lfloor N/2 \rfloor+1}^{N} \varepsilon\{c_i(n)\}$$

Where $\varepsilon\{c_i(n)\}$ represents the energy of $c_i(n)$, simply the sum of squared values of $c_i(n)$:

$$\varepsilon\{c_i(n)\} = \sum_{t=-\infty}^{\infty} c_i^2(n)$$

The definition of E_1 and E_2 is reasonable because the frequency of each IMF gradually approaches zero as the EMD process proceeds [13].As is shown in figure 1, the first several IMFs have higher frequency relatively than the last several IMFs. And the frequency of the last IMF is nearly zero.

Also, we extracted the 12th percentile, skewness, kurtosis and mode of the last IMF as the members of the feature set obtained from IMFs.

III. METHODS

Before the algorithm GP-RF that combines GP and RF is proposed, the basic ideas of GP and RF are briefly introduced as follows.

A. Genetic Programming

GP is basically a Genetic Algorithm (GA) applied to a population of computer programs [15][16]. The main difference between GA and GP is the representation of the individual. The individual in GA is coded with bit strings, while the representation of GP's individual is tree structured programs, with an example shown in Figure 2. The operators in the individual are those internal nodes which are called functions such as Q (means square root), +, - and *, etc, as shown in Figure 2, while the operands are the leave nodes called terminals such as a, b, c, d, etc. Compared with GA, GP allows evolving of much more complicated structures and can therefore be applied to a greater diversity of problems [16]. J. Koza has described GP in his book "Genetic Programming, on the programming of computers by means of natural selection" (1992) [15].

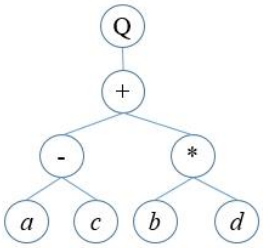

Figure 2. An example of GP program tree

The example of GP program tree in Figure 2 can be translated into mathematical formula as follows.

$$\sqrt{(a-c)+b*d}$$

A basic flow chart of GP is shown in Figure 3, which is almost the same as that of GA.

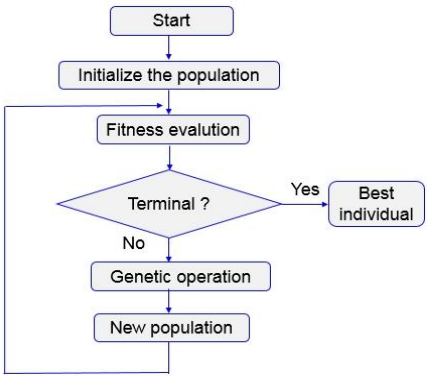

Figure 3. The flow chart of GP

First, a population should be initialized. Then, a fitness evaluation to every individual is conducted. In our case, the fitness is calculated by the accuracy of the classification. After the fitness evaluation, the best individual which has the highest fitness in the population is selected as the result of the algorithm if the evolution is finished. Otherwise, the genetic operat ions are applied on the population to generate a new population for the next generation. Generally, the genetic operations include crossover, mutation and replication. The crossover operation is done through exchanging the subtrees of a pair of selected individuals at random positions to generate two new offsprings (as shown in Figure 4). Mutation replaces one subtree with a randomly generated tree structure to result in a new individual (as shown in Figure 5). Replication selects individuals to be cloned into the new generation.

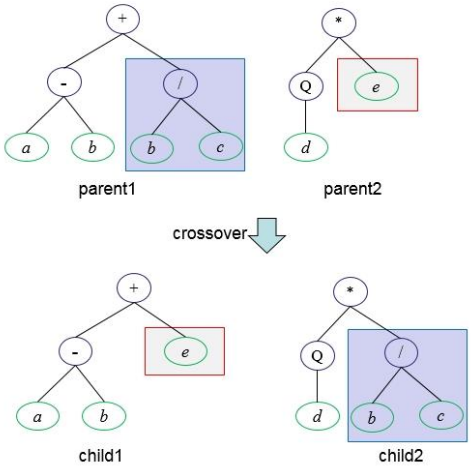

Figure 4. The crossover operation

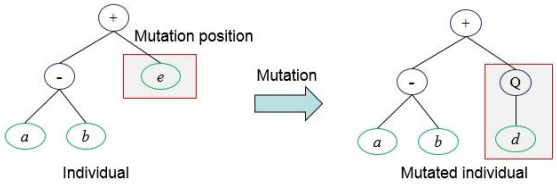

Figure 5. The mutation operation

Due to its powerful and effective global search ability for tree structures, GP is used to evolve the decision functions in this paper.

B. Random Forest

Random forest classifier is one of the most successful ensemble learning algorithms which has been proven to be a very popular and powerful technique in the pattern recognition and machine learning community for high-dimensional classification problems[10][17]. A decision forest is an ensemble of decision trees, which can be seen as one classifier containing several decision trees with various parameters [17].

The process of the algorithm includes three main parts: grouping the dataset and feature set, training the decision tree, and voting for the final result.

In the grouping part, the bootstrap idea proposed by L. Breiman [18] is adopted. The data records of the patient were grouped into t groups by re-sampling as shown in Figure6.In this figure, the input is the data set and output is the subset $\boldsymbol{D_i}$.The length of the subset $\boldsymbol{D_i}$ after grouping is the same as the original dataset. It is noteworthy to point out that some data in the subset can be the same due to the re-sampling process.

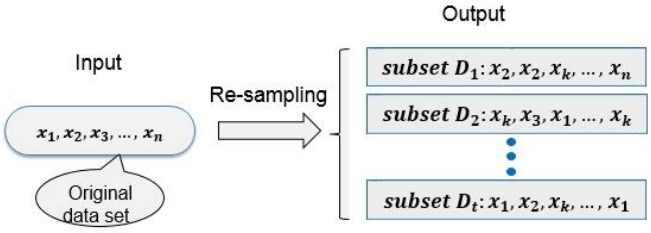

Figure 6. Grouping the dataset

The feature set is also categorized it into t groups as shown in Figure7. In this figure, the input is original feature set, and the subset of the original feature is obtained after sampling. One significant difference with grouping the data set is that the element in the subset of the feature cannot be the same, and the length of the subset can be smaller than the size of original feature set.

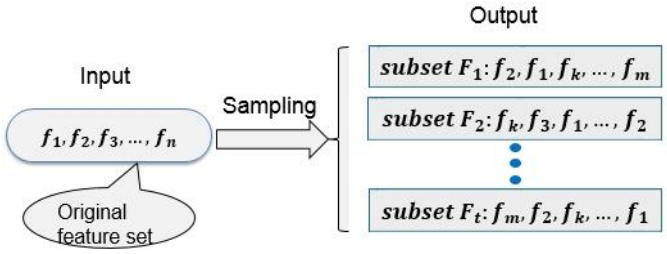

Figure 7. Grouping the feature set

After the dataset and the feature set are grouped, the subsets are used to train the decision tree. First, the subsets are combined as below:

$$G = \{G_1, G_2, \dots, G_i, \dots, G_t\},\ where\ G_i = \{D_i, F_i\}$$

The traditional method of using G_i to train the decision tree is CART algorithm [10]. For the powerful global searching ability of GP, and as is shown in the experiment part, the new method has a better result than the traditional method. So the decision tree here is obtained by GP and will be described soon.

After the random forest model is formed, for the classification problem, when a new data record comes, each decision tree will return a label. According to the majority voting method, the final class result of the new data record is determined by the majority. The flow chart of random forest algorithm is shown in Figure 8.

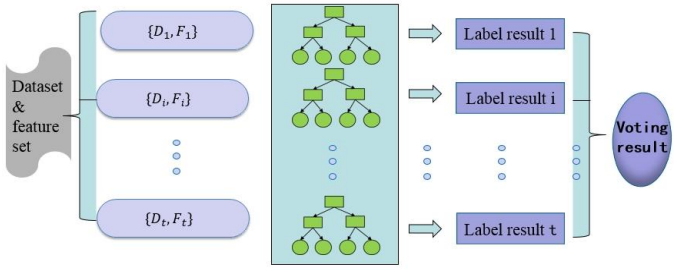

Figure 8. The flow chart of random forest

C. GP-RF

As previously discussed, the method of getting decision functions is using GP to evolve. In the phase of evolving decision functions, an individual of GP should be first defined, an example of which is shown in Figure 9.

In this individual, *+, sin, *, /* is the function symbols from function set, and *f2, f5, f3, f7* is the feature symbols from terminal set, where *sin* means sinusoidal function, +, *, / means the basic mathematical operator, and *f2, f5, f3, f7* means the feature 2, feature 5, feature 3 and feature 7, respectively. For a data record in this paper, each feature is represented by a value.

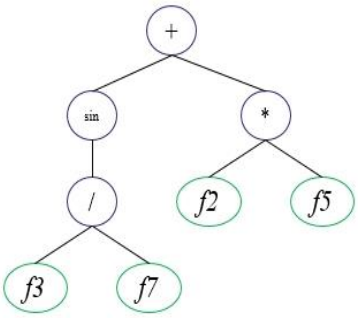

Figure 9. An example of an individual of GP used in this paper to present a decision function

Accordingly, Figure 7 can be expressed into the following formula (3).

$$y = \sin\left(\frac{f3}{f7}\right) + f2 * f5 \qquad (3)$$

Now, when a new data record comes, we can get a y value according to function (3). Then a threshold value can be set by the user to help a decision-making. For example, if we set the threshold value to zero and let the condition y>=0 indicate a normal record without AHE occurrence (corresponding to a label of NO_AHE), and y<0 an abnormal record with AHE onset (corresponding to a label of AHE), we actually get a classifier that can give each data a label of either AHE or NO_AHE. Then the fitness of an individual can be defined as the accuracy of the resulting classifier obtained by calculating the output labels of all the data and comparing them to the 'ground-truth' labels.

Up to now, we have explained the building and decoding of a decision function. We can then evolve the decision functions using GP following the flowchart of Figure 3. First, a population is initialized. After population initialization, fitness evaluation for each individual is carried out. The fitness of the individual is defined as the accuracy of the classification result. For every individual, a classification label can be obtained when a data record comes. When all the training data records are collected, we can obtain the accuracy of the individual by taking into account of all the classification result. After fitness evaluation, GP conducts other genetic operations and reproduce a new population. The elitism selection strategy is also used to keep the best individual in each generation.

The pseudo-code of the GP-RF is shown in Table 1.

GP-RF Algorithm
Input :
-t: the number of trees,
-D: the training data set,
-F: the feature set,
-m: the size of subset of the feature
-g: the generation of GP's evolution
Output:
-A random forest RF
Method:
1. ***For*** *i=1* ***to*** *t* ***do***
2. *Grouping the data set as in Figure 6.*
3. *Grouping the feature set as in Figure 7.*
4. *Tree(i) =* ***evolveTrees****(D_i,F_i,g);*
5. ***End for***
6. *RF is obtained with the set of trees with the number of t*
Function ***evolveTrees()***
1. *Initialize the population*
2. ***For*** *i=1* ***to*** *g* ***do***
3. *Evaluate the fitness of individuals in the population*
4. *Keep the best individual through elitism strategy*
5. *Apply genetic operations*
6. ***End for***
7. *Obtain a tree from the best individual of all the generations*

Table 1. Pseudo-code of the GP-RF

In the GP-RF algorithm, the input is the number of trees (decision functions) in the forest, the training data, the feature set , and the generation of the evolution in GP, the output is a RF forecasting model. Among the input, the number of trees in the forest, and the feature set are both predefined by the users. In step 2 of grouping the dataset as in Figure 6,we obtain the D_i through bootstrap method, whereas the D_i has the same length of the training data. In step 3 of grouping the feature set as in Figure 7, we obtain the F_i from the feature set, whereas the size of the F_i is *m*,which can be smaller than the size of the feature set. In step 4 a tree is evolved for each dataset and feature set combined using GP. After *t* trees have been obtained, a final RF model can be obtained in step 6.

When trees are evolved, a population is first initialized as in step 1 of Function evolveTrees(). Then a loop is conducted to evaluate the fitness of all individuals, and apply genetic operations to produce a new population. The Elitism strategy is also adopted to obtain the best individual in each generation. Finally, the best individual (an evolved decision tree) from all generations is selected as the output of the algorithm.

IV. EXPREIMENT RESULTS

A. Parameter Setting

The database used is MIMIC-II, and the size of the data records in this paper is 1599, where 799 pieces of records are AHE patient, and the other 800 pieces are normal records.

In order to get reliable estimates for classification accuracy on each classification, every experiment has been performed using 5-fold cross-validation. The number of decision trees in the random forest is 5. The size of the feature set extracted from IMFs produced by EMD is 77. In the grouping stage, the size of subset (F_i) of feature, *m* is set to 30.

B. Experiment Results

The obtained results can be described in terms of accuracy (AC), sensitivity (SE) and specificity (SP). AC, SE, SP are given by the following equations.

$$AC = \frac{TP + TN}{TP + TN + FP + FN}$$

$$SE = \frac{TP}{TP + FN}$$

$$SP = \frac{TN}{TN + FP}$$

where TP, FP, TN, and FN were defined as true positive, false positive, true negative and false negative events detected respectively.

The results of AC, SE and SP in the training set when 5-fold cross validation is used are shown in Table 2.

Training set						
Time / Precision	First validation	Second validation	Third validation	Forth validation	Fifth validation	**Average validation**
ACC	0.7852	0.7648	0.8016	0.7727	0.7945	**0.7837**
SE	0.8225	0.8146	0.8029	0.7910	0.8094	**0.8081**
SP	0.7553	0.7292	0.8003	0.7575	0.7798	**0.7644**

Table 2. The result of training set

The results of AC, SE and SP in the testing set when 5-fold cross validation is used are shown in Table 3.

Testing set						
Time / Precision	First validation	Second validation	Third validation	Forth validation	Fifth validation	**Average validation**
ACC	0.7524	0.8245	0.7524	0.7962	0.7524	**0.7755**
SE	0.7635	0.8786	0.7607	0.8786	0.7462	**0.8055**
SP	0.7427	0.7821	0.7436	0.7318	0.7566	**0.7514**

Table 3. The validation on testing set

We can see from the above result that the algorithm can obtain the ACC 78.37%, SE 80.81% and SP 76.44% in the training phase, and the ACC 77.55%, SE 80.55% and SP 75.14% in the validation phase.

The results are also compared with those using only GP method to train a decision function without random forest, and the traditional method of RF which the decision tree is obtained by CART method. After 5-fold cross validation in the training set, the comparison results are shown in Table 4.

Training set			
Algorithm / Precision	GP	GP-RF	RF
ACC	0.7803	**0.7837**	0.6938
SE	**0.8089**	0.8081	0.6870
SP	0.7575	**0.7644**	0.7009

Table 4. The comparison on training set

In the training phase, we can see that the GP-RF method is better than GP algorithm in ACC and SP indexes, but GP-RF outperforms GP in SE index. Compared with the RF method, GP-RF is better in the entire three indexes.

But in the validation phase, the results of GP-RF in ACC, SE and SP are all better than GP and RF, according to comparison results shown in Table 5.

Testing set			
Algorithm / Precision	GP	GP-RF	RF
ACC	0.7705	**0.7755**	0.7289
SE	0.7975	**0.8055**	0.7297
SP	0.7478	**0.7514**	0.7284

Table 5. The comparison on testing set

From the results, we can conclude that the GP-RF method proposed in this paper can obtain a performance of ACC 77.55%, SE 80.55% and SP 75.14% in the testing set, and outperforms both GP and RF methods.

For a better comparison between the three methods, T-test has been done. The T-test result is shown in Table 6.

p-value of T-test result		
Algorithm / Precision	GP-RF &GP	GP-RF &RF
ACC	0.3864	**0.02251**
SE	0.4175	**0.03923**
SP	0.3876	0.1211

Table 6. The T-test result

From the table 6, we can conclude that the method GP-RF is significantly better than RF method in ACC and SE for which the p-value is less than 0.05. It is also slightly better than GP in ACC and SE for which the p-value is more than 0.05. More tests need to be done and data to be collected to validate the performance of GP-RF in a more comprehensive base.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented a novel algorithm called GP-RF which combines the random forest and GP. RF is adopted as a classification model due to its strong predictive power, and GP is used to evolve the function trees in RF due to its strong global search capability. The proposed method got the results of ACC 77.55%, SE 80.55% and SP 75.14% in the validation phase, which outperforms both RF and GP algorithms.

It is also of great research interest to apply evolutionary algorithm to select an optimal set of features for classification, which will be investigated in the next step of our research.

REFERENCES

[1] Teresa Rocha, Simao Paredes, "Prediction of acute hypotensive episodes by means of neural network multi-models",*Computers in Biology and Medicine* 41 (2011) 881 - 890.

[2] J.Bassale, "Hypotension Prediction – Arterical Blood Pressure Variability", Technical Report, 2001.

[3] C.Crespo,et al., "Precursors in the arterial blood pressure signal to episodes of acute hypotension in sepsis", Proceedings of the 16th international EURASIP Conference BIOSIGNAL, vol.16,2002, pp, 206 - 208

[4] M. Saeed, "Temporal Pattern Recognition in Multiparameter ICU Data." Doctoral dissertation, Department of Electrical Engineering and ComputerScience, MIT, Cambridge, MA, 2007

[5] L.Lehman, M.saeed, "Similarity-based searching in multiparameter time series databases", Computers in Cardiology 35(2008) 653-656

[6] X Chen, D Xu, G Zhang, R Mukkamala, "Forecasting Acute Hypotensive Episodes in Intensive Care Patients Based on a Peripheral Arterial Blood Pressure Waveform.", Computers in Cardiology, Sept. 2009, pp. 545-548.

[7] MA Mneimneh, RJ Povinelli, "A rule-based approach for the prediction of acute hypotensive episodes," Computers in Cardiology, Sept. 2009, pp. 557-560.

[8] PA Foumier, JF Roy, "Acute hypotension episode prediction using information divergence for feature selection, andnon-parametric methods for classification," Computers in Cardiology, Sept. 2009, pp. 625-628.

[9] Vaibhav Awandekar, A.N. Cheeran, "Predicting Acute Hypotensive Episode by Bhattacharyya Distance." International Journal of Engineering Research and Application, 3(2):370-372, 2013.

[10] L.Breiman, Random forests, Mach. Learn. 45 (2001) 5-32

[11] MIMICII, http://physionet.org/physiobank/database/mimicdb/.

[12] M.Saeed, C.Lieu, et al, "MIMIC II: A Massive Temporal ICU Patient Database to Support Research in Intelligent Patient Monitoring", Computers in Cardiology 2002;29.641-644

[13] Abdollah Arasteh, et al, "Application of Empirical Mode Decomposition in Prediction of Acute Hypotension Episodes", ICBME2010, 3-4 November 2010

[14] Huang Norden E,et al.,"The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", Proc. R. Soc. Lond. A,1998

[15] Koza, J.R., “Genetic Programming, On the Programming of Computers by means of Natural Selection”, MIT Press, Cambridge, MA, ISBN 0-262-11170-5
[16] S.Sette, L.Boullart, “Genetic programming: principles and applications”, Engineering Applications of Artificial Intelligence 14 (2001) 727-736
[17] Ahmad Taher Azar, et al. “A random forest classifier for lymph disease”, Computer Methods and Programs in Biomedicine 113 (2014) 465-473
[18] L. Breiman, ”Bagging predictors”, Technical Report 421, Department of Statics, University of California, Berkeley, USA, 1994