Journal of Computers Vol. 30 No. 1, 2019, pp. 197-205
doi:10.3966/199115992019023001019

Retinal Blood Vessel Segmentation Using Pareto Ensemble Prunning with Diversity

Pei-Zhong Liu[1,4], Ze-Fan Liu[1*], Yu-Zhao Zhang[1], Yong Zhang[2], Hsuan-Ming Feng[3]

[1] College of Engineering, Huaqiao University, Quanzhou, 362021, China, Quanzhou, 362021, China
pzliu@hqu.edu.cn

[2] Quanzhou medical college, Quanzhou, 362021, China

[3] Department of Computer Science and Information Engineer, NationalQuemoy University,
No. 1 University Rd, Kin-Ning Vallage Kinmen, 892 Taiwan, ROC
hmfeng@nqu.edu.tw

[4] Research Center for Applied Statistics and Big Data, Huaqiao University, No. 668, Jimei Avenue, Xiamen 361021, China

Received 30 September 2017; Revised 29 October 2017; Accepted 10 January 2018

Abstract. Ensemble learning way is effectively determined to form a group of base learners with the provided hypotheses combination, which improves performance of the traditional single classifier systems. How to improve the correct responses of base learners is one of the fundamental challenges in ensemble learning systems. Although Pareto Ensemble Pruning (PEP) can concurrently maximize the generalization performance and minimize the number of base learners in an ensemble system, it still has a question to be elucidated: diversity. Diversity is a necessary machine for high generalization capability in classifier ensemble. In this study, a novel Pareto Ensemble Pruning with Diversity (PEPD) approach is proposed on the basis of negative correlation learning and NSGA-II multi-objective genetic algorithm. There are three goals in PEPD: minimizing the error and size; maximizing the diversity. Experimental results on DRIVE data sets illustrate that the PEPD algorithm achieves the better performance than the other state-of-the-art approaches.

Keywords: diversity, ensemble learning, negative correlation learning, retinal blood vessel segmentation

1 Introduction

Ensemble learning is regarded as a learning paradigm which integrates multiple base learners to complete a specific task. It firstly obtains multiple (homogeneous or heterogeneous) base classifiers, and then a certain strategy is determined from them. To improve the generalization ability of the decision model is one of the best issues by correcting the differential between these base classifiers. Lebanon [1] pointed out that they are enough weak classifiers in theory and practice while recognition rate are slightly better than random guess to be integrated into strong classifiers for achieving the high accuracy. These base learners are selected from some existing algorithms, such as C4.5 decision tree, BP neural network algorithm and the other classified algorithms SVM, etc.

Although an ensemble often has significantly better performance than a single learner, theses computation and storage overheads are sharply increasing as the number of base learners increases. In the other consideration, variation between the base learners is difficult to guarantee improvement [2-8]. Thus, in 2002, Zhou et al. proposed the concept of the selective ensemble [2]. Theoretical analysis and experiments show that it is better to eliminate some base classifiers with the acceptable low precision and

* Corresponding Author

little effect. In practice, the ensemble system is optimized by maximizing the difference between base classifiers based on trade-off of diversity and generalization.

Most recently, many algorithms are proposed to treat the ensemble issue as a multi-objective optimization problem. And the multi-objective evolutionary algorithms have been widely used in the Neural Network ensemble, such as GASEN (Genetic Algorithm based Selective Ensemble), ERNE (evolutionary random ensemble with negative correlation learning). In [3], the Pareto Ensemble Pruning (PEP) is taken as a Pareto optimization problem. It solves a bi-objective formulation by maximizing the generalization performance and minimizing the number of base learners in an ensemble system. But it still has a question to be elucidated: diversity. Diversity is a necessary condition for high generalization capability in classifier ensemble. In order to design the diverse classifiers, the base classifiers should be as unique as possible and the decision boundaries of individuals are adequately different from those of others.

In this paper, we investigate the explicit multi-objective formulation of ensemble pruning, and propose the Pareto Ensemble Pruning with Diversity (PEPD) algorithm on the basis of negative correlation learning and NSGA-II [12] multi-objective genetic algorithm. The goal of PEPD is used to optimize three objectives for accuracy, size and diversity, and to improve the performance of the ensemble. It's widely believed that the generalization of ensemble system depends on diversity and accuracy, and the base learners in ensemble should be accurate and diverse. At the same time, reducing the number base learners can accelerate speed prediction and save the storage space.

In the remainder of this paper, we briefly mention a number of existing elitist ensemble learning algorithms in Section 2. Thereafter, in Section 3, we describe the proposed PEPD algorithm in details. Section 4 presents simulation results of PEPD and compares them with other algorithms. Finally, conclusions for this approach are drawn in Section 5.

2 Related work

The ensemble technologies have two categories: (1) take diversity into account in the beginning process of generating base learners, namely generation method; (2) obtain a group of base classifier, and then select a subset from them, namely ensemble pruning method. There are three strategies to prune base learners [11-17].

Clustering based methods: the clustering algorithm is used to cluster the output of base classifiers, and then the best base classifier is selected from each cluster for ensemble. In [5], base classifiers are grouped by K-means clustering, and they are pruned by the diversity and accuracy of them.

Sorting based methods: first, the base learners are sorted according to their accuracy and diversity, and then the front of the base classifiers are selected for ensemble. For example, the Boosting idea are used in [6] to prune the Bagging classifier and in the AdaBoost classifiers based on the Kappa-Error map [11].

Optimization based method: the intelligent optimization algorithm is used to eliminate the unnecessary base classifier, and a base classifier subset is selected according to the precision and diversity. For example, in a greedy optimization method is proposed based on discrepancy [4]. The GASEN (Genetic Algorithm based Selective Ensemble) method is proposed in [2], which determined a genetic algorithm to search for the subset of classifiers with the highest integration accuracy on the validation set. DEIVCE combines the ideas from both MPANN and NCL, addressing the ensemble problem within an evolutionary multi-objective framework [10].

3 Pareto Ensemble Pruning with Diversity

In this approach, the proposed PEPD (Pareto Ensemble Pruning with Diversity), combined their advances from both PEP and NCL algorithms. For diversity, we use the negative correlation penalty function of NCL as one of the evaluated index for the multi-objectives. Then, the NSGA-II is proposed for the evolutionary computation. Three approaching objectives are considered at the same time to optimize the performance of the ensemble are accuracy, size and diversity.

Given a data set $D=\{(x_i,y_i)\}_{i=1}^{N}$ with N instances $y_i \in \{-1,1\}$ and a set of T trained base classifiers $H=\{h_t\}_{t=1}^{T}$, let Hs be a pruned ensemble with selector vector $s \in \{0,1\}^T$, where $s_t=1$ means that the base classifier h_t is selected. Hs^i is the output of the pruned ensemble for a training sample i. Then for each iteration the three goals are：

Objective 1 – Accuracy. In this work, we use mean square error to calculate the misclassification between each instance's class label $y_i \in \{-1,1\}$ and its corresponding ensemble prediction Hs^i :

$$\arg\min_{s\in\{0,1\}^T}(Accuracy = \frac{1}{N}\sum_{i=1}^{N}(Hs^i - y_i)^2) \quad \textbf{(1)}$$

Objective 2 – Size: The size of Hs is simply counted as:

$$\arg\min_{s\in\{0,1\}^T}(|s| = \sum_{t=1}^{T} s_t) \quad \textbf{(2)}$$

Objective 3 – Diversity. From NCL, the correlation penalty function is used as the third objective to optimize the ensemble performance. For each base classifier t, the following term gives an indication of how different it is from other classifiers:

$$\arg\min_{s\in\{0,1\}^T}(Diversity = \sum_{i=1}^{N}\left(h_t^i - Hs^i\right)\left[\sum_{k\neq t}^{T}\left(h_k^{\,i} - Hs^i\right)\right]) \quad \textbf{(3)}$$

Eq. 3's purpose is to negatively correlate each base learner's error with error from the rest of the ensemble system, through minimizing the mutual information between two base learners. It has been shown that NCL, due to the use of the penalty function, has achieved empirical successes and varies applications [9-12]. Hence this penalty function in PEPD is validated to ensure diversity in ensemble.

Rather than to optimize the mixture of these three objectives, we formulate the multi-objective problem as follows:

$$\arg\min_{s\in\{0,1\}^T}(Accuracy, |s|, Diversity) \quad \textbf{(4)}$$

The NSGA-II is utilized to solve this ensemble pruning problem, which is called PEPD. The algorithm randomly selects from {0,1} to initial population P = {s} of base learners, each of which can be obtained from Boosting, Bagging or many other methods, then evolves to diversify populations to improve the generalization of individual classifier.

The whole PEPD algorithm is described as follows.

(1) Initialize the values of NSGA-II algorithm parameters and randomly select from {0,1} to initial parent population P = {s}.

(2) Use fast-non-dominated sorting of P to determine the rank (front) of each member of population P.

(3) Calculate the crowding distance and sort based on front number and crowding distance.

(4) Select two solutions from P uniformly at random. Then select one with higher front number and crowding distance.

(5) Generate a new offspring through crossover and mutation. The crossover probability and mutation probability are all set to 1/T (T is the number of base learners). Then add the offspring to the population.

(6) Combine the parent population and offspring population. The combined population is increased as a size of 2T.

(7) Sort through fast-non-dominated sorting for combined population according to the front number and crowding distance and elect T base learners from the combined population. This procedure makes sure that all fronts can make contributions to the new population. What's more, the new population obtained will maintain diversity among population members.

(8) Stop this process if the expected number of generations is reached. If it's not satisfied then return back to step4 after creating the new population from the parent population.

(9) Stop and output ensemble: $Hs = sign(\sum_{s_t==1} h_t)$

The architecture of proposed PEPD algorithm is presented in Fig. 1. The differences between PEPD and PEP are obvious: (1) NCL is employed in PEPD to maintain the diversity in ensemble, ensure the higher accuracy of the base classifiers, and ultimately to achieve the improved performance of the ensemble. (2) NSGA-II is utilized to select the optimal set of classifiers by optimize accuracy, size and diversity. The computational complexity is O (MN2) that is significantly faster than genetic algorithm.

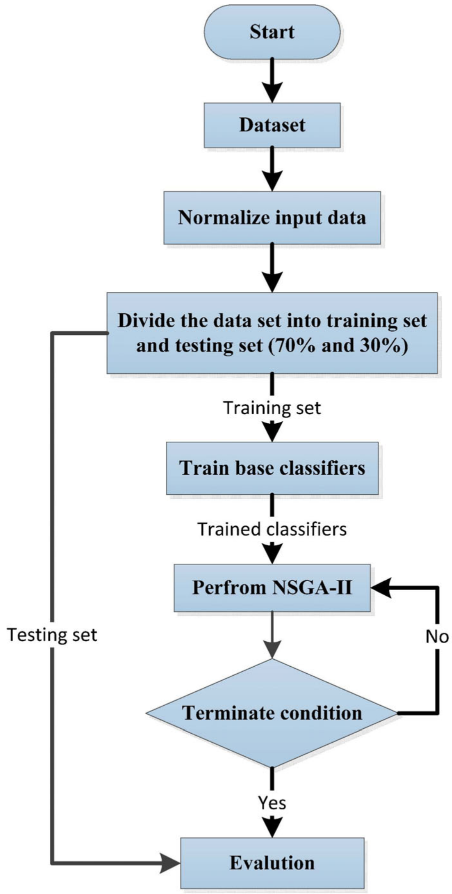

Fig. 1. Circuit for computing P'(x)

4 Simulation Result

This section proposed experiments to validate the advance of the proposed algorithm.

4.1 Data Sets Experiment

This section will evaluate the performance of PEPD on 10 real-world data sets (the dimension of features vary from 8 to 60 and the number of examples vary from 208 to 20 000) from UCI benchmark and Statlog data set [18]. They are clearly described in Table 1.

Table 1. Data sets used in this paper

Data set	#Examples	#Attributes	#Classes
Sonar	208	60	2
Ionosphere	351	34	2
Diabetes	768	8	2
Breast-w	699	9	2
Heart	270	13	2
Soybean	307	35	19
Vehicle	846	18	4
Segment	2310	19	7
Glass	214	10	7
Letter	20 000	16	26

In addition, since this paper will mainly talk about two-class classification problem, we make some modification on the multi-class data sets to generate two-class data sets: Soybean class "1-9" as one class while collapsing the remaining classes into another class; Vehicle "1-2" against "3-4"; Segment "1-3" against “4-7”; Glass “1-3” against “4-7”, Letter “1-13” against “14-26”.

The data set is uniformly and randomly divided into two parts: training set and testing set. In order to ensure the accuracy of the experiment, we use 10-fold cross validation if the number of instances is bigger than 500, or 5-fold cross validation if it’s less than 500. Each data set is normalized to zero mean and unit standard deviation. And the Bilateral Estimation of t-statistic with confidence level of 0.95 is employed here to calculate the final test error rate as follows:

$$\frac{|\bar{x}-\mu|}{\sigma/\sqrt{n}} \geq t_{0.025}(n-1) \tag{5}$$

The mean error rate and variance of n-fold cross validation are μ and σ, and $t_{0.025}(4)=2.776$, $t_{0.025}(9)=2.262$. The base classifiers used in this paper are from PRTool (http://www.prtools.org) toolbox. We choose stable learning algorithm Support Vector Machine (SVM) and unstable learning algorithm Decision Tree (DT) as base classifiers. The SVM used here is RBF kernel SVM with C = 1000, and σ is set as paper [7]. The tests were performed on MATLAB R2012b in a computer with processor i3 2.53 GHz and 4GB of RAM.

In this experiment, we have compared proposed PEPD with other state of art ensemble methods: single classifier, Bagging, AdaBoost, GASEN and PEP. A Bagging of 100 SVM or DT are used in PEP with PEPD as base classifiers, and the number of iterations is also set to 100 in Bagging, AdaBoost, and GASEN.

The results of average values of accuracy and standard deviations are shown in Table 2, when SVM as base classifiers. As we can see, except Diabetes, Heart and Glass, PEPD achieves the smallest test error. Compared to SVM, Bagging, AdaBoost, GASEN and PEP, the mean error rate have decreased 6.24%, 4.64%, 3.68%, 4.21% and 2.53%. From the experimental results in Table 2, it also can be seen that the Bagging ensemble method, which only relies on the diversity of perturbation construction of the training set, has no significant advantage compared with the single SVM classifier.

Table 2. The accuracy of SVM as base classifiers

Data set	Single	Bagging	AdaBoost	GASEN	PEP	PEPD
Sonar	76.54±1.65	76.10±3.80	77.68±2.52	76.61±3.40	78.15±3.02	81.49±2.62
Ionosphere	86.18±1.70	87.75±1.10	88.74±1.88	87.57±2.65	91.19±2.64	94.92±2.06
Diabetes	75.43±2.75	76.14±1.35	76.32±1.95	76.03±3.14	80.23±2.01	78.22±1.99
Breast-w	92.98±1.87	94.42±1.70	95.55±2.29	95.16±1.44	96.15±1.87	97.56±1.98
Heart	71.85±6.02	73.33±6.70	75.19±6.50	76.17±6.37	78.98±5.79	76.56±4.91
Soybean	86.76±3.66	91.71±2.23	92.77±2.52	91.43±2.50	92.79±2.39	94.72±2.09
Vehicle	79.69±3.19	77.41±1.59	78.19±1.08	77.85±1.33	80.36±1.46	83.40±1.67
Segment	95.89±0.91	95.45±0.91	95.93±0.78	95.89±1.18	96.09±1.05	97.41±0.91
Glass	72.52±1.08	73.07±1.40	73.61±1.60	72.63±1.57	81.21±1.56	76.52±1.08
Letter	68.33±0.49	76.86±0.91	77.82±0.37	76.43±1.18	78.03±1.47	79.84±1.36

The results of average values of accuracy and standard deviations are shown in Table 3, when DT as base classifiers. As we can see, except Diabetes, Vehicle and Letter, PEPD achieves the smallest test error. Compared to DT, Bagging, AdaBoost, GASEN and PEP, the mean error rates have decreased 6.04%, 3.32%, 2.02%, 2.96% and 1.81%.

Table 3. The accuracy of DT as base classifiers

Data set	Single	Bagging	AdaBoost	GASEN	PEP	PEPD
Sonar	67.32±8.20	73.11±6.66	75.08±5.95	74.46±6.57	75.46±7.02	76.67±6.13
Ionosphere	89.22±3.38	89.18±2.47	92.30±2.90	91.79±3.57	92.46±3.04	94.47±3.20
Diabetes	68.09±4.58	65.77±4.79	69.14±4.85	68.75±4.56	69.43±3.46	68.30±2.52
Breast-w	94.71±1.19	95.71±1.52	95.98±1.67	95.85±1.46	96.12±1.98	96.77±1.69
Heart	64.81±4.88	65.67±4.50	65.93±3.91	64.07±5.31	66.79±4.56	68.71±7.09
Soybean	55.76±1.54	69.02±2.37	70.41±2.12	65.56±2.15	71.32±2.36	76.27±1.67
Vehicle	70.25±3.89	73.63±1.90	73.01±2.22	73.28±2.86	76.42±2.34	75.88±1.73
Segment	94.32±1.37	95.88±1.24	96.04±0.98	95.94±1.32	96.25±1.35	96.73±1.06
Glass	73.05±1.44	74.66±2.18	75.59±1.62	74.67±2.30	77.98±2.34	78.20±2.38
Letter	72.72±0.47	74.86±0.87	77.05±0.53	76.71±0.36	80.37±0.79	78.72±0.67

ReferThen we analysis above results by rank-sum test, and the rank is defined by

$$R_j = \frac{1}{J}\sum_i r_i \tag{6}$$

where J is the number of experiment, and r_i is the rank of the i-th data set. The average errors of rank sum are shown in Table 4.

Table 4. Comparison of average errors of rank sum

Algorithm	Single	Bagging	AdaBoost	GASEN	PEP	PEPD
Error (SVM)	6.6	5.8	3.8	5.3	2.1	1.6
Error (DT)	6.9	5.7	3.8	5.2	2.1	1.6
Global	6.75	5.75	3.8	5.25	2.1	1.6

From Table 4, PEPD achieves the minimum average errors of rank sum. Therefore, the differences in achieving classification results between PEPD and others methods are statistically significant.

4.2 Retinal Vessel Segmentation Experiment

To further verify the effectiveness of the proposed algorithm, this part adopts the method proposed above to construct SVM ensemble classifier and applies it to retinal vascular segmentation. The retinal vascular is the only part of the human body that can be directly observed by noninvasive. It can assist in the diagnosis of diabetes, glaucoma, hypertension and other diseases by detecting the changes in the width, angle, and branches of the vessels. Retinal vascular network is the tree structure with numerous branch structure, and the contrast of small blood vessels and background is small, and the margin is ill-defined. Those make the automatic segmentation of small blood vessels more difficult. To solve the segmentation problem of small blood vessels, we firstly use multi-scale line detectors to extract image features and construct feature vectors. Then the manually labeled blood vessel pixels are records as positive samples, and the nonvascular pixels as negative samples, and a series of based SVM classifiers are trained. Finally, the PEPD method is applied to build an ensemble classifier and test its performance.

To reduce the detection error rate of edge pixels and improve the detection accuracy of the small blood vessels, we need to preprocess the fundus images. As the contrast of green image and background is the largest, so the blood vessel is extracted from the image with the green channel inversion.

The linear detector [19] defines a 15 × 15 window as a mask to calculate the average gray value I^W_{avg} corresponding to any pixel P in the retina image, and then calculate the average value of the pixels I^W_l through the detection line which makes the pixel P as the center to along the length of L (L = 15). I^W_l of 12 different directions are calculated, 15° for the step. When the direction of the test line is the same as that of the vessel, there is a maximum value of I^W_l, expressed in $I^W_{\max}$. Thereby we obtain the response function of the blood vessel: $R_W = I^W_{\max} - I^W_{avg}$. $I^W_{\max}$ maximum does not require maximum pixel gray value at the center of the detected line, which is why the basic linear detector can handle the central reflex of the vessel. Due to the single size and the fixed size of the detection line, the following two problems can easily arise: (1) fusion of two adjacent vessels into a single vessel. (2) dilatation occurs at the intersection of the two vessels.

To solve the above problems, Nguyen et al. [20] change L from 15 to 1, 3, 5, 7, 9, 11, 13, 15 and so on, and take their weighted average as the final test result, as shown in equation (7).

$$R_c = \frac{1}{n_L+1}\left(\sum_L R^L_W + I_{gc}\right) \tag{7}$$

where n_L is the number of all scales, I_{gc} is the pixel value corresponding to the original green channel image. Then, the equation (2) is used to fuse the response functions at different scales, and the final results are obtained. This method solves the expansion phenomenon that occurs when the blood vessel is merged into a single vessel and the blood vessel, but the method takes the same weight of the test result of the scale and does not take the different results of the different scale processing into account, resulting in the loss of part of the image or excessive use, reducing the accuracy of vascular segmentation.

Aim at this problem, the SVM ensemble classifier is introduced to distinguish the vascular and non vessel pixels. The feature vector $\mathbf{x} = [R_W^1, R_W^2, \cdots, R_W^L, I_{gc}]$ is constructed using the detection results of different scales line and detectors, L = 15. Then, a series of SVM classifiers with different classification accuracy and diversity are trained by ultilizing different parameters. Finally, the decision classifier is obtained by the PEPD algorithm. The specific processing flow is shown in Fig. 2.

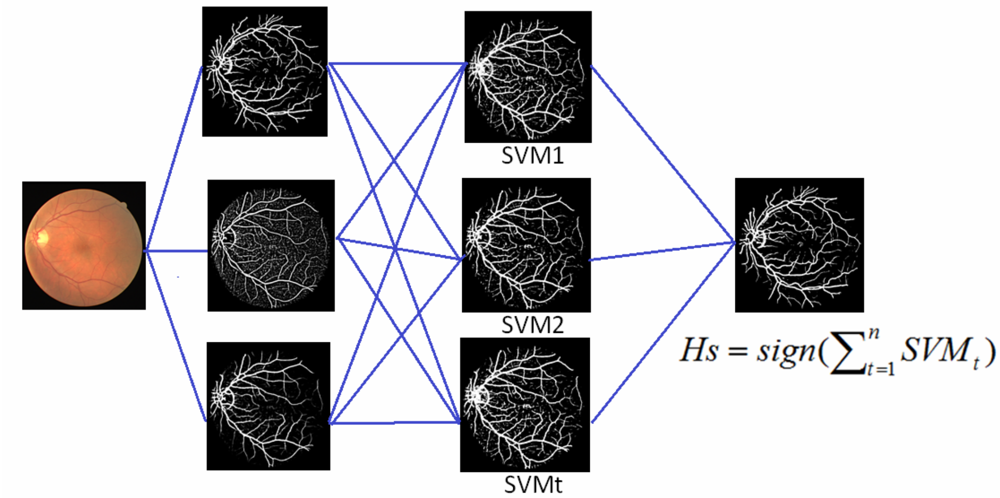

Fig. 2. The processing flow of the proposed algorithm

In this paper, an open DRIVE dataset is used to perform the simulation test, which consists of 40 color fundus images of 768×584 sizes, each of which consists of 20 training images and a test image. And the evaluation criteria of Sensitivity(Se) 、Specificity(Sp) 、Precision(Pre) F1-scole(F1) and accuracy(ACC) are used as main measures for evaluation and comparison. When the classifier is trained, 1000 samples are randomly selected from 20 feature graphs where 20 * 1000 samples are selected. And SVM adopts the radial basis function, The parameters setting as follows: $C \in [1,10]$. A total of 100 base SVM classifiers are trained.

The proposed algorithm is compared with the single-scale linear detection and multi-scale linear detection as shown in Fig. 3, performance indices of comparison results are shown in Table 5. It can be seen that the proposed algorithm has a good segmentation to the blood vessel and non vessel pixels, and the results contain more small blood vessels, and the whole blood vessel has good connectivity and high segmentation accuracy.

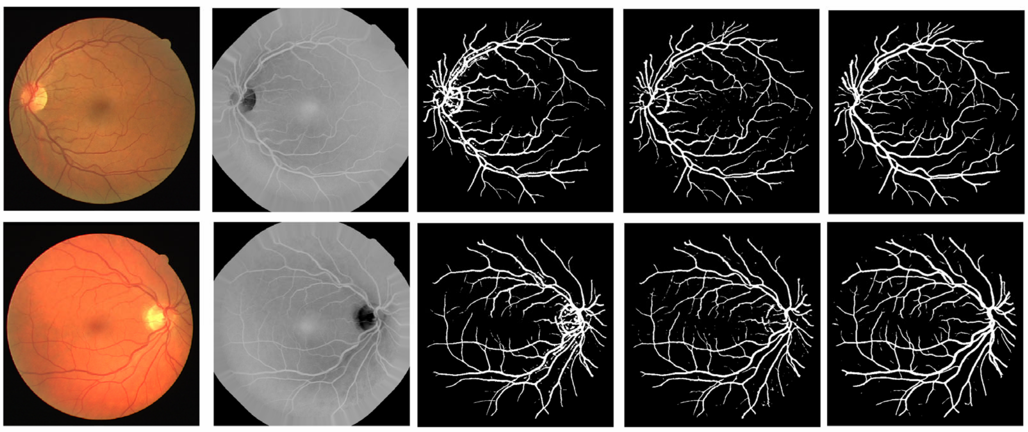

Fig. 3. The segmentation results of different algorithms from left to right: the raw fundus image, the preprocessing image, [19] segmentation results, [20] segmentation results and our segmentation result

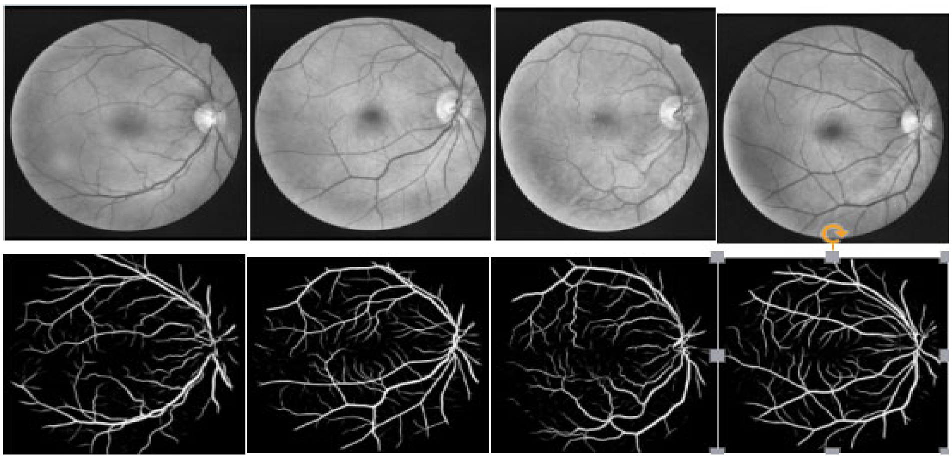

Fig. 4. Experience results with another fundus images

Table 5. Performance indexes comparison of different algorithms

	Literature [19]	Literature [20]	Our algorithm
Se	0.7273	0.7425	0.8673
Sp	0.9589	0.9701	0.9671
Pre	0.7196	0.7832	0.8230
F1	0.7217	0.7602	0.8446
ACC	0.9290	0.9407	0.9521

5 Conclusion

This paper proposes an efficient ensemble pruning algorithm that minimizes the error and size and maximizes the diversity at the same time. The proposed algorithm includes two topics. One is called the negative correlation learning, which is employed to calculate diversity. And the other is generated as NSGA-II multi-objective genetic algorithm to achieve the multi-objective optimization.

Several experiments are carried out in this paper to evaluate that PEPD performs on different training dataset in comparison with other ensemble algorithms. PEPD shows an excellent performance in the solution of these data sets. This study also shows that the selective ensemble also works well even the redundant or useless information between base classifiers are appeared. And PEPD not only keeps the benefits of PEP but also improves its performance through the optimizing diversity function.

Refrences

[1] G. Lebanon, J. Lafferty, Boosting and maximum likelihood for exponential models, in: Proc. Advances in Neural Information Processing Systems, 2001.

[2] Z.H. Zhou, J. Wu, W. Tang, Ensembling neural networks: many could be better than all, Artificial Intelligence 137(1-2)(2002) 239-263.

[3] C. Qian, Y. Yu, Z.H. Zhou, Pareto ensemble pruning, in: Proc. AAAI Conference on Artificial Intelligence, 2015.

[4] S.-S. Ma, L.-C. Jiao, L. Xiong, S. Gou, Greedy optimization classifiers ensemble based diversity, Pattern Recognition 44(6)(2011) 1245-1261.

[5] A. Lazarevic, Z. Obradovic, Effective pruning of neural network classifier ensembles 2(2)(2001) 796-801.

[6] G. Martinez-Muñoz, A. Suárez, Using boosting to prune bagging ensembles, Pattern Recognition Letters 28(1)(2007) 156-165.

[7] L. Li, B. Zou, Q. Hu, Dynamic classifier ensemble using classification confidence, Neurocomputing 99(99)(2013) 581-591.

[8] Y. Freund, R.E. Schapire, Experiments with a new boosting algorithm, in: Proc. the 13th. Int. Conf. on Machine Learning, 1996.

[9] Y. Yu, X. Yao, Z.-H. Zhou, On the approximation ability of evolutionary optimization with application to minimum set cover, Artificial Intelligence 180-181(2012) 20-33.

[10] A. Chandra, X. Yao, DIVACE: diverse and accurate ensemble learning algorithm, in: Z.R. Yang, H. Yin, R.M. Everson (Eds.), Intelligent Data Engineering and Automated Learning-IDEAL, 2004, Springer, Switzerland, 2004, pp. 619-625.

[11] D.D. Margineantu, T.G. Dietterich, Pruning adaptive boosting, in: Proc. Fourteenth International Conference on Machine Learning, 1997.

[12] K. Deb, A fast elitist multi-objective genetic algorithm: NSGA-II, IEEE Transactions on Evolutionary Computation 6(2)(2002) 182-197.

[13] X. Li, L. Wang, E. Sung, AdaBoost with SVM-based component classifiers, Engineering Applications of Artificial Intelligence 21(5)(2008) 785-795.

[14] G. Valentini, T.G. Dietterich, Bias-variance analysis of support vector machines for the development of SVM-based ensemble methods, Journal of Machine Learning Research 5(2004) 725-775.

[15] N. Ueda, R. Nakano, Generalization error of ensemble estimators, in: Proc. IEEE International Conference on Neural Networks, 1996.

[16] Z.S. Chan, N. Kasabov, A preliminary study on negative correlation learning via correlation-corrected data (NCCD), Neural Processing Letters 21(3)(2005) 207-214.

[17] L.I. Kuncheva, C.J. Whitaker, Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy, Machine Learning 51(2)(2003) 181-207.

[18] A. Asuncion, D. Newman, UCI machine learning repository. <http://www.ics.uci.edu/~mlearn/MLRepository.html>, 2007.

[19] E. Ricci, R. Perfetti, Retinal blood vessel segmentation using line operators and support vector classification, IEEE Transactions on Medical Imaging 26(10)(2007) 1357.

[20] U.T.V. Nguyen, A. Bhuiyan, L.A.F. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, Pattern Recognition 46(3)(2013) 703-715.