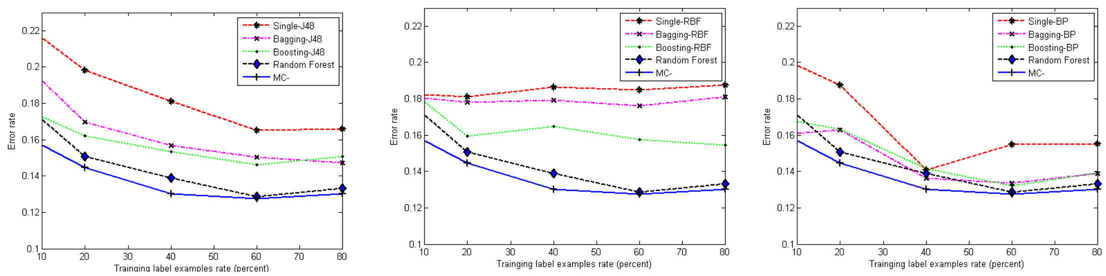
B. Ensemble classification experiment

The outputs of multiple single classifiers trained respectively are fused by combination methods. This ensemble classification will compensate the shortcoming of single classifiers, and boost the classification accuracy. The results of Bagging, Boosting, three multiple classifiers (base classifiers are J48, RBF and BP, named as MCS) and Random forest (combined with 10 decision trees) are illustrated in TABLE III.

From the TABLE III, the performance of ensemble classification such as bagging and boosting are better than that of single classifier. Comparing with the figure of TABLE II, as the training samples reduce down to about 20%, the classification accuracy still can reach 100% training sample of single classifier. It shows that ensemble methods can achieve classification performance with small training examples (compared with the single supervised classification). Fig.5 shows the performance of four approaches of ensemble comparing with the single classifier. With the improvement of training examples rate, in most cases the error rate is on the decline. The MCS and RF are stable and make better results than others. Boosting is better than Bagging in the J48 and RBF learn algorithm.



(a) J48 (b) RBF (c) BP

Fig 5 Performance of different classifiers

V. Conclusion

Ensemble methods as the way to obtaining highly accurate classifiers by combining less accurate ones, are learning algorithms that construct a set of classifiers and then classify new data by taking a vote of their predictions. According to the experimental results from the single classifier and ensemble methods including Bagging, AdaBoost, MCS as well as Random forest on environmental audio, the ensemble methods are able to outperform any single classifier. Even with the small training examples, the performance of classification and generalization can be guaranteed. It is difficult to obtain the labeled data as training examples for environmental audio. Ensemble methods provide an effective way to perform the classification. Among the ensemble methods, the Random forest performs so well in environmental audio classification and obtains strong generalization.

Further research work about deriving effective diversity controls for ensemble learning, and combining ensemble learning with and semi-supervised learning to build the better learning model are underway.

Acknowledgments

This work was supported by the Research Fund Projects of Yunnan Provincial Department of Education under the Grant No. 2012C098.

References

[1] Xiao-mei Ding-cai Yang. Environmental audio classification based on support vector machine. Electronic Measurement Technology. 2008,31(9):121-123
[2] Yi-wen Zheng. Typical Methods for Audio Classification. JiSuanJi Yu Xian Dai Hua.2007,(8):59-63
[3] Dietterich T G. Machine learning research: Four current directions. AI Magazine, 1997,18(4):97-136
[4] Zhi-Hua Zhou. Ensemble Methods-Foundations and Algorithms. CRC Press, Taylor & Francis Group.2012
[5] Cha Zhang, Yunqian Ma. Ensemble Machine learning Methods and Application. Springer.2012
[6] Breiman L. Bagging predicators. Machine Learning, 1996, 24(2):123-140
[7] Duda,R.O. pattern classification(2nd ed.). Mechanical industry press. 2003
[8] Ian H. Witten, Eibe Frank, Mark A.Hall. Data Mining: Practical machine learning tools and techniques (3rd ed.). Morgan Kaufmann,2011
[9] Breiman L. Random forests. Machine Learning,45(1):5-32,2001.
[10] Kuncheva L.I. Combining Pattern Classifications: Methods and Algorithms. John Wiely & Sons, Hoboken, NJ,2004