

Dimensionality Reduction by SVM-KNN Approach for Arecanut Classification

Ajit Danti^a and M Suresha^b

^aDepartment of Computer Applications, Jawaharlal Nehru National College of Engineering, Karnataka, India, Contact: ajitdanti@yahoo.com

^bDepartment of Computer Science, Kuvempu University, Karnataka, India, Contact: srit_suresh@yahoo.com.

In this paper, combination of Support Vector Machines(SVM) and k -Nearest Neighbour (k -NN) approach is employed for effective classification of arecanut. Combined approach reduces feature dimension considerably which inturn speed up classification of arecanut. The k -NN rule is a simple and effective method for multi-way classification that is much used in Computer Vision. However, its performance depends heavily on the distance metric being employed. k -NN classifiers suffer from the problem of high variance in the case of limited sampling. Alternatively, one could use support vector machines but they involve time-consuming optimization and computation of pairwise distances. We propose a combination of these two methods which deals with the multiclass problem, has reasonable computational complexity in classification and gives excellent results in practice. The basic idea is to find support vectors using Support Vector Machine Classifier. k -NN classifier uses only support vectors as a feature space which is given by SVMs in training phase. k is the most important parameter in the arecanut grading system based on k -NN. This method can be applied to large, multiclass data sets. Proposed method is experimented with large data set and determined the qualitative analysis of the classification approach.

Keywords: Arecanut, Classification Approach, k -Nearest Neighbour (k -NN), Support Vector Machines (SVM)

1. INTRODUCTION

ARECANUT (Areca catechu L.) is one of the important commercial crops of India. It plays a prominent role in the religious, social and cultural functions and economic life of people in India. Its cultivation is concentrated in South Western and North Western regions of India. The economic product is the fruit called *betel nut* and is used mainly for masticatory purposes. Today, the increasing technological development and sophistication of modern societies impose new quality standards to the crop producers. Consumers demand more and more information about the products they buy, demonstrating clear preferences for well-informed high-quality products. Arecanut has innovative in Ayurvedic and veterinary medicines. The habit of chewing arecanut is typical of the Indian sub-continent

and its neighborhood. Arecanut is grown in Bangladesh, China, Malaysia, Indonesia, Vietnam, Philippines and Thailand. India accounts for about 57 percent of world production. The quality, variety and types of arecanut vary from one place to another. Recent studies have shown that arecanut has pharmacological uses such as hypoglycemic effect, mitotic activity etc. Since, these classification activities are being done manually and is mainly labour dependent and hence automation is necessary. There are several computer based technologies for other crops but there is no computer vision based advanced technology in identifying grade for the arecanut. There is an increasing demand for computer vision based technology to address the above issue for arecanut farmers. The SVM, which is based on the theory of structural risk minimization in statistical

REFERENCES

1. Vladimir Vapnik. The Nature of Statistical Learning Theory, *Springer-Verlag*, 1995.
2. B Scholkopf and A J Smola. Learning with Kernels, *Cambridge, MIT Press*, 2002.
3. Cortes and V Vapnik. Support-Vector Networks, *Machine Learning*, pages 273-297, 1995.
4. J Shawe-Taylor and N Cristianini. Kernel Methods for Pattern Analysis, *Cambridge University Press*, 2004.
5. N Cristianini and J Shawe-Taylor. An Introduction to Support Vector Machines, *Cambridge University, Cambridge, U.K. Press*, 2000.
6. A Micchelli and M Pontil. Learning the Kernel Function via Regularization, *Journal of Machine Learning Research*, 6:1099-1125, 2005.
7. Argyriou R Hauser, C A Micchelli and M Pontil. A DC Algorithm for Kernel Selection, *In Proceedings of 23rd International Conference on Machine Learning*, Pittsburgh, pages 41-49, 2006.
8. F R Bach, G R G Lanckriet and M I Jordan. Multiple Kernel Learning, Conic Duality and the SMO Algorithm, *In Proceedings of 21st International Conference on Machine Learning*, pages 6-14, 2004.
9. G R G Lanckriet, N Cristianini, P L Bartlett, L El Ghaoui and M I Jordan. Learning the Kernel Matrix with Semi Definite Programming, *Journal on Machine Learning Research*, 5:27-72, 2004.
10. I Argyriou, C A Micchelli and M Pontil. Learning Convex Combinations of Continuously Parameterized basic Kernels, *In Proceedings of 18th Annual Conference on Learning Theory*, Bertinoro, Italy, pages 338-352, 2005.
11. I W Tsang and J T Kwok. Efficient Hyperkernel Learning using Second-Order Cone Programming, *IEEE Transactions on Neural Networks*, 17:48-58, 2006.
12. S Sonnenburg, G Ratsch, C Schafer and B Schalkopf. Journal on Machine Learning Research. *PPH*, 7:1531-1565, 2006.
13. Y Lin and H H Zhang. Component selection and smoothing in smoothing spline analysis of variance models, *Technical Report, University of Wisconsin - Madison and North Carolina State University*, 2003.
14. Lixin Duan, Tsang I W, Dong Xu. Domain Transfer Multiple Kernel Learning, *IEEE Transactions on PAMI*, 34(3):465-479, 2012.
15. Vedaldi A, Zisserman A. Efficient Additive Kernels via Explicit Feature Maps, *IEEE Transactions on PAMI*, 34(3):480-492, 2012.
16. Joachims T. Text Categorization with Support Vector Machines: Learning with Many Relevant Features, *In Proceedings of the European Conference on Machine Learning*, 1998.
17. Li Baoli, Chen Yuzhong, and Yu Shiwen. A Comparative Study on Automatic Categorization Methods for Chinese Search Engine, *In Proceedings of the Eighth Joint International Computer Conference*, Zhejiang University Press, Hangzhou, pages 117-120, 2002.
18. D S Guru, Y H Sharath, S Manjunath. Texture Features and K-NN in Classification of Flower Images, *IJCA on Recent Trends in Image Processing and Pattern Recognition*, 2010.
19. Manning C D and Schutze H. Foundations of Statistical Natural Language Processing, *Cambridge, MIT Press*, 1999.
20. Domingo Mery and Franco Pedreschi. Segmentation of Color Food Images using a Robust Algorithm, *Journal of Food Engineering*, 66:353-360, 2005.
21. Meftah Salem M Alfatni, Abdul Rashid Mohamed Shariff, Helmi Zulhaidi Mohd Shafri, Osama M Ben saaed and Omar M Eshanta. Oil Palm Fruit Bunch Grading System Using Red, Green and Blue Digital Number, *Journal of Applied Sciences*, 8(8):1444-1452, 1999.
22. Gyeongyong Heo, Klette R, Young Woon Woo, Kwang-Baek Kim, Nam Ho Kim. Fuzzy Support Vector Machine with a Fuzzy Nearest Neighbor Classifier for Insect Footprint Classification, *IEEE International Conference on Fuzzy Systems (FUZZ)*, pages 1-6, July 2010.
23. Pallabi Parveen and Bhavani Thuraisingham. Face Recognition Using Multiple Classifiers, *IEEE International Conference on Tools with Artificial Intelligence*, pages 179-186, Nov 2006.
24. Pingpeng Yuan, Yuqin Chen, Hai Jin and Li Huang. MSVM-kNN: Combining SVM and k-NN for Multi-class Text Classification, *IEEE International Workshop on Semantic Computing and Systems*, pages 133-140, 2008.
25. Xiaoqiao Shen. Gene Expression Data Classification using SVM-KNN Classifier, *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*, pages 149 - 152, 2004.
26. Rafael C Gonzalez and Richard E Woods. Dig-

- ital Image Processing, *PPH*, 2008.
27. Christopher J C Burges. Data Mining and Knowledge Discovery, *A Tutorial on Support Vector Machines for Pattern Recognition*, pages 121-167, 1998.



Ajit Danti is currently the Director and Professor in the Department of Computer Applications, JNN College of Engineering, Shimoga, Karnataka. He obtained his Bachelor of Engineering from University Visvesvaraya College of Engineering, Bangalore. He received his Masters degree in Computer Management from Shivaji

University, Maharashtra. He was awarded Ph.D in Computer Science and Technology from Gulbarga University, Karnataka, India.

M Suresha is currently the Assistant Professor in the Department of Computer Science, Kuvempu University, Shimoga, Karnataka. He obtained his Masters degree in Computer Applications from Kuvempu University, Shimoga, Karnataka, India. He is pursuing Ph.D degree in Computer Science and Technology in Kuvempu University.