Apples Grading based on SVM Classifier

Suresha M
Dept. of Computer Science
Kuvempu University,
Shankaraghatta, Shivamoga

Shilpa N.A
Dept. of Computer Science
Kuvempu university,
Shankaraghatta, Shivamoga

Soumya B
Dept. of Computer Science
Kuvempu University,
Shankaraghatta, Shivamoga

ABSTRACT
In this paper, effective automatic grading of apples is proposed. The apples RGB images is converted into HSV image and threshold based approach is used for segmentation of apples from the background. Average red and green color components of the apples are determined for classification of apples. With the help of support vector machines (SVMs), classification is done and found accuracy of 100%.

General terms
Image Processing, Pattern Recognition.

Keywords
Apple Grading, Classification, Pattern Recognition, Support Vector Machines.

1. INTRODUCTION
Apples are among the most familiar fruits, with a long history of cultivation. The apple is the pomaceous fruit of the apple tree, species Malus domestica in the rose family Rosaceae. It is one of the most widely cultivated tree fruits. The tree is small and deciduous, reaching tall, with a broad, often densely twiggy crown. The leaves are alternately arranged simple ovals 5 to 12 cm long and broad on a petiole with an acute tip, serrated margin and a slightly downy underside. Blossoms are produced in spring simultaneously with the budding of the leaves. The flowers are white with a pink tinge that gradually fades, five petaled, and in diameter. The fruit matures in autumn, and is typically diameter. The center of the fruit contains five carpels arranged in a five-point star, each carpel containing one to three seeds. The tree originated from Central Asia, where its wild ancestor is still found today. There are more than 7,500 known cultivars of apples resulting in a range of desired characteristics. Cultivars vary in their yield and the ultimate size of the tree, even when grown on the same rootstock. Different cultivars are available for temperate and subtropical climates. One large collection of over 2100 apple cultivars is housed at the National Fruit Collection in England.

A support vector machine (SVM) is a concept in statistics and computer science for a set of related supervised learning methods that analyze data and recognize patterns, used for classification and regression analysis. The standard SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the input, making the SVM a non-probabilistic binary linear classifier. Given a set of training examples each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. More formally, a support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier. Whereas the original problem may be stated in a finite dimensional space, it often happens that the sets to discriminate are not linearly separable in that space. For this reason, it was proposed that the original finite-dimensional space be mapped into a much higher-dimensional space, presumably making the separation easier in that space. To keep the computational load reasonable, the mappings used by SVM schemes are designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function $k(x, y)$ selected to suit the problem William H [11]. The hyperplanes in the higher-dimensional space are defined as the set of points whose inner product with a vector in that space is constant. The vectors defining the hyperplanes can be chosen to be linear combinations with parameters α_i of images of feature vectors that occur in the data base. With this choice of a hyperplane, the points x in the feature space that are mapped into the hyperplane are defined by the relation:

$$\sum_i \alpha_i K(x_i, x) = \text{constant} \quad (1)$$

Note that if $K(x, y)$ becomes small as y grows farther away from x, each element in the sum measures the degree of closeness of the test point x to the corresponding data base point x_i. In this way, the sum of kernels above can be used to measure the relative nearness of each test point to the data points originating in one or the other of the sets to be discriminated. Note the fact that the set of points x mapped into any hyperplane can be quite convoluted as a result, allowing much more complex discrimination between sets which are not convex at all in the original space.

2. LITERATURE SURVEY
A computer vision based system is introduced to automatically grade apple fruits. Segmentation of defected skin is done by global thresholding techniques. Statistical features are extracted from the segmented areas and then fruit is graded by a supervised classifier. Linear discriminant, nearest neighbor, fuzzy nearest neighbor, adaboost and support machines classifiers are tested for fruit grading and found 89% recognition rate using SVM Devrim Unay et al. [4]. Conducted an exhaustive survey of image thresholding methods, categorize them, express their formulas under a uniform notation, and finally carry their performance

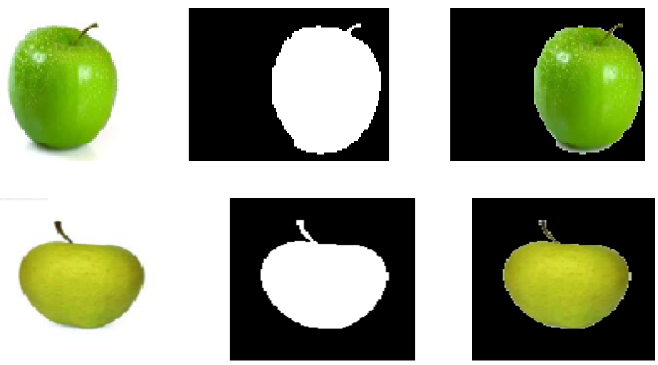
comparison. The thresholding methods are categorized according to the information they are exploiting, such as histogram shape, measurement space clustering, entropy, object attributes, spatial correlation, and local gray-level surface Mehmet Sezgin et al.[8]. Designed and implemented a prototypical computer vision based date grading and sorting system. Defined a set of external quality features. The system uses RGB images of the date fruits. From these images, it automatically extracts the aforementioned external date quality features. Based on the extracted features it classifies dates into three quality categories. A back propagation neural network classifier used to test results and the system has given 80% accuracy Yousef Al Ohali [12]. Work have been done on classification of fruits, flowers and seeds based on color Domingo Mery et al.[5], Meftah Salem M Alfatni et al.[7]. The SVM, which is based on the theory of structural risk minimization in statistical learning Vladmir Vapnik [10], has outperformed many traditional learning algorithms. It is now generally recognized as a powerful method for various machine learning problems B Scholkf et al.[2], J shave Taylor et al.[6], N Christianini et al.[9].

3. SEGMENTATION

Generally speaking, through the binarization of gray-scale images useful information for the segmentation of touched or overlapped images may be lost in many cases. If we analyze gray-scale images, however, specific topographic features and the variation of intensities can be observed in the image boundaries. In this paper, we have used threshold based segmentation to segment apple image from the background. The segmented image is converted into binary image and applied morphological operations to fill the possible holes inside the apple object as shown in the Fig. 1 (b) and 2 (b). Then, the binary image is multiplied with original image to segment the apple image as shown in Fig. 1 (c) and 2 (c).

(a) original images (b)mask images (c) segmented images

Fig 1.:sample segmentation results for red apple images



(a)original images (b) mask images(c)segmented images

Fig 2: sample segmentation results for green apple images

4. SUPPORT VECTOR MACHINES FOR CLASSIFICATION

The original SVM algorithm was invented by Vladimir N. Vapnik and the current standard incarnation (soft margin) was proposed by Vapnik et al. [10].

4.1. Motivation

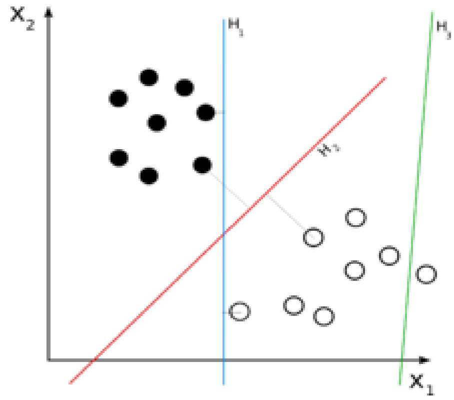


Fig 3: Linear classifier

In the Fig. 3 line H3 doesn't separate the two classes where as H1 does, with a small margin and H2 with the maximum margin. Classifying data is a common task in machine learning. Suppose some given data points each belong to one of two classes, and the goal is to decide which class a new data point will be in. In the case of support vector machines, a data point is viewed as a *p*-dimensional vector (a list of *p* numbers), and we want to know whether we can separate such points with a $(p - 1)$ dimensional hyperplane. This is called a linear classifier. There are many hyperplanes that might classify the data. One reasonable choice as the best hyperplane is the one that represents the largest separation, or margin, between the two classes. So we choose the hyperplane so that the distance from it to the nearest data point on each side is maximized. If such a hyperplane exists, it is known as the maximum-margin hyperplane and the linear classifier it defines is known as a maximum margin classifier; or equivalently, the perceptron of optimal stability.

4.2. Linear SVM

We are given some training data, *D* a set of *n* points of the form

$$D = \{ (x_i, y_i) \mid x_i \in R^p, y_i \in \{-1, 1\} \} \qquad (2)$$

where the y_i is either 1 or −1, indicating the class to which the point x_i belongs. Each x_i is a *p*-dimensional real vector. We want to find the maximum-margin hyperplane that divides the points having $y_i = 1$ from those having $y_i = -1$. Any hyperplane can be written as the set of points *x* satisfying

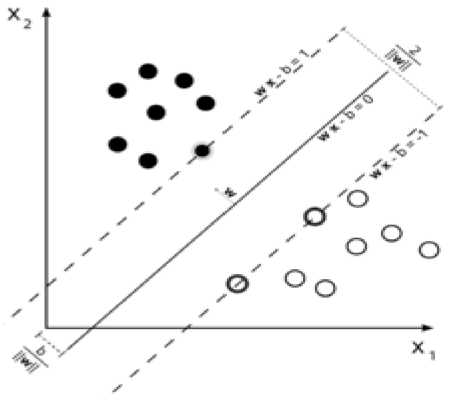


Fig 4: Linear classifier with maximize the margin

Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors.

$$w.x - b = 0 \quad (3)$$

Where $w.x$ denotes the dot product of the normal vector *w* to the hyperplane. The parameter $\frac{b}{\|w\|}$ determines the offset of the hyperplane from the origin along the normal vector *w*

We want to choose the *w* and *b* to maximize the margin, or distance between the parallel hyperplanes that are as far apart as possible while still separating the data. These hyperplanes can be described by the equations

$$w.x - b = 1 \quad \textbf{(4)}$$

and

$$w.x - b = -1 \quad \textbf{(5)}$$

Note that if the training data are linearly separable, we can select the two hyperplanes of the margin in a way that there are no points between them and then try to maximize their distance. By using geometry, we find the distance between these two hyperplanes is $\frac{2}{||w||}$, so we want to minimize $\|w\|$. As we have to prevent data points from falling into the margin, we add the following constraint: for each *i*

$$w.x_i - b \geq 1 \ \ for\, x_i \ \text{ of the first class} \quad (6)$$

or

$$w.x_i - b \leq -1 \ \ for\, x_i \ \text{ of the second class} \quad (7)$$

This can be rewritten as:

$$y_i(w.x_i - b) \geq 1 \quad (8)$$

We can put this together to get the optimization problem:

Minimize in (w,b) $||w||$

Subject to (for any $i = 1......n$)

$$y_i(w.x_i - b) \geq 1 \quad (9)$$

4.3 **.Kernel Machine**

The original optimal hyperplane algorithm proposed by Vapnik in 1963 was a linear classifier[10]. However, in 1992, Boser et al. suggested a way to create nonlinear classifiers by applying the kernel trick to maximum-margin hyperplanes Boser et al [3] (originally proposed by Aizerman et al. [1]). The resulting algorithm is formally similar, except that every dot product is replaced by a nonlinear kernel function. This allows the algorithm to fit the maximum-margin hyperplane in a transformed feature space. The transformation may be nonlinear and the transformed space high dimensional; thus though the classifier is a hyperplane in the high-dimensional feature space, it may be nonlinear in the original input space.

If the kernel used is a Gaussian radial basis function, the corresponding feature space is a Hilbert space of infinite dimensions. Maximum margin classifiers are well regularized, so the infinite dimensions do not spoil the results. Some common kernels include:

Polynomial (homogeneous):

$$k(x_i, x_j) = (x_i . x_j)d \quad (10)$$

Polynomial (inhomogeneous):

$$k(x_i, x_j) = (x_i . x_j + 1)d \quad (11)$$

Gaussian radial basis function:

$$k(x_i, x_j) = \exp(-\gamma\, ||x_i - x_j||2) \quad (12)$$

for $\gamma > 0$ Sometimes parameterized using

$$\gamma = 1/2\sigma^2$$

Hyperbolic tangent:

$$k(x_i, x_j) = \tanh(kx_i . x_j + c) \quad (13)$$

for some (not every) $k > 0$ and $c < 0$. The kernel is related to the transform $\varphi(x_i)$ by the equation $k(x_i, x_j) = \varphi(x_i)$. The value *w* is also in the transformed space, with $w = \sum \alpha_i y_i \varphi(x_i)$. Dot products with *w* for classification can again be computed by the kernel trick, i.e. $w.\varphi(x) = \sum_i \alpha_i y_i k(x_i, x)$. However, there does not in general exist a value *w'* such that $w.\varphi(x) = k(w', x)$

5. RESULTS AND DISCUSSION

The database contains 90 images. All the images were downloaded from world wide web. Images were resized into 89 X 142 pixel resolution for reasonable computation speed. The proposed method efficiently classify apples as red apple and green apple. We have used SVM classifier for classification of apples. We have considered linear kernal function for the SVM classifier and found 100% accuracy and average red color against green color of the apples are plotted as show in the Fig. 5.

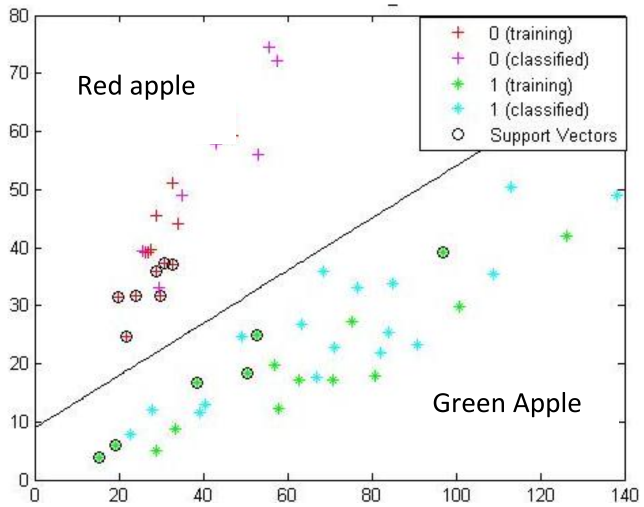


Fig 5: SVM classifier

6. CONCLUSION

In this paper, we have used threshold based segmentation to segment the apple image form the background. In the segmented region, we have used only red and green components to classify the apples. The classification of apples is done using SVM classifier. This method can be extended to other objects such as classification of seeds, flowers etc.. where there is an human intervention is existed for classification.

7. REFERENCES

[1] Aizerman. Mark, A. Braverman. Emmanuel, M. Rozonoer. Lev, I. 1964. Theoretical foundations of the potential function method in pattern recognition learning. Automation and Remote Control, 821–837.

[2] B, Scholkopf. A, J, Smola. 2002. Learning with Kernels. Cambridge, MA, MIT.

[3] Boser. Bernhard, E. Guyon, Isabelle M. Vapnik, Vladimir, N. 1992. A training algorithm for optimal margin classifiers. ACM Workshop on COLT, ACM Press, 144–152, Pittsburgh, PA.

[4] Devrim, Unay. Bernard, Gosselin. 1998. Thresholding based Segmentation and Apple Graidng by Machine Vision. Computers and Electronics in Agriculture,117-130.

[5] Domingo, Mery. Franco, Pedreschi. 2005. Segmentation of color food images using a robust algorithm. Journal of Food Engineering, 353–360.

[6] J, Shawe, Taylor. N, Cristianini. 2004. Kernel Methods for Pattern Analysis. Cambridge University Press.

[7] Meftah, Salem. M, Alfatni. Abdul, Rashid, Mohamed, Shariff. Helmi, Zulhaidi. Mohd, Shafri. Osama, M, Ben saaed. Omar, M, Eshanta. 2008. Oil Palm Fruit Bunch Grading System Using Red, Green and Blue Digital Number. Journal of Applied Sciences, 1444-1452.

[8] Mehmet, Sezgin. Bulent, Sankur. 2004. Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging, 146–165.

[9] N, Cristianini. J, Shawe, Taylor. 2000. An Introduction to Support Vector Machines. Cambridge University, Cambridge, U.K. Press.

[10] Vladimir, Vapnik. 1995. The Nature of Statistical Learning Theory. Springer-Verlag, New York.

[11] William, H. Teukolsky. Saul, A. Vetterling. William T, Flannery. B, P. 2007. Support Vector Machines, Numerical Recipes: The Art of Scientific Computing, 3rd edition Cambridge University, New York.

[12] Yousef Al Ohali. 2010. Computer vision based date fruit grading system: Design and implementation. College of Computer and Information Science, King Saud University, Riyadh, Saudi, 29-36.