ISSN (ONLINE): 2279-0055, ISSN (PRINTED): 2279-0047

International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

International Association of Scientific Innovation and Research (IASIR)
(An Association Unifying the Sciences, Engineering, and Applied Research)

ISSN (Print): 2279-0047
ISSN (Online): 2279-0055

International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

www.iasir.net

A Comparative Analysis of Feature Extraction Methods for Fruit Grading Classifications

P.Deepa[1] Dr.S.N.Geethalakshmi[2]
[1]Ph.D Research Scholar [2]Professor
Department of computer science,
Avinashilingam Institute for Home Science and Higher Education for Women,
Coimbatore-641042, Tamil Nadu, India

Abstract: Fruit plays an important role in our life. India is exporting large number of fruits to abroad in that mosambi is one among. Before exporting the fruits, the fruits have to be graded according to their quality. This paper presents an evaluation and comparison of the performance of three different extraction methods for classification of defect and non defect fruits. Three different feature extraction methods are GLCM (Grey Level Co-occurrence Matrix), shape features and intensity based features. The performance of each feature extraction method is evaluated and compared based on PNN (Probabilistic Neural Network) classifier. The experimental results suggest that shape feature outperformed than other two methods.

***Keywords**: PNN, Fruit grading, GLCM, Shape, Intensity and PNN.*

I. INTRODUCTION

India being an agricultural country, exports a huge quantity of fruits abroad. But still in our country @40% produce is wasted for want of facilities for preservation and processing. The fruits have to be graded before being packed and sent to the international market for pricing. Grading is done on the basis of various criteria like weight, shape, color, size etc. prior to export; mosambi should be free from damages and defects. Presence of defects and without defect can be seen as in fig 1. Currently the quality inspection of mosambi is done manually. Human inspection involves labor intensive work and judgment made can be very subjective depending on the mood, condition, and knowledge about the fruit grading person involved. Further more, the manual process can be very time consuming and inefficient especially when dealing with high production volume.

a) With defect (b) without defect

Fig 1. Example of mosambi with defect and without defect

Various supervised classifiers such as PNN are used for classification problems in the area of fruit grading. Image feature extraction is important step in fruit classification. These features are extracted using image processing techniques. Several features are extracted from mosambi, such as texture feature, shape feature and intensity based features. Textures are one of the important features used for many applications. Texture features have been widely used in fruit quality classification. The texture features are ability to distinguish between defect and non defect mosambi. Texture is an alteration and variation of surface of the image. In general, texture can be characterized as the space distribution of gray levels in a neighborhood.

Texture feature have been proven to be useful in differentiating defect and non defect pattern in the fruit. Extracted texture features provide information about textural characteristics of the image. Different classifier used in agricultural applications including neural network, support vector machine and fuzzy classifier. Neural network have been widely used for fruit grading. There are two types of texture measure: first order and second order [4]. In the first order, texture measures are statistics calculated from an individual pixel and do not consider pixel neighbor relationships. In the second order, measures consider the relationship between neighbor pixels. The GLCM and shape features is a second order texture calculation and intensity based features are first order texture calculation. Texture features has been and used as parameter to enhance the classification result. This paper presents a comparison among three types of features used in fruit classification. A texture is a method of capturing pattern in the image. These features are calculated using statistical measures such as entropy, contrast and uniformity etc. Automatic classification into defect and non defect pattern is based on the texture features extracted from the fruit images.

The paper is structured as follows. In section 2 related works are discussed. Section 3 deals with the proposed methodology. In section 4 performance measures are explained in detail. Section 5 is the experimental results, followed by conclusions at section 6.

II. RELATED WORK

In the literature, various numbers of techniques are described to detect and classify the presence of defects in the fruit images.

DevrimUnay et.al,[1] used statistical features to identify the defected areas in the apple images. And they are classified using various classifiers such as Linear Discriminant, Nearest Neighbour, Fuzzy Nearest Neighbor and SVM. Using these classifiers they obtained 90% recognition of defected fruit and non defect fruit.

Slamet Riyadi et.al,[2] used centroidal profile to extract features of mean, diameter, maximum of diameter was extracted from papaya object. These features sets are then fed separately to the neural network for the size grade classification. The proposed technique has shown satisfactory result with classification accuracy of more than 95%.

S.Arivazhagan et.al, [3] used intensity, color, shape and texture features for fruit recognition. The recognition is done by the minimum distance classifier based upon the statistical and co-occurrence features derived from the Wavelet transformed sub- bands.

In order to classify mangoes we need to be aware of the mango grading standard. Color and the size are the most significant criteria that are used to sort fruits. However, for sorting of mangoes there is another major factor which is the skin texture of mangoes that can improve the accuracy of the classification system [4].

The attributes chosen for the oranges, peaches and apples [4] are size, color, stem location and detection of external blemishes. Size has been identified using machine vision by measuring area or diameter of fruits. The averaged surface color is a good indicator for these types of fruits. To determine the stem location, they take the images of the random fruit from oranges, apples and peaches. Then, the images analysis algorithms were applied and shown centroid of the stem in the computer. The external blemishes were detected using the combination of infrared and visible information. [5].

Devrim Unay, [6] used statistical, texture and geometric features are extracted from apple images and various classifiers are used to grade the fruit according to the European standards and they obtained 93.5% accuracy.

III. METHODOLOGY

The flowchart for proposed fruit classification is shown in figure 2.

A. Image Acquisition

To evaluate the proposed method, 200 mosambi fruits were collected from the market. The images of mosambi were then captured at random orientation from perpendicular views. The camera was setup in a fixed position to get an appropriate silhouette of the object as shown in figure 3.

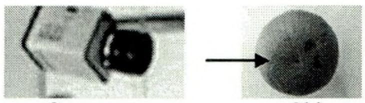


Fig 3. Image acquisition set up

1. Image preprocessing

The pre-processing task involves some procedures to prepare the images to be ready for image processing. The images were initially normalized to produce uniformity in terms of image size and to reduce the processing time. The original image with dimensions of 640×480 pixels was resized to one third of its normal size. An appropriate silhouette should be obtained to get an accurate processing result. For this task, the RGB image was converted to a grayscale.

B. Feature extraction

Feature extraction is a method of capturing visual content of an image. The objective of feature extraction process is to represent raw image in its reduced form to facilitate decision making process such as pattern classification. A variety of technique used for texture feature extraction such as GLCM, shape and intensity based features. Feature extraction step is important step to get high classification rate. A set of features are extracted in order to allow a classifier to distinguish between defect and non defect pattern. The non defect fruit can be identified on the basis of textural appearance. Extracted features are used in neural classifier to train it for the recognition of particular class either defect or non defect. The ability of the classifier to assign the unknown object to the correct class is dependent on the extracted features.

1. GLCM

A GLCM element P θ, d (i, j) is the joint probability of the gray level pairs i and j in a given direction θ separated by distance of d units. In this work seven features are determined for texture discrimination: Autocorrelation, contrast, cluster shade, entropy, sum of squares, and sum of entropy. Their definitions are given by the equations (1-5). Each subdivided block is an independent ROI. Multi-distance and multi-direction can be used to extract a large number of features. In our method we extract GLCM features using one distance d = {1}, and four direction θ = {0^0, 90^0, 180^0, 270^0}, which result in 20 i.e. (1 × 4 × 5) features extracted for each block.

$$\text{Autocorrelation} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(p_x-\mu_x)(p_y-\mu_x)\sigma_x\sigma_y \quad (1)$$

$$\text{Contrast} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)(i-j)^2 \quad (2)$$

$$\text{Cluster shade} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)(i+j-\mu_x-\mu_3)^3 \quad (3)$$

$$\text{Energy} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)^2 \quad (4)$$

$$\text{Entropy} = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)\log(p(i,j)) \quad (5)$$

$$\text{Sum of entropy} = \sum_{i=0}^{2G-2}p_{x+y}(i)\log p_{x+y}(i)) \quad (6)$$

$$\text{Sum of squares} = \sum_{i=0}^{G-1`}\sum_{j=0}^{G-1}p(i,j)(i-\mu)^2 \quad (7)$$

2. Shape Features

The shape features are also called geometric or morphological features. The shape of the defect is valuable feature. These types of features are based on the shape of ROIs. This feature does not consider the intensity of pixels in the region; take the shape of the segmented region.

$$\text{Area} = \sum_{x,y} I(x,y)$$

$$\text{Centroid} = \frac{1}{N}\sum_{i=1}^{N} I(x_i, y_i)$$

$$\text{Diameter} = \max_{x_k, x_1 \in R} d(x_k, x_1)$$

$$\text{Perimeter} = \sum_{i=1}^{N-1}|x_i - x_{i+1}|$$

Minor axis = Minimum axis: it is the smallest distance connecting one point along the region boundary to another point on the region boundary going through the center of the region.

Maximum axis = It is the largest distance connecting one point along the region.

Maximum axis = It is the largest distance connecting one point along the region boundary to another point on the region boundary going through the center of the region.

3. Intensity based features

Pixel intensities are simplest available feature useful for pattern recognition. Intensity features are first order statistics depends only on individual pixel values. The intensity and its variation inside the fruit images can be measured by features like: median, mode, standard deviation and variance. Average intensity of every column in the image is m x n, then the total number of mean is n. Sample features for the three feature extraction method is shown in Table 4.

C. CLASSIFICATION

The proposed method used probabilistic neural network. The schematic representation of neural network with 'n' inputs 'm' hidden units and one output unit [7].The extracted features are considered as input to the neural classifier. PNN is a feed forward neural network. It is a supervised neural network that is widely used in the area of pattern recognition, nonlinear mapping, and estimation of the probability of class membership and likelihood ratios [8]. The neural network trained by adjusting the weights so as to be able to predict the correct class. The desired output was specified as 1 for defect and 2 for non defect. The input features are normalized between 1 and 2.The classification process is divided into the training phase and the testing phase. In the training phase the known fruit images are given. In the testing phase, unknown images are given and the classification is performed using the classifier after training. The accuracy of the classification depends on the efficiency of the training.

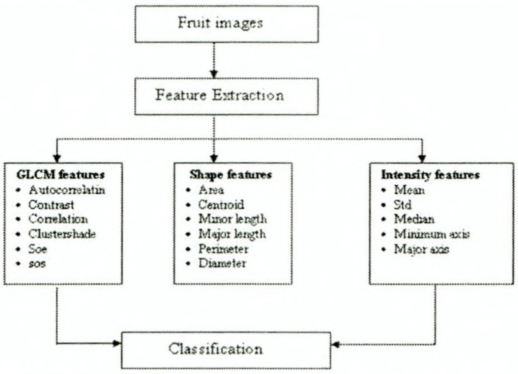


Fig 2. Flow chart for proposed method

Features	Image1	Image2	Image3
Autocorrelation	2317.721	1535.732	1464.386
Contrast	1.072433	0.664854	0.606031
Cluster shade	-11154.5	-6288.98	-2687.15
Energy	0.020463	0.01611	0.018337
Entropy	4.752221	4.659628	4.531491

Table 1: GLCM feature values for defect fruit

Features	Image1	Image2	Image3
Autocorrelation	1771.022	2208.575	2202.072
Contrast	0.438494	0.80648	0.821715
Cluster shade	-7656.26	-12728.9	-11715.5
Energy	0.024761	0.026217	0.033455
Entropy	4.327687	4.625607	4.573867

Table 2: GLCM feature values for non defect fruit

Features	Image1	Image2	Image3
Area	65536	65536	65531
Centroid	128.5	128.5	124.3
diameter	288.8651	288.8651	278.8651
perimeter	1020	987	1000
Major length	295.6033	295.6033	283.5022

Table 3; shape features for defect fruit

Features	Image1	Image2	Image3
Area	75536	72678	74313
Centroid	228.3	214.3	200.2
diameter	400.5671	401.4673	407.4231
perimeter	1100	1103.1	1110
Major length	400.5033	404.5044	401.5

Table4: shape features for non defect fruit

Features	Image1	Image2	Image3
Mean	222.8555	213.6836	214.9063
Std	62.0699	0	56.5944
Median	225	127	217
Maximum	17	35	17
Minimum	212	200	200

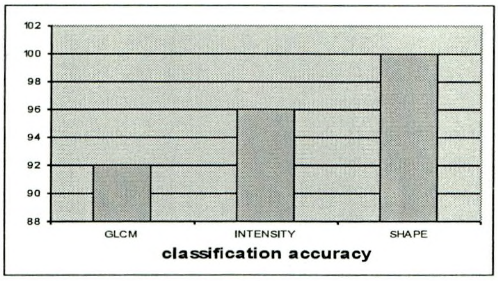
Table 5: Intensity based feature for defect fruit

Features	Image1	Image2	Image3
Mean	208.0313	237.2148	236.3008
Std	45.9514	237.2148	54.2519
Median	209	238	237
Maximum	49	35	40
Minimum	197	231	232

Table 6: Intensity based feature for non defect fruit

IV. EXPERIMENTAL RESULTS

The efficiency of three different feature extraction methods is trained and tested using PNN classifier. The dataset used for this experiment is composed of 200 mosambi images. Which includes 100 with defect and 100 without defect, 80% (150 out of 200) set of images are used for training and 20% (100 out of 200) used for testing. The effectiveness of the three different feature extraction methods were evaluated and compared. Three experiments are conducted. In each experiment, the architecture of the PNN, training and testing samples are same. In the first experiment GLCM features were extracted and its classification was done using PNN classifier. In the second experiment shape features were extracted and its classification was done and in the third experiment the intensity features were extracted and was classified. The results shows that GLCM features based PNN is giving 96% classification rate, shape feature is giving 100% classification rate and intensity features is giving 92% classification rate.



V. CONCLUSION

This paper examined the three different types of feature extraction method. The result proves that shape feature based PNN is giving higher classification rate of 100%. The shape gives a better performance when compared with GLCM and intensity features. In future, the fruits can be graded with Indian standards with different classifiers to find out the optimum classification procedure

.

VI. REFERENCES

1. Devrim Unay et al, 2010, Automatic grading of Bi-colored apples by multispectral machine vision, Journal of Computers and Electronics in Agriculture, Elsevier.
2. S.Arivazhagan et al, Oct 2010, Fruit Recognition using Color and Texture Features, Journal of Emerging Trends in Computing and Information Sciences, Vol. 1, No. 2.
3. Slamet Riyadi, et al, 2007, Papaya Size Grading using Centroidal Profile Analysis of Digital Image, 6th wseas international conference on circuits, systems, electronics,control & signal processing.
4. Blasco, J., Aleixos, N. and Molto, E(2003). Machine Vision System Quality Grading of Fruit, *Biosystems Engineering*, 415-423.
5. Leemansa, V., Mageinb, H. and Destain, M.-F.(2002). On-line FruitGrading According to their External Quality using Machine Vision. *Journal of Biosystems Engineering, 83,* 397–404.
6. D. F. Specht; H. Romsdahl, 1994, .Experience with adaptive probabilistic neural network and adaptive general regression neural network, In *Proceedings of the IEEE International Conference on Neural Networks*, volume 2, pp.1203.1208.
7. N.K. Bose and P. Liang. Neural network fundamentals with graphs, algorithms and application. New Delhi: McGraw-Hill. 1998
8. R.O. Duda, P.E. Hart and D.G. Stork. Pattern Classification. CA: John Willey & Sons. 2001