IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

Fruit Identification and Classification Techniques A Review Using Neural Networks Approach

Ms. S.Mahalakshmi[1], Harini Srinivas[2], Meghana S[3], C Sai Ashwini[4]

Assistant Professor, Dept of Information Science & Engg, BMS Institute of Technology & Mgmt., Bengaluru, India[1]

Student, Department of Information Science & Engg., BMS Institute of Technology & Mgmt., Bengaluru, India[2,3,4]

Abstract: This paper gives an insight into the results of a survey based on fruit identification and classification. The survey depicts the credibility of choosing the appropriate classifier and the feature extraction methods for correct and exact recognition of the fruits. It also reports on the accuracy and performance of each method implemented in the papers taken into consideration. Morphological features, color features, intensity based features and other features are extracted from the fruit images and these are subject to various types of classifiers like Probabilistic Neural Network(PNN), Support Vector Machine(SVM), Back Propagation Network(BPN) and K-Nearest Neighbour(KNN) algorithm. The goal of this paper is an overview of the techniques implemented in fruit identification and classification.

Keywords: Fruits; feature; classifiers; SVM; ANN; KNN; PNN.

I. INTRODUCTION

The fruit industry plays a vital role in a country's economic growth. They account for a fraction of the agricultural output produced by a country. It forms a part of the food processing industry. Fruits are a major source of energy, vitamins, minerals, fiber and other nutrients. They contribute to an essential part of our diet. Fruits come in varying shapes, color and sizes. Some of them are exported, thereby yielding profit to the industry. Fruit sorting and grading are performed before export. This determines the quality of the fruits which is an important factor in the food processing industry.

The food industry has widely used machine vision for quality inspection of fruits, vegetable and processed food [1]. The first step in machine vision is image acquisition, followed by image processing and finally, obtaining the required output, which may or may not satisfy the requirements of the user. Computer vision uses the concept of feature detection, a low-level image processing operation. These features are then extracted resulting in a feature vector which identifies a fruit, based on color, shape, texture and intensity based features. They are then fed to various classifiers to classify them accordingly. The various classifiers include Probabilistic Neural Network (PNN), Support Vector Machine (SVM), Back Propagation Network (BPN), K-Nearest Neighbor (KNN) algorithm, K-Mean classifier, Naive Bayes Classifier, Linear Discriminant classifier and others.

The various applications of fruit recognition and classification are in agricultural and horticultural fields, in robotic fruit harvesting, where robots detect fruits on trees grown in large plantations, which can later be harvested. It also finds application in plantation science and in supermarkets to identify fruits. It can also be applied for educational purpose to enhance learning, especially for small kids and Down syndrome patients [2]. The performance and accuracy of each method implemented in the papers surveyed are discussed. This paper is organized into 3 sections. Section II describes the survey. Section III concludes the paper.

II. LITERATURE SURVEY

This section is divided into 9 sub-sections, A to I, each describing the methods used in the papers considered.

A. Discrete Curvelet Transform

Suchitra[1] discussed Fruit Skin Defect Identification System for Fruit grading based on Discrete Curvelet Transform. It implements the concept of computer vision. In this method, the fruit image is first decomposed using Curvelet Transform, and then, textural features such as Energy, Entropy, Mean and Standard deviation are extracted. These are then tested on classifiers, like SVM and PNN. Image pre-processing includes the conversion of an image in RGB color space to grayscale using the MATLAB function, 'rgb2gray'.The image is then resized to 256 by 256 pixels. The curvelet transform is then implemented, where, first, 2D Fast Fourier Transform (FFT) of the image is taken, and then, a digital implementation based on unequal-spaced Fast Fourier Transform (USFFT) is done. Thus, the statistical measures for texture classification are extracted. Only 4 of them are used for fruit grading. MATLAB v7.0 for image pre-processing, segmentation, color transformations, texture feature extraction, classification, and CurveLab Toolbox, Version 2.1.3 for the calculation of fast discrete Curvelet transform, are used. For fruit classification, SVM classifier, with radial basis kernel, and Probabilistic Neural Network classifier are implemented. According to the results obtained, SVM classifier with energy, entropy and mean texture features, achieved 91.42% and 91.72% for Guava and Lemon. PNN achieved classification rates of 82.85% and 89.45% for Guava and Lemon. Thus, SVM is found to be a better classifier due to its higher accuracy.

B. Colour and Texture Features

S. Arivazhagan [3] proposed a fruit recognition method based on the fusion of color and texture features. It uses

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

the minimum distance classifier on the statistical and co-occurrence features obtained from the Wavelet transformed sub-bands. The original fruit image is first converted into its HSV representation, and the texture features are computed from the luminance 'V' channel, and color features are computed from the chrominance channels, 'H' and 'S'. Discrete Curvelet transform is used to decompose the 'V' component and a co-occurrence matrix is obtained from the approximation sub-band. Features like, contrast, energy, local homogeneity, cluster shade and cluster prominence are thus calculated from the co-occurrence matrix. Statistical features include mean, standard deviation, skewness, kurtosis, hue, saturation, luminance and discrete curvelet transform, and are derived from 'H', 'S' and 'V' components. A total of 13 features, 8 chrominance and 5 texture features are thus computed. For fruit classification, the images are divided into training and testing sets. Some images are used as training set, whereas the remaining images constitute the testing set. Background subtraction is performed on the images, in the S channel, to focus only on the object, rather than, on other parts of the image. Minimum distance classifier (k-means classifier) is used for the classification. A combination of color and texture features resulted in a good classification rate of 86.00%, than compared to individual rates of 45.49% and 70.85% using only color features and only texture features, respectively. Thus, fruit recognition rate can be improved by taking into account both color and texture features.

C. Artificial Neural Network

Dayanand Savakar [4] presents a technique to identify and classify bulk fruit images using Artificial Neural Network. A Back Propagation Neural Network (BPNN) is used to identify and classify the images. The images are pre-processed. Color and texture features are extracted from these images and stored for future use. They form the training set. When a new image is to be identified, features are extracted from it, and classified using the stored extracted features and an Artificial Neural network (ANN), trained using the feed forward rule. The algorithms developed, extract 18 color and 27 texture features. For color feature extraction, Hue, Saturation and Intensity components are extracted from RGB components and are calculated using the required formulae. The variation in the intensity of a fruit surface can be exhibited by the texture features, such as smoothness and regularity.

For texture feature extraction, gray level co-occurrence matrix is used to extract features such as mean, variance, range, energy, entropy, contrast, inverse difference moment, correlation and homogeneity. For fruit classification, the ANN model implements the back propagation algorithm. The number of neurons in the input layer is the number of input features. The number of neurons in the output layer is the number of categories of fruit samples. Five categories of fruits, Apple, Chickoo, Mango, Orange and Sweet lemon are considered. When only color features are used, the classification rate using BPNN, for Chickoo is 90% and that, for Mango is 87%. When texture features only are taken into consideration, the classification rates increase, that is, 93% for Chickoo and 90% for Mango. A combination of color and texture features gives the classification rates as, 94% for Chickoo and 92% for Mango. Thus, color and texture features should be taken into consideration for accurate identification and classification of fruits using BPNN.

D. Principal Component Analysis

Yudong discusses the classification of fruits using Computer vision and a Multiclass kernel Support Vector Machine[5][15]. The first step includes, image acquisition by a digital camera, and background removal by a split-and-merge algorithm. The second step involves the extraction of color features using a color histogram, texture features and shape features to obtain a feature space. In the third step, Principal Component Analysis (PCA) is performed, to reduce the dimensions of the obtained feature space. The multiclass SVMs constructed includes Winner-Takes-All SVM, Max-Wins-Voting SVM, and Directed Acyclic Graph SVM. The kernels used are linear kernel, Homogeneous polynomial kernel and Gaussian radial basis kernel. A color histogram gives the distribution of colors in an image. For color feature extraction, the original image is reduced to 64 colors and an image histogram is constructed. For texture feature extraction, Unser's texture feature vector is used. The sum and difference histograms define seven indexes, namely, mean, contrast, homogeneity, energy, variance, correlation, and entropy for texture. Shape features include area, perimeter, Euler, convex, solidity, Minor Length, Major Length and Eccentricity. 64 color, 7 texture and 8 shape features contribute to a total of 79 features. These should be first normalized before performing PCA. PCA transforms the original set of correlates variables into a new set of ordered variables according to the degree of importance. The output of the PCA is then fed to a kernel SVM. In a Winner-Takes-All SVM, first, a number of binary SVMs are trained to distinguish data in a single class from the data of the remaining classes. When this is applied to a new test data, all the classifiers are run, and the classifier which gives the largest value is chosen. In a Max-Wins-Voting SVM, a binary SVM is constructed for each pair of classes. When it is applied to a new test data, each SVM gives a vote to the winning class. The test data is then given the label of the class having the maximum labels. In a directed Acyclic Graph SVM, each SVM is constructed as the Max-Wins-Voting SVM. It can be represented as a tree with a root node, intermediate nodes and leaf nodes. Starting from the root node, the individual binary SVMs are evaluated either to the left or right edge.

The process continues till a leaf node is reached. All the SVMs are then trained using 5-fold cross validation with the input as reduced feature vector of 14 features. According to the results, the classification accuracy of Max-Wins-Voting SVM for the Gaussian Radial Basis kernel is the highest, with 88.2% and the Directed Acyclic graph SVM is the fastest compared to the other SVMs. Thus, best results are obtained, if the Max-Wins-Voting SVM is implemented with Gaussian Radial Basis kernel. If time consumption is taken into account, then the

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

Directed Acyclic Graph SVM performs better than the others.

E. FPGA based

Harsh [6] deals with a Field-programmable Gate Array (FPGA) based Efficient Fruit Recognition System using Minimum distance classifier. A FPGA is an integrated circuit whose configuration is specified using a Hardware Descriptive Language (HDL). The platform used is SPARTAN 3. JAVA ECLIPSE IDE generates the COE file of a test image. A Verilog platform is developed by burning this file onto the IP core. The Verilog simulations are done in Xilinx ISE 10.1 and ISIM. First, the 3D feature vector is obtained which includes the first order statistical features and the shape feature. MATLAB Analysis of the Minimum Distance classifier is then performed using the feature vector. For feature extraction, the first order statistical features, such as, mean and variance is extracted and, shape feature is calculated using the area and perimeter of the fruit. Image segmentation and Edge Detection is then performed. The former is done to separate the background from the fruit. A threshold based automatic segmentation technique is implemented by an Inter Means Algorithm, which is an iterative algorithm that finds the threshold value which splits the image histogram into two halves. After the threshold is calculated, the image is converted to a binary image. Edge detection identifies points in an image where there are discontinuities. The Approximate Robert edge detection method is used here. A minimum distance classifier is then used, which classifies unknown image data into classes and reduces the distance between the image data and the class. It uses the Euclidean distance formula and is implemented in two phases, training and testing. The appropriate training and testing algorithms are performed. One of the hardware constraints encountered is, no provision for camera interfacing in the platform. Therefore, to solve this problem, the test image coefficients have to be burned on the memory. Training the data on FPGA is difficult. So, MATLAB Analysis is done and it is found that the Verilog and MATLAB outputs are almost the same. The FPGA output is obtained in the FPGA board where the glowing of the LEDs identifies the appropriate fruit. The classification rate for this method is about 85%. Thus, the paper implements a real time fruit recognition system resulting in hardware based approach.

F. Color, shape and Size

Woo Chaw Seng [2] presents a New Fruit Recognition System that processes, analyses, classifies and identifies the fruit images, based on color, shape and size features of the fruit. The K-Nearest Neighbors (KNN) Algorithm performs the classification, using the Euclidean distance metric to measure the distance between the attributes of the unknown fruit and the fruit examples stored. Using the KNN algorithm as a classifier, the Fruit Recognition System classifies the fruits based on mean color values, shape roundness value, area and perimeter values of the fruit. The color values are obtained by cropping the fruit area. RGB values are computed for each pixel using the mean function of MATLAB. To compute the roundness value, the fruit region properties are extracted to calculate the area and perimeter of the fruit. The KNN algorithm calculates the 'K' shortest distance or closest examples to the input fruit, and then assigns the input fruit to the class where majority of the 'K' closest examples is from.

For Fruit Recognition System, the value chosen for 'K' is '1', meaning the classification of input fruit sample is based on the class of the closest fruit example. The Fruit Recognition System consists of five main processing modules, which are, fruit input selection module, fruit color computing module, fruit shape computing module, fruit size computing module and classification or recognition module. The classification or recognition module is responsible to classify the input or user selected fruit using the KNN algorithm. It measures the distance between the feature values of the input fruit and the stored fruit and finds the fruit example that has the shortest distance with the input. The system will then identify and assign the class to the input fruit. The Fruit Recognition System developed recognizes all the test fruit images selected by user or system tester from the fruit selection menu on the system. The recognition results of the system are accurate up to 90%. This system serves as useful tool in a variety fields such as educational, image retrieval and plantation science.

G. Shape based Identification

Deepa [7] analyzes the evaluation and comparison of the performance of three different feature extraction methods for classification of defect and non-defect fruits. The methods include GLCM (Grey Level Co-occurrence Matrix) for texture features, Shape features and Intensity based features. The performance of each feature extraction method is evaluated and compared based on Probabilistic Neural Network (PNN) classifier. Methodology consists of Image Acquisition, where the images of mosambis are captured at random orientation from perpendicular views to get an appropriate silhouette of the object, Image preprocessing, where the original image with dimensions of 640*480 pixels is resized to one third of its normal size and the RGB image is converted to grayscale, Feature Extraction, where raw images are represented in its reduced form to facilitate decision making process. The GLCM features include Autocorrelation, Contrast, Cluster shade, Entropy, Energy, Sum of entropy, and Sum of squares. They are extracted using a distance d= {1}, and four directions, θ= {0, 90,180,270}, which result in 20 features for a block. The Shape features are based on the shape of the Region of Interest (ROI) and take the shape of the segmented regions. It includes area, centroid, diameter, perimeter, minor axis and maximum axis. The Intensity features are first order statistics and depend only on individual pixel values. Intensity is measured by features like median, mode, standard deviation and variance. The features extracted are used to train the neural network classifier. PNN is a feed forward neural network whose input is the extracted features and the desired output is specified as 1 for defect and 2 for non defect. The classification process is organized into training and testing

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

phase. Unknown images in the testing phase are classified using the trained classifier. The effectiveness of the three different feature extraction methods on PNN are evaluated and compared. The results show that the GLCM features gives a classification rate of 96%, Shape feature gives 100% classification rate and Intensity features gives 92% classification rate. Thus the result concludes that the Shape feature extraction is better than GLCM and Intensity feature extraction as it gives the highest accuracy in classification using PNN. PNN is used for classification problems in the area of fruit grading.

H. Fruit Detection using multiple features algorithm
The paper by Patel [8] reports on the fruit detection technique using improved multiple features based algorithm to detect fruits on trees. It does not involve the classification of fruits. The algorithm is developed to calculate the different weights for features such as intensity, color, orientation and edge for the input image. The weights give the location of the fruit in an image. The first step in this technique is feature extraction. The original image is converted to a HSV color model, where 'H' and 'S' represent the color features and 'I' gives the intensity features of the input image. The orientation feature maps are obtained from the intensity feature map by filtering it using four gabor filters at 0, 45, 90, and 135 degrees. Edge features involves identifying discontinuities in an image. The gradient magnitude is as a measure of the edge features. The extracted features are thus integrated using weights to form an integrated map. Feature maps are constructed from the binary maps which, are generated when the integrated map is segmented using global threshold. Fruit regions are then, extracted from the feature maps. The steps involved are first, the image is reconstructed with phase values of frequency spectrum to obtain the region of interest. Local and Global points of the fruit region are computed. The rare feature values are found out by using the histogram of the feature map. A conspicuity map is generated by combining the local, global and rare feature points. The feature points of the extracted features are used to calculate the weights. Thus, the fruit region is extracted. The extracted fruit region is used for generating the binary mask by comparing each pixel values with global threshold. The threshold has to be selected properly for successful extraction of the fruit region. The efficiency of the algorithm depends on the weights and the extracted features. This process is automatic and does not need user intervention. Thus, a multiple feature analysis is used in fruit detection technique to recognize fruits on trees, with an efficiency of up to 90%.

I. Combined Features
Bhanu Pratap [9] and Basvaraj [12] proposes an algorithm for fruits classification based on the shape, color and texture features. Artificial neural network (ANN) is used as a classifier that recognizes and classifies the fruits to the class to which they belong. The input layer of the neural network depends upon the number of inputs. There is a hidden layer, which consists of neurons that process the information and generates the output. It has five output layers because fruits are classified in five different classes. The methodology consists of image capture and preprocessing, feature extraction and classifying the extracted features based on ANN. The images are taken from a Digital Camera (Nikon) placed at a height of 1 foot. Edge detection is then, applied, as a preprocessing technique, to remove high frequency noise by using a low pass Gaussian filter. After the image is made noise free, different features like shape, color and texture features are extracted. For shape based classification of fruit, area, perimeter, major axis length and minor axis length are calculated by object segmentation with the background using edge detection techniques. The RGB image is converted into a gray scale image, and using a threshold, a binary image is obtained to compute the shape features. For color based classification, mean and standard deviation are calculated for the color space like HSI, HSV. The RGB components are separated from the original image to extract the 'H', 'S' and 'I' components. The Gray Level Co-occurrence Matrix (GLCM) is used to calculate texture features. It gives a description of pixel brightness value (gray levels) in an image. Different texture features such as entropy, energy, homogeneity and dissimilarity, are computed from the matrix. All these features are stored in a database as training set, for future use. The features extracted from an unknown image are compared with the stored features, and are classified by the ANN into a suitable class. An Algorithm is developed to extract 4 shape features, 16 color features and 22 texture features from fruit image. The neural network performs the classification based on shape, color, texture and both color and texture features. MATLAB/ SIMULINK software is used to obtain result. From the result, Shape based classification gives 83.2% accuracy, Color based gives 90%, and Texture based gives 89.60%. When color and texture features are combined, results are improved to 96 %. Hence it can be concluded that color and texture together give a better result. If shape is used along with the color and texture features, there may be complexity in the computation of the results. Thus, best results are obtained for a combination of color and texture features, than compared to the classification using individual features such as shape, color and texture.

III. CONCLUSION

The literature survey describes the various techniques implemented and the classifiers used. Out of these, the best method chosen is the one which uses ANN as the classifier to obtain a classification rate of 96% for the combination of color and texture features. Thus, Neural Networks can be pre-dominantly used for the classification of fruits as they result in a higher degree of accuracy, than compared to other classifiers. Neural Networks find applications in predicting weather, complex problem solving, pattern recognition, image matching, face recognition, and many others. We conclude that Neural networks classifiers are best for fruit recognition with the combined features as though it gives more complexity will yield good accuracy. These are no proper methods to identify the defection from inside. Even though soft x ray

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

methods are used it will have an effect on fruits. Our research is based on identifying the defect of fruits based on reduced features set.

REFERENCES

[1] Suchitra A. Khoje, Dr. S. K. Bodhe, Dr. Alpana Adsul, "Automated Skin Defect Identification System for Fruit Grading Based on Discrete Curvelet Transform", International Journal of Engineering and Technology (IJET), Vol 5 No 4 Aug-Sep 2013.

[2] Woo Chaw Seng and Seyed Hadi Mirisaee, "A New Method for Fruits Recognition System", MNCC Transactions on ICT, Vol. 1, No. 1, June 2009.

[3] S. Arivazhagan, R. Newlin Shebiah, S. Selva Nidhyanandhan, L. Ganesan, "Fruit Recognition using Color and Texture features", Journal of Emerging Trends in Computing and Information Sciences(CIS Journal), Vol. 1, No. 2, Oct 2010.

[4] Dayanand Savakar, "Identification and Classification of Bulk Fruits Images using Artificial Neural Networks", International Journal of Engineering and Innovative Technology (IJEIT), Volume 1, Issue 3, March 2012.

[5] Yudong Zhang and Lenan Wu, "Classification of Fruits using Computer vision and a Multiclass Support Vector machine", Sensors 2012, 12.

[6] Harsh S Holalad, Preethi Warrier, Aniket D Sabarad, "An FPGA based Efficient Fruit Recognition System Using Minimum Distance Classifier", Journal of Information Engineering and Applications, Vol. 2, No. 6, 2012.

[7] P.Deepa, Dr. S. N. Geethalakshmi, "A Comparative Analysis of Feature Extraction Methods for Fruit Grading Classifications", International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS), March-May 2013.

[8] Hetal N. Patel, Dr. R.K.Jain, Dr. M.V.Joshi, "Fruit Detection using Improved Multiple Features based Algorithm", International Journal of Computer Applications, Volume 13, No. 2, January 2011.

[9] Bhanu Pratap, Navneet Agarwal, Sunil Joshi and Suriti Gupta, "Development of ANN Based Efficient Fruit Recognition Technique", Global Journal of Computer Science and Technology: C Software and Data Engineering, Volume 14 Issue 5, Version 1.0, Year 2014.

[10] Jasmeen Gill, Dr. Parvinder Singh Sandhu and Dr. Tejwant Singh, “A Review of Automatic Fruit Classification using Soft Computing Techniques”, International Conference on Computer, Systems and Electronics Engineering (ICSCEE), April 2014.

[11] Manali Kshirsagar and Parul Arora, “Classification Techniques for Computer Vision Based Fruit Quality Inspection: A Review”, International Journal of Recent Advances in Engineering and Technology (IJRAET), Volume-2, Issue-3, 2014.

[12] Basvaraj .S. Anami, J.D. Pujari, Rajesh Yakkundimath, “Identification and Classification of Normal and Affected Agriculture/horticulture Produce Based on Combined Color and Texture Feature Extraction”, International Journal of Computer Applications in Engineering Sciences, Vol. I, Issue III, September 2011.

[13] Jyoti A Kodagali and S Balaji, “Computer Vision and Image Analysis based Techniques for Automatic Characterization of Fruits – a Review”, International Journal of Computer Applications, Volume 50, No. 6, July 2012.

[14] Seema Paragi, H. Girish, ”A Comparative Study of Image Classifiers in a Fruit Recognition System”, International Journal of Advanced Research in Computer Engineering and Technology (IJARCET), Volume 3 Issue 6, June 2014.

[15] Kondo N, Ahmad U, Monta M and Murasc H, “Machine vision based quality evaluation of Iyokan orange fruit using neural networks”, Computers and Electronics in Agriculture, Vol. 29, 2000.