Abstract — Automation in food processing comes into play to increase productivity, quality and profitable growth of the country. Fruit grading is an important process for producers which affects the fruits quality evaluation and export market. Although the grading and sorting can be done by the human, but it is slow, labour intensive and tedious. Hence, there is a need of an intelligent fruit grading system. In recent years, researchers had developed numerous algorithms for fruit sorting using computer vision. Colour, textural and morphological features are the most commonly used to identify the diseases, maturity and class of the fruits. Techniques include clustering and color based segmentation, artificial neural network and different classifiers based classification of diseases. The main focus of our work is obtaining the analysis of different fruit diseases detection techniques. Subsequently, these features are used to train soft computing technique network. In this paper, use of image processing in agriculture has been reviewed so as to provide an insight to the use of vision based systems by highlighting their advantages and disadvantages.

Keywords— Fruit diseases, SVM classifier, segmentation, Feature extraction, Automation, Fruit classification, computer vision, colour, morphological, texture.

I. INTRODUCTION

In recent years, use of image processing has been increasing day by day in different areas such as industrial image processing, medical imaging, real time imaging, texture classification, object recognition, etc. Image processing and computer vision in agriculture is another fast growing research field. It is an important analysing tool for pre-harvest to post-harvest of crops. It has lots of applications in agriculture. The cultivation of crops can be improved by the technological support. Fruits and vegetables losses are caused by disease. Diseases are seen on the leaves and fruits of plant, therefore disease detection plays an important role in cultivation of crops. Pathogens, fungi, microorganism, bacteria and viruses are types of fruit diseases also unhealthy environment is responsible for diseases. There are lots of techniques to identify diseases in fruits in its early stages. The old method of disease detection in fruit is naked eye observation and it’s not effective. Using digital method, the disease detection can efficient, accurately, time consuming is less, saves time. Different image processing techniques and lots of algorithms have been developed by researchers with the help of MATLAB for accurate fruit disease identification. In land identification, image processing is used for identification of land that will be suitable for agriculture [1]. In plant nitrogen identification, image processing can be used for estimation of plant nitrogen identification and chlorophyll identification [2]. In pest control, image processing is a good tool for identification of pest infected areas because it favors to build up the pest population [3] [4].

Using computer vision, image processing is used to automatic detection and classification of plant disease from colour, texture and shape [5]. Food quality can be improved by quality inspection using computer vision [6].

II. BASIC STEPS OF IMAGE PROCESSING

Step1: Image Acquisition: This is the first step of image processing in which camera is used for capturing fruits images in digital form and store in any digital media.

Step2: Image Pre-processing: This section removes noise, smoothen the image also perform resizing of images. RGB images are converted to the grey images also contrast of image is increased at certain level.

Step3: Image Segmentation: Segmentation is used for partitioning an image into various parts.

Step4: Feature Extraction: This section is used for obtaining features like color, texture and shape which reduce resources to describe large set of data before classification of image.

Step5: Classification: This section analyzes numerical property of image features and organize its data into categories. It use neural network which performs training and classification of fruits diseases.

III. COMPUTER VISION

Computer vision is used to gather the information from the images which are captured from the real time world. It is a field that includes methods for image acquisition, processing, analyzing and understanding the images in order to gather symbolic and numerical information. Basically its aim is to duplicate the effect of human vision by electronically perceiving, understanding and classification of images[7]. Computer vision is widely used in post-harvest industries for quality inspection and grading of fruits and vegetables.

A. Computer Vision Setup

A computer vision system constitutes of five components : illumination, a camera, an image capture board also known as a frame grabber or digitizer, computer hardware and software [8].

B. Colour Features

Images are captured in RGB colour models. It is most common Colour model in image processing and based on primary colours red(R), green (G), blue (B). Basically, for colour features, every image is separated into red, green and blue planes, respectively, and through these planes, mean, median, standard deviation are calculated [9]. NTSC or YIQ colour space consists of three components luminance (Y), which represents gray scale information, hue (I) and saturation (Q), which carry information of a signal.
This colour model is used in television in United State. YCbCr colour space is used in digital video. Here, Y contains luminance information, Cb contains colour information between blue component and reference value and Cr contains colour information between blue component and reference value [10]. HSV colour space is used by people to select colours from a colour wheel or palette. In this colour space, H means hue that refers to tint; S means saturation that represents shade; V means value that refers to tone. HSI colour space means hue, saturation and intensity. HSI is the best tool for developing image processing algorithm based on colour that are natural and perceived by humans. Image acquisition method is capturing of digital images, after that the paper deals image pre-processing techniques which includes feature extraction. Three feature vectors namely: morphology, texture and color are used for feature extraction. Image components for boundaries are extracted using morphology. Various visual patterns are described by texture feature. RGB colour space is converted to HSI colour space in color feature extraction also histogram of image is computed. ANN neural network and back propagation algorithm is used for classification diseases in network. Lastly fruit grading is determined by the calculation of weight and spread of disease on the fruit[36]. Then color based segmentation techniques such as clustering, YCbCr, RGB, La*b*, HSV are used. In feature extraction three features are extracted. Morphology, color, texture features are helpful for the classification purpose. In texture feature extraction, Gabor filter is used and in morphology feature extraction, boundary of image is obtained. The eroded images are subtracted from original image to extract shape vector from healthy fruit image. Using minimum distance classifier (MDC), diseased and non-diseased fruit are classified after training and testing of images[44].

C. Morphological Features

The most commonly used features for classification of fruits is morphological features, i.e. shape and size. Size features are physical dimensional measurements that tell about the appearance of an object. Area, perimeter, major and minor axis lengths, and aspect ratio are commonly used as morphological features. Morphological features are widely used in automatic sorting purpose in industries [9]. The area is a scalar quantity that actual number of pixels in the region. Perimeter is a scalar quantity and it is the distance around the boundary of the region. Major axis length is scalar quantity that is specifying the length of major axis (in pixels) of the major axis of the ellipse that has the same normalized second central moment as the region. Minor axis length is also scalar that is specifying the length of minor axis (in pixels) of the minor axis of the ellipse that has the same normalized second central moment as the region. Shape features are measured by roundness \((4\pi \times \text{Area}/\text{Perimeter}^2)\), aspect ratio (Major Axis/Minor Axis) and compactness (Perimeter/2/Area)[12].

D. Texture Features

Texture represents the surface appearance and distribution of elements. It is an important feature in machine vision and predicts surface in form of contrast, roughness, orientation, entropy, etc. A wide variety of techniques have been proposed for describing image texture. In model based approach, a set of parameters is derived from variation of pixel elements that is used to define an image model such as Gaussian Markov Random Field (GMRF), fractional Brownian motion (fBm). In GMRF, conditional probability of a certain pixel depends upon the value of the neighbouring pixels, while fBm exploits the self-similarity of texture at varying scales. Statistical based methods are derived by analysing the spatial distribution of pixel gray level values such as Gray level co-occurrence matrix (GLCM), run-length and auto covariance function [13].

E. Application of Computer Vision in Fruit Grading System

Sorting of agriculture products automatically is more efficient as compared to the current manual system which is very slow, tedious, labour intensive and error prone. However, there is a need of an automatic sorting system which can identify the agricultural products based on their characteristics. Computer vision has been widely used in fruit grading system. Grading is based upon shape, size, colour, intensity and texture of the fruits. Savarak had graded five different types of fruit images(Apple, Chiccoo, Orange, Mango and Sweet Lemon). Total 5000 sample images had been captured, i.e. 1000 images of each type of fruit. The algorithm had been developed by extracting 18 colours and 27 texture features. The colour features were calculated by separating RGB (Red, Green and Blue) components. The RGB image was then converted into HIS model and its components were separated. Mean, variance and ranges were calculated for each RGB and HSI component individually. Texture features were calculated using Gray Level Co-occurrence Matrices (GLCM). The study revealed that classification of chickoo, apple, sweet lemon, orange and mango were 94%, 93%, 93%, 92%, 92% respectively [14]. Deepa proposed a method to evaluate the extracted features used for grading and classification of defected and non-defected fruits. The image database included 200 mosambi fruits. Shape features, intensity features and texture features were calculated. This database was then classified based on PNN and the result showed that shape, intensity and texture features gave 100%, 92%, 96% identification rate respectively [15]. Mustafa et al. presented a novel approach for fruit grading system. In this paper five fruits (apples, bananas, carrots, mangoes and oranges) were analysed. Shape and colour features were extracted from fruit sample images. Morphological features were used to distinguish between almost similar shapes and sizes such as apple and orange or bananas and carrot. Colour features were used to remove misclassification between apple and orange or banana and carrot and increase the accuracy to 79-90% [9]. Khojastehnaz and et al. proposed an algorithm for sorting and classification of lemon fruits based upon the colour and size in Visual Basic 6. Volume of sample image had been calculated and RGB images were converted into HSI images. HSI values were estimated and these data were stored in a database. During the sorting stage, calculated volume and colour are compared with the saved information in the database. The overall system gave 94.04% accuracy [16]. Jack man et al. proposed a system for food quality assessment based on computer vision. In this paper, to extract surface texture feature of food, three approaches had been suggested which were pixel co-occurrence, Run length and Difference Histogram methods. Fourier Transform was also proposed as a method to extract texture feature in term of convolution of sinusoidal waves but it had a limitation that it is used only where a small number of frequencies can reproduce the surface image.

Another classical approach Wavelet Transform was also proposed for texture analysis. An artificial intelligence was provided to computer to analysis the texture which was efficient and robust [17]. Liminget al. presented a system for automatic grading of strawberry. In this RGB image was changed into L*a*b* colour model. The size was calculated using the major axis length and colour features were extracted.
from the dominant colour model on a* channel. K means clustering method was used for classification purpose and it gave 90% accuracy for shape classification and 88.8% accuracy for colour grading [18]. Alavi proposed a system for grading of Mozafati dates. Total 100 fruits were taken and three quality parameters were calculated such as quantity of juice, size and freshness. These fruits were graded using both fuzzy interference system and human experts for comparison and graded using fuzzy showed 86% conformity result as compared to human experts [19]. Suresha et al presented an automatic grading of apples with the help of support vector machines (SVM). In this, apple images were captured into RGB colour model and threshold based segmentation was used to extract the region of interest from the background. The RGB colour model was then converted into HSV model and average red and green colour components were determined for classification. This classifier gave 100% accuracy in grading [20]. Leemans et al. proposed apples grading method and two types of apple were used—Golden Delicious and Jona gold. Features were extracted from shape, colour, texture and stem position. This method for apple external quality grading gave 72% accuracy for Golden Delicious and 78% Jona gold apples. The grading of healthy fruits was better and an error rate decreases to 5 and 10%, respectively [21]. Nozari et al. presented an algorithm for grading of Mozafatidates. Total 100 fruits were taken and classified based on length, width and thickness. These fruits were graded using both ANFIS and human experts for comparison and ANFIS showed 93.5% conformity result as compared to human experts [22]. Razak et al. presented automatic grading of mango using fuzzy analysis. In this size, colour and skin features were extracted. Size of mango was determined by calculating area of sample image. Then RGB component was extracted from image and mean of three colour components was detected. For shape analysis, edge detection algorithm had been implemented. Fuzzy inference rules were applied for mango grading into different classes and it gave 80% overall accuracy [23]. Kavdir et al. proposed a method of apple grading in which some quality features were extracted such as colour, size and defects of apples. These features were gathered and evaluated using the fuzzy system and this gave 89% accuracy in classification [24]. Zhang proposed that a new method to differentiate apple stem-end/calyx from true defect according to their depth information on apple surface. The projector was designed to generate NIR structured light based on position encoding. The NIR structured light was used to sense the depth information of apple surface, then the stem-end/calyx region was identified. It gave an overall 95.24% detection rate [25].

IV. CLASSIFIER MODELS

This section includes the most popular classifier models that include fuzzy logic, artificial neural network, support vector machine and adaptive network-based fuzzy interference system.

A. Zadeh Logic

Zadeh proposed the theory of fuzzy sets. This theory introduced making of the membership function operations over the range of real numbers [0, 1]. To calculate logic membership functions, new operations were proposed and showed to be areas on able tool to generalize classic logic. Fuzzy logic is nonparametric classification procedure which can deduce nonlinear classification between input and output, and maintain flexibility in decision making even on complex problems. Fuzzy systems provide the means of translating the expert knowledge of humans about the process in terms of fuzzy (IF–THEN) rules. A fuzzy rule is the basic unit to gain knowledge in fuzzy systems.

Like a conventional rule in artificial intelligence, fuzzy system has two components: an ‘if’ part and a ‘then’ part which are also known as antecedent and consequent, respectively [19]. Fuzzy is easy to train and requires less number of samples. However, it gives a lower accuracy rate as compared to any other classifier [30]. Kavdir et al. proposed a method of apple grading in colour, size and defects of apples are extracted. These features were gathered and evaluated using fuzzy system and this gave 89% accuracy in classification [24]. Date fruits were graded using Fuzzy by extracting some features like quality of juice, size and freshness and it gave 86% accuracy [19]. Rokunuzzaman et al. presented an algorithm to classify tomatoes and it gave 84% accuracy rate [29].

B. Artificial Neural Network (ANN)

ANN is massively parallel distributed information processing system that is made up of artificial neurons has certain performance characteristics resembling biological neurons of the human brain [31]. A neural network is characterized by its architecture that represents in which pattern it is connected between nodes and which method is used to determine the connection weights and the activation function. Feed forward hierarchical architecture is the most commonly used ANN.

A typical three-layered feed-forward neural network is made of a multiple elements called nodes, and connection pathways that link them. When an input signal is received at a neuron, then neuron processes it, and sends an output signal to other interconnected neurons [31]. ANN is suitable to analyse complex problems. However, it has some demerits such as it needs large training set and it is a time consuming process. ANN gave high accuracy for training set but inadequate interpolation for testing data [32]. Mustafa et al. proposed a method to determine the size and ripeness of banana. Shape and colour features were extracted. Then ANN was used for classification and it gave accuracy of 79.90% [9]. Rokunuzzaman et al. presented an algorithm to classify tomatoes ANN also and it gave 87.5% accuracy rate which was more than the accuracy rate given by Fuzzy logic [29]. Alpasandi et al. proposed a method in which peach was classified and it gave 99.3% accuracy rate [33]. K-means clustering algorithm is used for labeling each pixel in the image also authors used SURF (Speed up Robust Feature) algorithm for extracting features such as color, structure, texture and morphology. SURF algorithm uses the blob detector and local descriptors.

Blob analysis consists of refinement, extraction and analysis for detecting the scope of interest of given images. Artificial neural network (ANN) is applied for pattern matching and classification of diseases [40].

C. Adaptive Neural Fuzzy Interference System

ANFIS was first introduced by Jang which is capable of approximating any real and continuous function on a compact set to any degree of accuracy. ANFIS is a fusion of artificial neural network and fuzzy interference system. In ANFIS, a fuzzy inference system (FIS) has been constructed, then membership function parameters are tuned using either a back propagation algorithm alone or in combination with a least squares method. This adjustment allows to FIS to train from the extracted data. Normally, an ANFIS is made up of the membership layer, the fuzzification layer, the normalization layer, the defuzzification layer and the fifth layer, the output layer [34]. ANFIS is represented faster than ANN. However,
when the trained parameters are applied to checking data, total error of ANFIS is larger than ANN Nozari et al. presented an algorithm for grading of Mozafatidates and classified these based on weight, length, width and thickness. These fruits were graded using both ANFIS and human experts for comparison and ANFIS showed 93.5% accuracy as compared to human experts [22].

D. Support Vector Machine

In machine learning, support vector machines are supervised learning systems based on the statistical learning theory for data classification and pattern recognition in classification and regression analysis problems.

This method is used for classification of both linear and nonlinear data. In support vector machine model, samples are presented as points in space, and mapped such that the samples of the distinct classes are separated by a clear boundary which is as wide as possible. In this method, an optimal boundary of separating the classes, a kernel function may be used. The kernel function evaluates the relationships within the data and makes complex divisions in the space [20]. SVM has the highest accuracy rate as compared to other techniques [35]. Suresha et al. proposed an Apple classifier using SVM and it gave 100% accuracy when only colour features are compared [20]. Zheng et al. presented an algorithm of mango grading. Fractal dimension and L*a*b* colour model are used for grading purposes. Using SVM, fractal dimension and colour gave 85.19% and 88.89% accuracy respectively [28]. K-means based segmentation is used to detect region of interest to select infected part only in the image. After that features are extracted from the segmented portion of the given images for training and to store in database. Then with the help of support vector machine and feature database, the image will be classified using feature derived from segmented part of the image. Due to the learning involved, this system detects type of diseases with greater accuracy also the system suggest way to prevent the occurrence of the disease and helps farmers by identifying the problem in the fruit [43].

CONCLUSION

In this paper, different image processing based classification techniques for fruit grading and sorting system is reviewed. Machine vision based fruit grading systems are capable of replacing labour work for inspection of fruit grading. Different researchers used algorithms for image segmentation, feature extraction, training and classification of fruit disease. Out of morphological, colour and texture feature, morphological gave highest accuracy rate. In colour model, HIS (Hue, Saturation, Intensity) colour model is commonly used for grading because it is related to human perception. In machine learning techniques, SVM (Support Vector Machine) gave highest accuracy, but ANFIS (Adaptive Neuro Fuzzy Interference System) showed the best result out of these techniques. Further, Fuzzy gave lowest accuracy rate result, but it is easy to implement.

References


