Abstract: The demand for the quick and automated quality assessment system for fresh fruits are increased constantly to avoid the problems with manual classification such as inaccuracy, inconsistency, time consuming and human resource. This research work is focused towards automatic grading of mangoes so that the good quality mangoes can reach the market by the mango traders. Nowadays, these grading of mangoes has been done with the different methods namely, fuzzy logic and neural network. In this paper, an automatic Multiclass support vector machine (SVM) based mango grading system is proposed. The proposed methodology contains the following process, pre-processing, segmentation, feature extraction and then the mangoes are sorted into three grades as very good, good and bad using the multiclass SVM classifier. We achieved an accuracy of about 97%. The performance results prove that proposed system provides good performance in grading when compared to the existing systems such as vision based sorting, artificial neural networks, fuzzy logic, SVM.

Keywords: Multi-class SVM classifier, geometrical, feature extraction, grading

1. Introduction

Mangoes are the fruit which is largely produce in India with the approximation of 63% is formed through India of entire production. Other than India of the major countries where production of the mango can be Philippines, Mexico, Pakistan, Indonesia and China. The availability of the mangoes in the several import markets is constant. Trading of mangoes in winter are done by some countries like Venezuela, Brazil, and Peru whereas trading of mangoes in spring and summer are done by Hatti, Mexico, India and Philippines. The consumption of the mangoes is largely done in its trading company. Though, demands are growing gradually in the zones like Middle East, Japan, Europe and North America [1]. Production of mangoes in India is 18.00 million of tones. There are different types of mangoes according to size, textures and color of a mango. Generally the mangoes which is ripe has a yellow-orange or reddish peel and are juicy on eating, while the unripe mangoes has green peel and unripe mangoes does not have same juicy or flavour as a fresh fruits[2]. Techniques like image processing and machine vision are more and more valuable in the industries of fruit, especially for quality inspection and defect sorting application [3].

The objective of this paper is to grade the mangoes automatically with help of the image processing algorithm and SVM classifier for the classification. The traders of the mangoes can use this system to grade the mangoes and separate them according to the quality automatically. In industry time and efficiency are the major factors that can’t be achieved by the human inspection system because it is difficult for the human to inspect large number of mangoes.

The rest of the paper is organized as: Section II focusses towards review of literature, Section III elaborates the proposed methodology, Section IV and V demonstrates experimental results and comparative study respectively. Finally, conclusions are drawn in Section VI.

2. Review of Literature

The classification techniques play its vital role in classifying various fruits such as grapes, date fruits, mangoes olives for its quality [22]. This results in better grading of fruit. This section mainly explains the recent research on fruits grading system.

The modelling and controlling of the drying process of grapes by applying the artificial neural network and machine vision were used in [4]. There, the experiment was done using the Matlab and the comparison was done
between two different algorithms i.e. multiple linear regression (MLR) and artificial neural network (ANN). ANN based algorithm achieved better results than MLR. For the experiment the sample of seedless sultana grapes of Iran and inverter, SSR relay, D/A transducer, Computer, Camera, Florescent lamp, Fiber glass, Thin layer grapes, Balance, Temperature sensor, Heater, Accumulator fan was used. Image is analysed using image processing and the ANN algorithm i.e. Multi-layered feed forward neural network was used where 5 input, 4 hidden layer i.e. Tansig and 1 output i.e. purlin for the outcomes. Where Tansig \( n = \frac{2}{1 + e^{-(2n)}} - 1 \) and purlin \( n = n \). MLR also performed on the same process of drying grapes and on the basis of the result MFFNN is better than the MLR with hot air drying system.

The Classification of date fruits using the feature of size, shape and description of texture was explained in [5]. First, the color image with different type of date is disintegrated in dissimilar colour modules. The texture descriptor is applied to encode the best features and then FDR (Fisher Discriminator Ratio) applied to reduce the features. Finally, for classification SVM is used. [6]

A model proposed in [7], to see the broadness of the olives. By investigating the RGB photographs of the olive tree the certification has been performed. This data could be helpful for surveying the best gathering time. To check the likelihood of package of strata histogram examinations from an olive tree picture has been utilized and the information set under investigation for the ordinariness. Feedforward neural system is used with backpropagation learning estimation. Regardless neural structure is unfathomable with 97% right assembling and second system achieves execution of 88.8%. This model is used to scorn the hugeness of the olive. A key variable for Figureuring the vastness of the olives is the equilateral breadth and it is not truly influenced when olives are transversal to the photograph.

The model proposed in [8], Region of Interest (ROI) is partitioned from the data using Otsu Method and shape exhibiting is performed on it. Territory based descriptor is used on the segmented image, where huge centre point length and minor centre point length isolated. The breaking point based descriptor is used and the edge of the divided things is registered. The gathering has been done with the Bayes classifier using Bayes rule i.e. Posterior = (prior*likelihood)/confirmation. The system has been attempted three changes of mango and is found 90% exactness. [18, 19]

Fruit skin defect identification system using Machine vision is proposed in [9]. Here, from the input image patch local feature is extracted using the colour histogram. The linear SVM is trained with this features and same is used to identify the defect of the fruits skin. They achieved 96.7% accuracy and 1.7% false slip away distinguished.

Fleecy Figureuring for Automatic Fruit Grading structure using machine vision system according to their maturity level and quality is proposed in [10]. Fuzzy rule based algorithm is used to inspecting the four assessments like poor (G1), medium (G2), extraordinary (G3) and awesome (G4) in perspective of improvement level and quality. This framework saw to be brisk, ease and furthermore canny. [17] Mango's sorting and highlight extraction using 2D/3D Vision-Based it can isolate 2D and 3D visual properties of mango, for instance, size (length, width, and thickness), foreseen extent, volume, and surface zone from use them in sorting. The photos are at first divided to expel the layout districts of mango. To begin with the cameras are changed in accordance with get the normal and outward camera parameters and after that the 3Dvolume voxels are made in light of layout photos of the natural item in different points of view. In the wake of craving it is get the coarse 3D, volume and surface locale of natural items. For sorting of mangoes, the neural framework is used. The segments are described and discovered for physical properties. These fuse foreseen zone (A), length (L), width (W), thickness (T), 3D volume (V), and 3D surface extent (S) [11].

3. Proposed Methodology

In this section, the proposed system for the classification of mangoes is explained in detail. It is shown in Figure 1. It is explained as, the input mango image is pre-processed to remove the noises and improve the quality of the image. Then, the enhanced mango image is segmented from the background. The features such as histogram, texture and geometrical are extracted from the resultant segmented image. Later, multi-class SVM is trained with the extracted features and trained SVM is used to classify and grade the mango as Very Good, Good or Bad.

3.1 Pre-processing

Image processing is done for manipulation of image data and the task which can influence is distributed into categories of three i.e. understanding, analysis, and processing done on image. Pre-processing of an image is done by the median filtering to reduce the noises and sharpen contrast, highlight contour, edge detection by using sobel edge detector operation [12]. The median filtered output is shown in Figure 1.
3.2 Segmentation

The segmentation has been done with the help of the thresholding algorithm [20] and then the boundary has been extracted. Then the features are extracted from the segmented image for classification of the mango. The segmented output is shown in Figure. 3

3.3 Feature Extraction

Features which are extracted comprise of the appropriate information from the data which has been input and then the desired task with the help of using data which are reduced in its place of comprehensive original data. Extraction of features is used for reducing the resource in the suitable amount from the set of large data. Image processing is the very important area of the feature extraction. In which algorithm are used to detect various desire portion of the image. Feature extraction is done so that the feature from the object can be extracted to use in the Multi-class SVM. Features are extracted with respect to various aspects namely, Geometrical extraction of feature, texture extraction of feature, colour extraction of feature, and histogram extraction of feature.

3.4 Geometrical feature extraction

In geometrical feature extraction first extraction of the boundary has been done. In the boundary extraction the binary image has been eroded and subtracted from the grey scale [13]. It is shown in Figure.4. After the extraction of the boundary, the regional properties of the mango has been extracted on the aspect of area (A), eccentricity (E), perimeter (P), major axis length (Maj), minor axis length (M) and orientation (O) and it is tabulated in Table. 1. Geometric feature is a technique for the combination of machine learning and computer vision to solve the visual task. Genetic algorithm has been use to learn the feature. This technique is widely used in the artificial intelligence. [14]
3.5 Histogram Feature Extraction

Statistical description is in image processing is commonly used. Standard deviation is unbiased estimation of the brightness within the region is called simple standard deviation. The other histogram features are mean, skewness and kurtosis. It is defined as,

\[
\mu = \frac{1}{N \times M} \sum_{i=1}^{N} \sum_{j=1}^{M} p(i, j)
\]

(1)

\[
\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} [p(i, j) - \mu]^2}
\]

(2)

\[
\text{Skewness} = \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} \left( \frac{p(i, j) - \mu}{\sigma} \right)^3
\]

(3)

\[
\text{Kurtosis} = \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} \left( \frac{p(i, j) - \mu}{\sigma} \right)^4
\]

(4)

where, \(p(i, j)\) is the intensity value of the pixel at the point \((i, j)\). \(M \times N\) is the size of the image. It is tabulated in Table 2.

The corresponding intensity graph is shown in Figure 5.

3.6 Texture Extraction

Texture extraction has been used to find the smooth, coarse or periodic of the mango image. The texture is
find using Max Probability (M), correlation (CO), contrast (C), energy (E), homogeneity (H), entropy (EN).

\[
\text{Energy} = \sum_{i=1}^{N} \sum_{j=1}^{M} p^2(i, j) \tag{5}
\]

\[
\text{Contrast} = \sum_{i=1}^{N} \sum_{j=1}^{M} (i - j)^2 \cdot p(i, j) \tag{6}
\]

\[
\text{Correlation} = n = \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{p(i, j) - \mu_i \mu_j}{\sigma_i \sigma_j} \tag{7}
\]

\begin{align*}
\text{Homogeneity} & = \sum_{i=1}^{N} \sum_{j=1}^{M} \frac{p(i, j)}{1 + |i - j|} \tag{8} \\
\text{Entropy} & = -\sum_{i=0}^{N} \sum_{j=0}^{N} p(i, j) \cdot \log p(i, j) \tag{9}
\end{align*}

where, \(p(i, j)\) is the intensity value of the pixel at the point \((i, j)\). \(M \times N\) is the size of the image.

<table>
<thead>
<tr>
<th>E</th>
<th>Autocorrelation</th>
<th>Contrast</th>
<th>Correlation</th>
<th>Cluster Prominence</th>
<th>Cluster Shade</th>
</tr>
</thead>
<tbody>
<tr>
<td>5923</td>
<td>0.6311</td>
<td>512.7845</td>
<td>397.7776</td>
<td>1.34E+03</td>
<td>0.0975</td>
</tr>
<tr>
<td>6343</td>
<td>0.5089</td>
<td>423.0178</td>
<td>364.1446</td>
<td>1211.526</td>
<td>0.117846</td>
</tr>
<tr>
<td>5594</td>
<td>0.599929</td>
<td>523.9159</td>
<td>419.1605</td>
<td>1388.462</td>
<td>0.082127</td>
</tr>
<tr>
<td>4794</td>
<td>0.554118</td>
<td>553.3291</td>
<td>460.6122</td>
<td>1376.178</td>
<td>0.068158</td>
</tr>
<tr>
<td>5256</td>
<td>0.487423</td>
<td>486.6586</td>
<td>424.9336</td>
<td>1212.254</td>
<td>0.069439</td>
</tr>
</tbody>
</table>

### 3.7 Feature Selection using Firefly optimization

The firefly algorithm (FA) is a nature-inspired metaheuristic optimization algorithm which is inspired by the fireflies' flashing behaviour. The crucial purpose for a firefly's flash is to act as a signal system to attract other fireflies. Xin-She Yang [24] formulated this firefly algorithm using following rules:

1. All fireflies are unisexual, so that any individual firefly will be attracted to all other fireflies;
2. The brightness of a firefly is determined by the encoded objective function.
3. Attractiveness is proportional to their brightness, and for any two flashing fireflies, the less bright one will be attracted by the brighter one; however, the intensity decrease as their mutual distance increases; If there are no fireflies brighter than a given firefly, it will move randomly.

In general, the attractiveness of a firefly is determined by its brightness which in turn is proportional to the encoded objective function. The objective used here to select the extracted feature is,

\[
F_i = \frac{(FP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) (TP + FN) (TN + FP) (TN + FN)}} \tag{10}
\]

where, \(TP\) – True positive, \(TN\) – True negative
\(FP\) – False positive, and \(FN\) – False Negative

To obtain the optimized features, the extracted features of histogram namely, mean, standard deviation, skew, kurtosis, geometrical features such as area (A), eccentricity (E), perimeter (P), major axis length (Maj), minor axis length (M) and orientation (O), and texture features namely, autocorrelation, contrast, correlation, cluster prominence, cluster shade, energy. These 16 features are given as initial solution to the firefly optimization algorithm. The optimization algorithm is used to select the features which are appropriate for classification using the fitness function as given in equation 10. First, random selection of features is done and evaluates the fitness function. The fitness function is based on the Mathew Correlation Coefficient (MCC), which performs the correlation between the True positive rate and False positive rate. The MCC value varies between -1 to +1 [23]. This process of computing the fitness function is repeated until the decided accuracy has been obtained with optimized features. As a result, this further improves the selection of features for classification. For our dataset, the optimized features determined by the proposed cuckoo search optimization algorithm are given as, area, convex area, energy, contrast, entropy, mean.

### 3.8 Classification

The training points which are belongs to any of the one from N different classes and its aim is to make a function from the new specified record which in future can be sad right the class where the record belongs. After the extraction of the features grading of mangoes has been performed by the Multi-class SVM. Once the SVM is trained, it is fit to use with the actual data set that is to be classified. The SVM classifier will read all the
recorded data features and classify them according to its knowledge that it has after training. With the help of features extracted, the grading has been done in the three ways: Very good, good and bad. Multi-class SVM has been taken as it can classify in more than two class and fast [15].

The classification of the mango has been done with the help of the multi-class SVM where the data has been trained, label and test so as to classify the images of the mango and grading accordingly. The Multi-class classifier reduce the train data and label which is use to give the result according to the classes here three classes has been taken very good, good and bad. And the final result has been come with the decision making of the classes the image of the mango basically belong to which class.

The classes are label as if 0 is the result than it is very good, if 1 is the result than it is good and if 2 is the result than it is bad. It is shown using Figure. 6

![Figure 6. Grading of mangoes in three categories (a) Very Good, (b) Good, and (c) Bad](image)

4. Simulation Results

The performance of the proposed system is tested with the database of 100 mango images which is collected from the website of the COFILAB team [16]. In our experimental result analysis 80% images used for training and 20% images used for testing. The same is compared with the existing method in terms of the specificity, sensitivity and accuracy. The sample database images are shown in Figure. 7

![Figure 7. Sample Database Images](image)

The quality of the mangoes are classified accurately in the proposed method using Multi-class SVM and it is experimented using the Matlab and achieves about 97% accuracy and capable for more than two class classification which in result has classified in three categories Very good, good and bad. It is shown in Figure. 7

![Figure 8. GUI Result grading mango as (a) very good (b) good (c) bad](image)

5. Comparative Study

To evaluate the performance of the proposed system, here in this section, we compared the proposed systems with the related conventional systems and the results are tabulated in table. 4. It is also shown in Figure.9. Here performance of the proposed system is compared in terms of the specificity, sensitivity and accuracy with the existing systems.
Table 4. Comparative Analysis

<table>
<thead>
<tr>
<th>Technique</th>
<th>Classifier</th>
<th>Feature Extraction</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>[4]</td>
<td>MFFNN (Multi-Layered feed forward Neural Network)</td>
<td>Geometrical feature extraction, RGB space</td>
<td>90.45%</td>
<td>85%</td>
<td>90%</td>
</tr>
<tr>
<td>[5]</td>
<td>Local Texture descriptor, classifier</td>
<td>Shape and size feature</td>
<td>98.45%</td>
<td>90%</td>
<td>98%</td>
</tr>
<tr>
<td>[6]</td>
<td>Artificial Neural Network</td>
<td>Histogram feature and geometrical feature and color</td>
<td>98%</td>
<td>89%</td>
<td>96.47%</td>
</tr>
<tr>
<td>[7]</td>
<td>Neural Network</td>
<td>Histogram feature, RGB color, geometrical feature</td>
<td>89%</td>
<td>87%</td>
<td>88%</td>
</tr>
<tr>
<td>[8][18]</td>
<td>Fuzzy Logic</td>
<td>Geometrical, color, texture feature</td>
<td>91.45%</td>
<td>86%</td>
<td>90%</td>
</tr>
<tr>
<td>[9]</td>
<td>SVM</td>
<td>Color, histogram feature</td>
<td>98%</td>
<td>91%</td>
<td>96.7%</td>
</tr>
<tr>
<td>[10][17]</td>
<td>Fuzzy Logic</td>
<td>Boundary tracing, texture feature</td>
<td>88%</td>
<td>80%</td>
<td>86%</td>
</tr>
<tr>
<td>[11]</td>
<td>Vision based sorting</td>
<td>Geometrical</td>
<td>98%</td>
<td>91%</td>
<td>96.47%</td>
</tr>
<tr>
<td>Proposed Approach</td>
<td>Multi-class SVM</td>
<td>Geometrical feature, texture and histogram feature</td>
<td>97.28%</td>
<td>93%</td>
<td>97%</td>
</tr>
</tbody>
</table>

From the table, 4 and Figure 9, we infer that the results were found to be good when compared to the existing conventional systems such as such as vision based sorting, artificial neural networks, fuzzy logic, SVM.

![Comparison plot of the various Approach](image)

Figure 9. Comparison plot of the various Approach

6. Conclusion

In this research work, machine vision based automatic grading system for mango fruit has been presented. The Multi-class SVM classifier has been applied on the database of 100 still images, 80% images for training and 20% images testing has been done. We achieved an accuracy of about 97%. On the basis of the result and comparison with the existing method such as such as vision based sorting, artificial neural networks, fuzzy logic, SVM, it is concluded that the proposed algorithm is best suitable for the Classification and grading of the mango. The future work will focus on the optimized system for grading of mangoes by incorporating optimization algorithm in the feature selection part for improving the accuracy.

References


[16] CoFILAB Team Website www.Cofilab.com


