Pertanika J. Sci. & Technol. 27 (4): 1693 - 1708 (2019)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

# Hybrid Feature Extraction and Machine Learning Approach for Fruits and Vegetable Classification

**Nimratveer Kaur Bahia[1], Rajneesh Rani[1]*, Aman Kamboj[1] and Deepti Kakkar[2]**

[1]*Department of Computer Science and Engineering, Dr. B. R. Ambedkar National Institute of Technology Jalandhar, Jalandhar, Punjab 144011, India*

[2]*Department of Electronic and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology Jalandhar, Jalandhar, Punjab 144011, India*

## ABSTRACT

Manual fruits and vegetables detection become easy when it is done in small amount, but it is a tedious process and more labor is required when gigantic amount is considered. So, automatic detection of these comes into usage. This study took the images of fruits and vegetables as input to the very first stage of processing from where detection was done. The entire process constituted three stages: Background subtraction, extraction of color as well as texture features, and then classification. Background subtraction was performed using k mean clustering technique. Color features were identified using statistical features. To identify texture features Histogram of Oriented Gradient (HOG), Local Binary Pattern (LBP) and Gray Level Co-occurrence Matrix (GLOM) were used. For training and classification, Support Vector Machine (SVM) classifier had been used and performance of this classifier had been compared with K Nearest Neighbor (KNN) classifier. After comparing the results, it shows that accuracy of SVM was higher than that of KNN. The accuracy obtained by SVM with quadratic kernel function was 94.3%.

*Keywords:* Color, gray level co-occurrence matrix, K mean clustering, K nearest neighbor, support vector machine, texture

ARTICLE INFO

*Article history:*
Received: 23 October 2018
Accepted: 04 June 2019
Published: 21 October 2019

*E-mail addresses*:
nimratbahia@gmail.com (Nimratveer Kaur Bahia)
ranir@nitj.ac.in (Rajneesh Rani)
amankambojj@gmail.com (Aman Kamboj)
kakkard@nitj.ac.in (Deepti Kakkar)
* Corresponding author

## INTRODUCTION

Detection system is a 'magnificent challenge' posed to the computer vision for attaining recognition of the near human levels. An object can be very well detected using image processing techniques. There exists abundance of techniques used for the objects

ISSN: 0128-7680
e-ISSN: 2231-8526

detection but still lots of improvements are required to achieve high accuracy. Content based image retrieval can be applied for the object detection. It detects the image's contents that include color, texture, shape and size (Rocha et al., 2008). There are many applications of this work when used in the form of a mobile application. It helps the cashier during the billing and useful for the kids to increase their learning power (Zawbaa et al., 2014). The detection of fruits as well as vegetables is useful at places like supermarkets where cost involved for fruits purchased by any customer can be done automatically. Fruits as well as vegetables recognition may be brought into use in computer visualization for the automatic categorization of fruits and vegetables from any random badge, containing various fruits. The aim of this section is to represent a comprehensive overview of research development activities in the field of automatic fruit and vegetable classification.

Seng and Mirisaee (2009) had used the combination of color, size and shape features with KNN classifier to classify fruits and vegetables. The experiments done by them were on only seven categories and accuracy obtained was 90%. Ninawe and Pandey (2014) studied the fruit detection completion system by doing the experiments on a dataset of six varieties of fruit which contained red apple, green banana, green guava, green melon, orange and watermelon. They used four features: color, shape, texture, size and then used their combination for obtaining the better results. They used the geometrical properties to calculate the area as well as perimeter and computed roundness and entropy values for extracting different features. The classification was done using the KNN classifier. This method has a drawback as shape of each fruit or vegetable is different so we cannot recognize it on the bases of shape and size as it decreases the accuracy. Arivazhagan et al. (2010) had tried to do recognition of fruits and vegetables using minimum distance classifier based upon the statistical and co-occurrence features derived from the Wavelet transformed sub-bands. They achieved an accuracy of 87% approximately by doing experiments on a database of about 2635 fruits from 15 different classes. In this method they used 50% of images as training to maintain the high accuracy by using less number of color and texture features. Zhang and Wu (2012) had proposed a novel classification method based on a multi-class kernel support vector machine (kSVM) with the desirable goal of accurate and fast classification of fruits. The experimental results demonstrated that the Max-Wins-Voting SVM with Gaussian Radial Basis kernel achieved the best classification accuracy of 88.2%. In this method they had used 60% of images as training to achieve the high accuracy, but results were more accurate if training and testing images were in equal proportion. Dubey and Jalal (2015) proposed the method that recognized various numbers of different fruits as well as vegetables. They had extracted different color along with texture features. Global color histogram (GCH), Color coherence vector (CCV), Color difference histogram (CDH) had been used to extract color features and Structure Element Histogram (SEH), Completed Local Binary pattern (CLTP), Local Binary Pattern

(LBP) and Local ternary pattern (LTP) had been used to extract the texture features. For classification, they had used Multiclass SVM classifier for training and testing purpose and achieved 93.84% accuracy by doing experiments on a dataset of 2312 images. In this they used different methods for feature extraction but some of the methods did not help gain the accuracy. Jhawar (2016) sorted the orange fruit by using a pattern recognition technique. They took the 160 images which were gathered from various locations in Vidarbha area of Maharashtra. Only four features were extracted using maturity level. For doing the classification, three techniques were used: one was Edited Multi seed Nearest Neighbor Technique, second was linear regression technique while the third was nearest prototype. All were based on the technique of pattern recognition. The maximum accuracy obtained was 97.98% by using linear Regression classifier. The techniques used in this paper gave high accuracy, but they used very few numbers of images so, there was a need to test the method on large dataset to get accurate results. Shukla and Desai (2016) proposed the model which used machine learning for the automated recognition of fruit. In this, they used the color, texture, shape features followed by their combination for obtaining the better results. Color Coherence Vector (CCV) was used to extract the color features and GLCM and LBP techniques were used for texture features extraction and some statistical features were used for shape analysis. Two classifiers were used: KNN and SVM for the classification and the results were compared. The best accuracy obtained was 91.3% with KNN whereas by using the SVM classifier the accuracy was 86.96%. The technique is tested on very few images. Moreover, the images contain only one category of fruit. So, it is easy to detect fruits and vegetables in the images. Moallem et al. (2017) conducted a study which showed the grading of the golden delicious apple. To do this, they used different techniques. Stem end detection, clays detection, primary defect segmentation, refinement of defect regions was used to do the segmentation. After the segmentation was complete, they detected the defected regions corresponding to an image. After this statistical, textural and geometric feature were extracted. On finding the feature vector, various classifiers are applied. SVM, KNN, and MLP (multi-layer perceptron) classifiers were used. After classifying they divided the fruit into healthy and detected part. After that, healthy parts were further classified into first rank and second rank. SVM Classifier outperformed the other two type of classifiers with recognition rate of 92.5% and 89.2% for two categories. In this study, the author used only one type of fruit, which did not match with other fruits and vegetables due to this the method had achieved good accuracy. Another study by Wang et al. (2018) was on comfortable footwear design for patients with diabetic conditions. The features like HSV and HOG and GLCM were extracted and fed to the Fuzzy Support Vector Machine (FSVM) for the training of diabetic plantar pressure images. Their proposed system obtained an accuracy of 84.3% which is much better than in comparisons to SVM and LSVM. Ansari and Ghrera (2018) suggested a novel intuitionistic fuzzy feature extraction method to extract the local texture, in this method the authors had incorporated a new intuitionistic fuzzy set theory for the representation of pattern in the images. Their proposed method also contributes to more than one bin in the distribution, which was used as a feature vector. The proposed approach had shown much better results over the local binary pattern.

From the study of literature, it has been concluded that there is requirement to explore more state-of-the-art feature extraction and classification techniques to detect fruits and vegetables as majority of the methods used in literature have been tested on small datasets and images were having one type of fruit. So, in this paper we have utilized texture, color and hybrid features (color + texture) of an image for fruits and vegetable classification and compared their results using SVM and KNN classifier. Our method is evaluated on a large dataset and each image in the dataset contains multiple instances of the fruit or vegetable. The rest of paper is planned as follows: Section 2 explains the proposed methodology and the dataset used is implicated in the proposed work. Section 3 describes the various results obtained including the dataset of raw images obtained for research and compares the results with earlier approaches. Section 4 gives the conclusion and some future direction.

## MATIRAL AND METHODS

### Dataset

The dataset of fruits and vegetables used in this presented work is the same as used by Dubey and Jalal (2015). It consisted of 15 categories: Spanish Pear (159), Asterix Potato (182), Cashew (210), Nectarine (247),Plum (264), Onion (75), Granny-smith Apple (155), Orange (103), Tahiti Lime (105), Kiwi (151), Fuji Apple (212), Watermelon (192), Diamond Peach (211), Agata Potato (201) and Honeydew Melon (145): total of 2612 images. All this data was collected from supermarket where there were different types and varieties of fruits and vegetables. Figure 1 represents the dataset of different kind of Fruits and Vegetables (Dataset is accessible at – http://www.ic.unicamp.br/~rocha/pub/downloads/tropical-fruits-DB-1024x768.tar.gz). The Dataset contains more than one fruit in one image.

### Methodology

The methodology used in the present work is described in this section as shown in Figure 2. The set of raw images had been first passed through K mean clustering technique for background subtraction and then various color and texture features had been extracted from the image as a feature vector and SVM classifier had been trained for the classification purpose.

During the training phase, both raw images (fruits and vegetables) and their corresponding labels were fed to the classifier. At the time of testing the same procedure was followed and classifier returned the labels corresponding to the input image of fruit or vegetable.

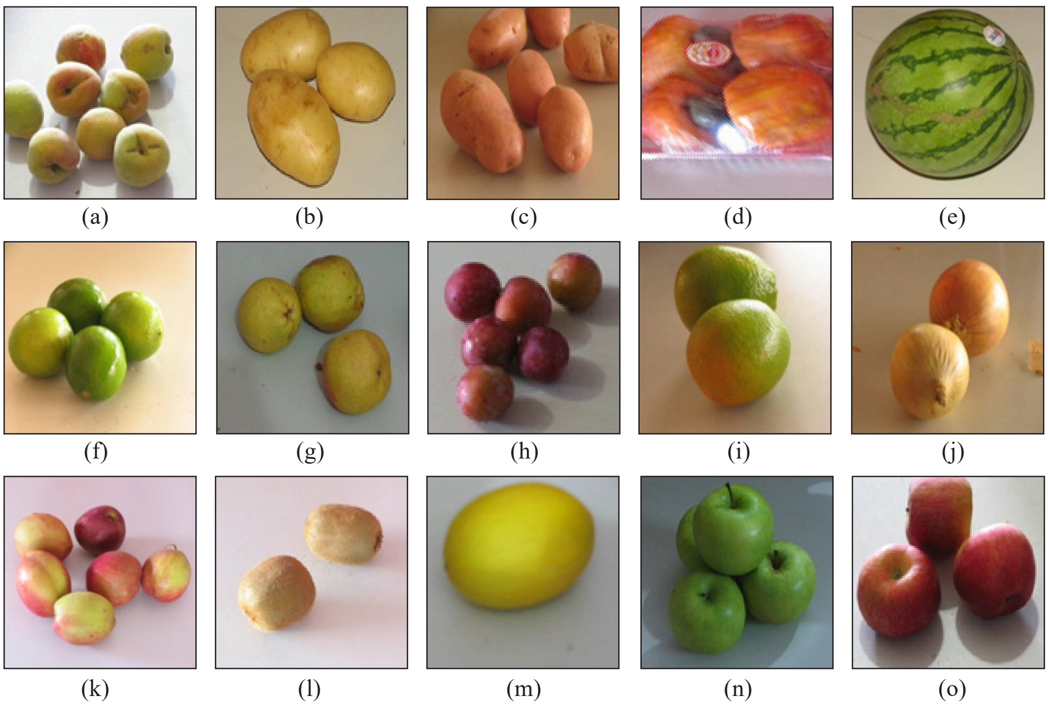



*Figure 1*. Data set of 15 different kinds of fruit and vegetable. (a) Diamond peach; (b) Agata Potato; (c) Asterix Potato; (d) Cashew; (e) Watermelon; (f) Tahiti Lime; (g) Spanish Pear; (h) Plum; (i) Oranges; (j) Onions; (k) Nectarine; (l) Kiwi; (m) Honeydew melon; (n) Granny Smith Apple; and (o) Fuji Apple

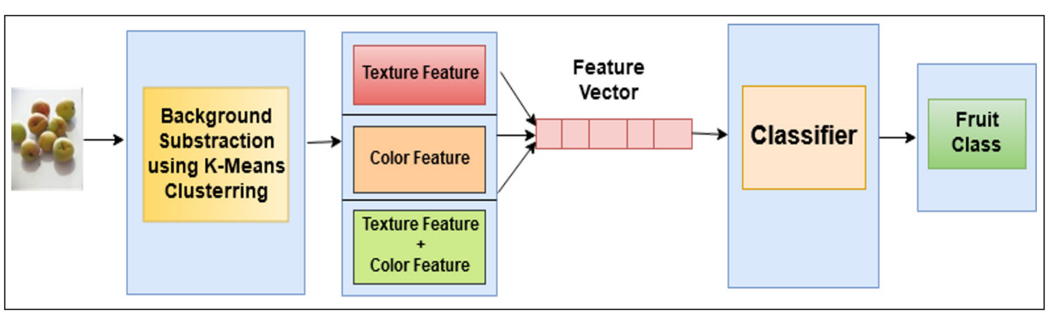



*Figure 2.* Flow chart of the proposed methodology

**Background Subtraction.** For performance enhancement of the proposed approach, supermarket dataset of 15 different categories of fruits along with vegetables was used. These images were used as the input. As the dataset used in the present work had been gathered from the supermarket so the data contained noisy or blurred images. For this type of raw data, background subtraction was made for extraction of the region of interest from the images. Background subtraction is the technique in which only foreground object is displayed and rest all another object which comes in the background is black. This will remove the background and extract foreground.

In the present work, for background subtraction, K-Mean Clustering technique had been used. The image was divided into k segments. K was taken as 2 because we had to divide the image into 2 segments in one part only the fruit or the vegetable would be shown and rest all the background part would be black. Figure 3 and Figure 4 show an example of K-Mean Clustering.

(a) (b) (c)

*Figure 3.* Before K-Mean Segmentation

(a) (b) (c)

*Figure 4*. After K-Mean Segmentation

The Figure 3 shows the images collected from supermarkets. In these figures, images have unwanted objects. In Figure 3(a) the image has a shadow of the fruit and in Figure 3(c) hand is captured in the image. Images after the preprocessing are displayed in Figure 4. After doing the k means clustering, we removed all the types of noise present in the images which got captured from the supermarket.

**Feature Extraction.** The data obtained after the background subtraction contains all the information that is required for extracting the desired results. The main motive of feature extraction is to acquire the most appropriate information from the data and constitute that information in a lower dimensionality space. When the input data to an algorithm is very large to be operated and it is imagined to be unnecessary then input data is turned into a few sets of features. Feature Extraction is the conversion of input images into different features suggested by (Zaitoun & Aqel, 2015). Many fruit detection and classification systems are based on color features along with shape features. But there are some types of fruits which have the samecolor and shape features. In the present work, extraction has been done by using two kinds of features: color and texture. Shape features are not used, because, in our dataset, there are more than one fruit, which is present in one image and cannot define its shape.

***Color Feature Extraction.*** Color is the best feature to distinguish between fruits and vegetables. Because we see that almost all fruits and vegetables are of different colors, so we can easily distinguish. The various color spaces exist like RGB (Red, Green, and Blue) and HSI (Hue, Saturation and Intensity) and many more. Each color space has its own importance and provides color information in a more intuitive way. So, in the presented work both RGB and HSI color spaces had been used to extract the color features. After doing background subtraction image was in RGB color space. Six Color features had been extracted by using RGB color space: Kurtosis, standard deviation and skewness for the given RGB image and mean of all components that was mean of Red, Green and Blue component of the image. Three features were extracted after converting the image into HIS color space. Mean values of hue, saturation and intensity were calculated. Thus, a total of nine color features had been extracted.

***Texture Feature Extraction.*** The texture is an important feature for detection of any object or to recognize some part from an image. A single pixel can never reveal the texture of the surface of an object. When there are a group of different pixels, then it becomes the texture element (Abdelmounaime and Dong-Chen, 2013; Lalibertea and Rango, 2008). The texture is calculated by the outer part of an object which measures the roughness, coarseness, and smoothness (Reddy et al., 2009). The neighborhood of an image is spatially distributed and specifies its texture (Pujari et al., 2013; Clausi, 2002). In this work, seven textural features had been extracted: five features such as contrast, correlation, homogeneity, energy, entropy with the help of GLCM while two other features were extracted with another two techniques HOG and LBP.

***GLCM (Gray Level Co-occurrence Matrix).*** The textural feature is a type of gray-tone special dependencies which is helpful for the recognition of an image. The gray level matrix consists of a two-dimensional histogram, which is divided by a permanent spatial relationship. It is the statistical method that gives the spatial correlation of pixels. These can be calculated by first converting the RGB to gray-scale image (Sonka et al., 2014). In the proposed method first GLCM had been created, and then different statistics were calculated, which gave the information about the texture of an image. Different statistics used in this work are Entropy, Homogeneity, Contrast, Correlation, and Energy.

***LBP (Local Binary Pattern).*** Prakasa (2016) worked with LBP technique for texture feature extraction. LBP extracts the surface of the image. In this, the texture pattern probability was computed and represented into histogram. They concluded the two type of LBP; one was symmetric and other was natural. For pattern classification, the LBP texture features could also be used. It is a very effective visual distributor technique for extracting the texture feature (Shukla & Desai, 2016). LBP is calculated by equating the adjacent pixels of an image (Dubey & Jalal, 2015). Ansari and Ghrera (2016) had proposed intuitionistic fuzzy local binary for extracting texture features from an image, suggesting extended fuzzy local binary pattern by incorporating intuitionistic fuzzy local binary pattern. The method was applied on various images and results indicated the effectiveness of the proposed method. The studies indicated that an LBP was a great texture feature used for categorization of the objects. In the present work, after extracting the LBP for each pixel, a histogram was created which represented the texture and mean of the histogram for an image was calculated and was labeled as LBP feature.

***HOG (Histogram of oriented gradients).*** HOG utilizes overlapping local contrast normalization, while gets calculated on an evenly spaced cells' dense grid for enhanced accuracy (Dalal & Triggs, 2005). This texture feature had been calculated by taking the mean of luminance gradient component of each pixel of an image after converting it into gray scale image.

### Classification

The process of classification with the help of feature vectors aids in the detection of fruits and vegetables. It defines boundaries between special targets in feature space with the help of extracted features as independent variables. Recent research has used a variety of machine learning models for example, KNN, SVM, decision trees and Neural Networks (NN) and their variants for this purpose. Linear and non-linear hyperdimensional data can be classified with the SVM which is a non-linear mapping of data with the help of kernel functions. KNN is an instance based non-parametric similarity measure learning for data of infinite dimensions and a decision tree is a probability-based graph for multi-class classification. SVM and KNN have been widely used for fruit and vegetable classification and a comparable classification effectiveness with respect to Multi-layer Perceptron (MLP) and Radial Bias Functions (RBF) has been reported (Hameed et al., 2018). That is why, in the present work, experiments had been done using Support Vector Machine (SVM) classifier and results had been compared with KNN classifier.

**Support Vector Machine (SVM).** The SVM is a machine learning tool for the use of data classification. SVM is used as the classification tool as it is a multiclass classifier with good accuracy and has got the ability to find a hyper plane with the widest margin that divides the samples into classes using kernel functions. There is one target value and various attributes for each instance in the training set (Prakash et al., 2012). The aim of the SVM is designing and producing such a model that will be able to predict the target value of data instances in the testing set where only attributes are being provided.

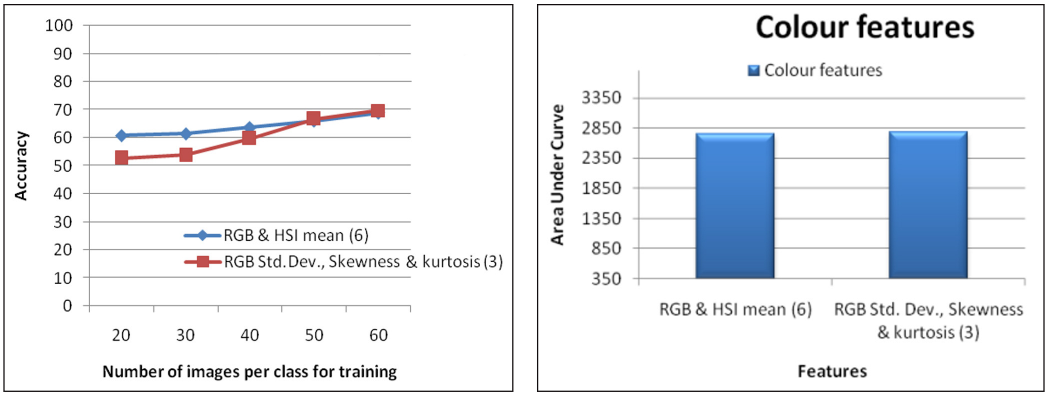
**KNN (K Nearest Neighbor).** KNN is the simplest classifier among all the other classifiers. It is the non-parametric method which is used for regression as well as for classifier. It does the classification on the basis of distance by measuring the distance matrix (Teoh et al., 2004). In this, K is the value used to make the boundaries of each and every class. When the value of K increases or decreases it affects the boundary values of class and error rate. When the value of K is 1 then the error rate is zero for the training sample. As the value of k increases, the boundary becomes smoother and error rate also increases.

## RESULTS AND DISCUSSIONS

The study compares the various color features and texture features and then uses a combination of both, to find optimized features for the classification of fruits and vegetables. The experiments had been done using a different number of training images per class (20, 30, 40, 50 and 60) and the rest of the images had been taken as test images.

### Analysis of Color Features

Figure 5(a) shows the average accuracy of fruits and vegetables using a different combination of color features. First, six color features which are the statistic mean a featurethat is red mean, blue mean, green mean, hue mean, saturation means, brightness mean. Second, a combination of color features: standard deviation, skewness, the kurtosis of RGB image has been used. Quadratic SVM is used to compute the accuracy.




(a) (b)

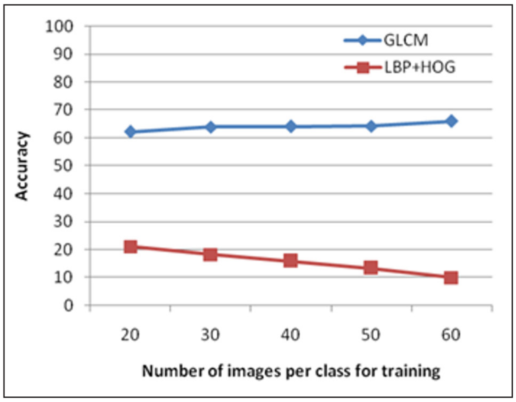
*Figure 5.* Average accuracy and area under curve of color features

Figure 5(b) shows the area under the curve for the same combination of color features. From Figure 5(a), accuracy increases as the number of training images per class are increased and maximum accuracy obtained is approximately 70%.

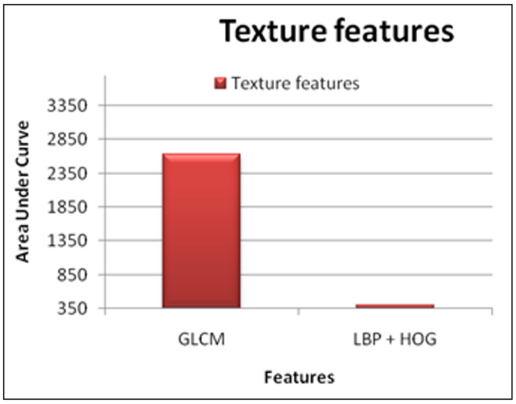
## Analysis of Texture Features

The experiments had been done using a different combination of texture features. First, the combination of two texture features LBP and HOG and the second combination of five statistical GLCM features had been considered for experimentation. Statistical features of GLCM are energy, homogeneity, contrast, entropy, correlation. Average accuracy had been calculated by using Quadratic SVM and is represented in Figure 6(a). The area under the curve was also calculated for different features and is represented in Figure 6(b).

From the figures 6(a) and 6(b), it depicts that GLCM features perform very well than other two types of features in all cases.




(a)



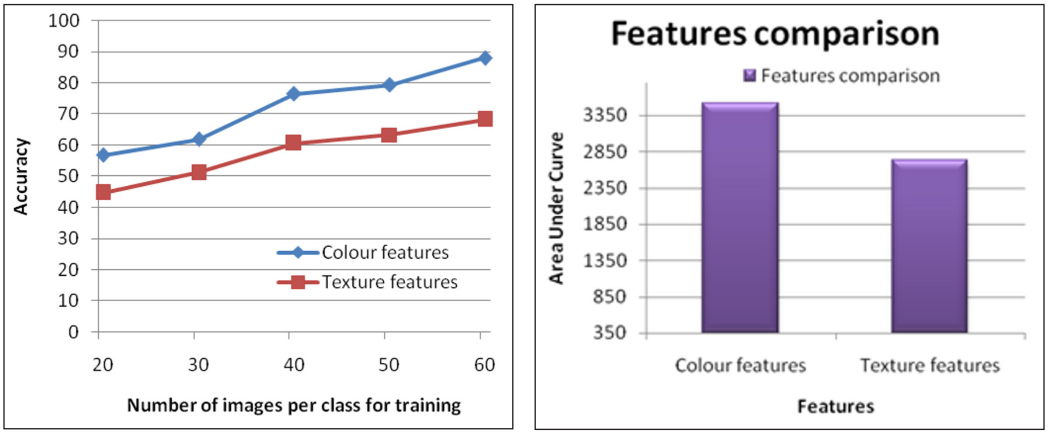

(b)

*Figure 6*. Average accuracy and area under curve of texture features

## Analysis of Hybrid features (Combination of Texture and Color features)

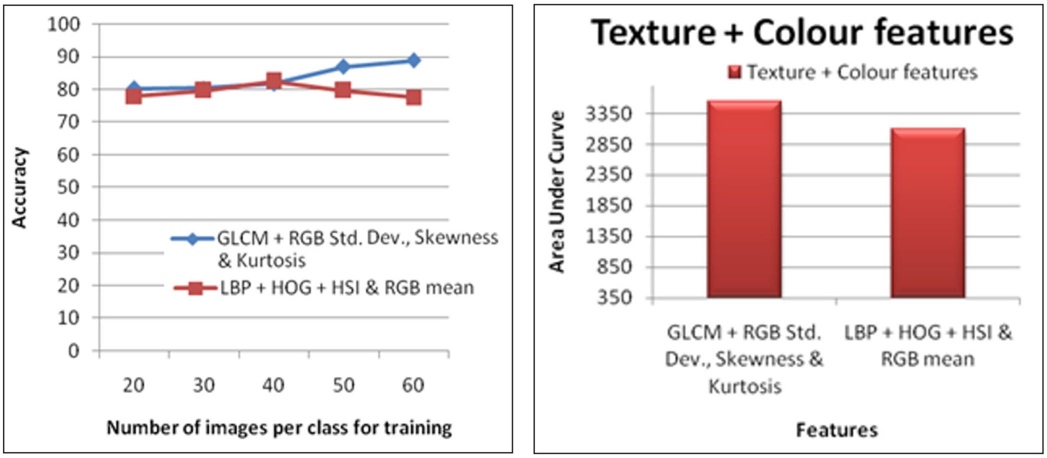
Different experiments had been done by using different combinations of color and texture features. In the first experiment, all the color features together and all texture features together had been considered and computed the accuracy and area under the curve by using Quadratic SVM. Results are shown in Figure 7.

In the second experiment, a combination of some texture feature and some color feature had been used to find its corresponding accuracies. Combination of GLCM texture features with some color features which were statistical mean of RGB and HSI. HOG and LBP texture features were combined with some other color features which were skewness, standard deviation, and kurtosis of RGB image. Figure 8 shows the average accuracy by using a different number of images per class for training and area under the curve when we used color and texture features.




(a) (b)

*Figure 7*. Average Accuracy and Area under curve of color and texture feature




(a) (b)

*Figure 8.* Average Accuracy and area under curve when using combination of texture and color features

In the third experiment we evaluated all the sixteen features using SVM classifier to compare its results with KNN. The Table 1 shows the testing accuracy of SVM classifier with different kernel functions and Table 2 shows the testing accuracy of KNN classifier of its different types.

The comparison of KNN classifier with SVM classifier is shown in Figure 9. After comparing, the best result was generated with the SVM model with Quadratic kernel function and one-against-all strategy with a testing accuracy of 94.3%. While for KNN it gave the best accuracy of 74.66%. We conclude that SVM is better classifier to classify the fruits and vegetables than KNN classifier. This may be quadratic kernel function used in the SVM classifier which clearly separates the boundaries of different fruit and vegetables.

Table 1
*Comparison of SVM and their corresponding accuracies*

| Sr. No. | Classifier | Accuracy |
|---|---|---|
| **1.** | Linear SVM | **80.19%** |
| **2.** | **Quadratic SVM** | **94.3%** |
| **3.** | Cubic SVM | **84.67%** |
| **4.** | Fine Gaussian SVM | **68.80%** |
| **5.** | Medium Gaussian SVM | **73.48%** |
| **6.** | Coarse Gaussian SVM | **52.45%** |

Table 2
*Comparison of KNN and their corresponding accuracies*

| Sr. No. | Classifier | Accuracy |
|---|---|---|
| **1.** | **Fine KNN** | **75.6%** |
| **2.** | Medium KNN | **68.98%** |
| **3.** | Coarse KNN | **48.89%** |
| **4.** | Cosine KNN | **70.38%** |
| **5.** | Cubic KNN | **68.28%** |
| **6.** | Weighted KNN | **74.59%** |

In the fourth experiment, all nine color features with all the seven texture features have been considered and computed the average accuracy using Quadratic SVM Classifier. The test accuracy of each type of fruits and vegetables by using only color, only texture, and the combination of both, with Support Vector Machine, is mentioned in Table 3. The average accuracy has been found at 94.3%.

Table 3
*Percentage of test accuracy of different fruits and vegetables*

| S. No. | Fruit or Vegetable name | Texture Features | Color Features | All Color + All Texture Features |
|---|---|---|---|---|
| **1** | Diamond Peach | 71.75% | 83.60% | **89.4%** |
| **2** | Agata Potato | 56.42% | 74.46% | **91.5%** |
| **3** | Cashew | 81.33% | 70.86% | **95.33%** |
| **4** | Fuji Apple | 58.55% | 89.33% | **90.8%** |
| **5** | Granny Smith Apple | 72.63% | 85.53% | **100%** |
| **6** | Honeydew Melon | 97.64% | 100% | **97.6%** |
| **7** | Kiwi | 49.54% | 92.94% | **91%** |
| **8** | Onion | 66.66% | 90.09% | **100%** |
| **9** | Orange | 62.79% | 100% | **97.7%** |
| **10** | Tahiti Lime | 65.21% | 86.04% | **85%** |
| **11** | Water melon | 62.12% | 71.73% | **98.4%** |
| **12** | Spanish Pear | 63.63% | 97.72% | **93%** |

Table 3 *(continue)*

| S. No. | Fruit or Vegetable name | Texture Features | Color Features | All Color + All Texture Features |
|---|---|---|---|---|
| **13** | Plum | 83.33% | 83.83% | **90.6%** |
| **14** | Asterix Potato | 59.35% | 89.21% | **92.6%** |
| **15** | Nectarine | 68.85% | 93.58% | **93.04%** |
| | **Overall Accuracy** | **68.22%** | **87.5%** | **94.3%** |

## Comparison with Existing Approaches

Results are also compared with other existing fruit or vegetable detection methods in Table 4 and concluded that the proposed method has achieved better accuracy. Dubey and Jalal (2015) performed fruit classification on same dataset (which was used in this research) and used different feature extraction techniques. Their method obtained 93% accuracy. Arivazhagan et al. (2010) also worked on the same dataset and they used different feature extraction techniques. For classification of fruits and vegetables minimum distance classifier was used and an accuracy of 87% was obtained.

The results are also compared with the approaches which work on a different dataset. In Zhang and Wu (2012), the color, shape and texture features are used with Kernel SVM and an accuracy of 88.2% was achieved. In Shukla and Desai (2016), similar type of features were used and with KNN classifier and an accuracy of 91.3% was achieved.

Table 4
*Fruit recognition accuracy comparison with existing approaches*

| Reference | Fruit/ Vegetable (Dataset) | Features Extraction | Classifiers | Accuracy |
|---|---|---|---|---|
| **Seng & Mirisaee (2009)** | 7 types of Fruits (50) | Mean of color image + shape features | K-Nearest Neighbors Algorithm | **90%** |
| **Arivazhagan et al. (2010)** | 15types of fruits (2612) | Statistical features, cluster shade, cluster prominence | Minimum Distance classifier | **87%** |
| **Zhang & Wu (2012)** | 18 types of fruits (1653) | Color features, texture and shape features | Kernel SVM | **88.2%** |
| **Dubey & Jalal (2015)** | 15typesof fruits (2612) | GCH, CCV, CDH, SHE, LBP, LTP, CLBP | Multiclass SVM | **93.84%** |
| **Shukla & Desai (2016)** | 9 types of fruits (115) | Color, texture, shape features | Multiclass SVM and KNN | **87% with SVM and 91.3% with KNN** |
| **Proposed work** | 15 types of fruits (2612) | Statistical features, GLCM, LBP, HOG | Quadratic SVM (one-against-all strategy) | **94.3%** |

Different classifiers were used to find accuracy. In our proposed method we have archieved an accuracy of 94.3, which is far better than existing methods. So, it is concluded that proposed work performs better than the existing approaches with less error rate.

## CONCLUSION AND FUTURE WORK

Detection of fruits and vegetables is done manually but it becomes a difficult task when it is done in the supermarket or in industry. Time taken by a human to detect the fruits and vegetables becomes high so to reduce the time and to increase the accuracy, automatic detection of fruits and vegetables comes into existence. This research work is used to recognize fruits and vegetables into different categories based on different color and texture features. First a background subtraction method is used to extract the desire region of interest and various color and texture features are extracted. Then two classifiers SVM and KNN were brought into use for training and testing of the images using the features extracted at the previous stage. It is concluded that color gives the superior result than the texture but when the color of some fruit or vegetables are same then it becomes difficult to classify them, so texture features are used to differentiate the fruit or vegetable which have the same color. To further improve the classification accuracy, color and texture features are hybridized and results are computed using SVM and KNN classifier and it is concluded that SVM gives the highest accuracy up to 94.3%.

In this proposed work a single image has multiple fruits or vegetables but of the same type. So, in future the same work can be extended for classification of different fruit or vegetable present in a single image. As the fruits and vegetables are efficiently classified on color and texture basis but in the further work a mobile based application can be devolved which capture images of the fruit and can identify its class. To further improve the accuracy a deep learning model can also applied, and the same work can be also extended by adding more fruits and vegetables images in the database.

## ACKNOWLEDGEMENT

The authors are thankful to acknowledge the support of Department of Computer Science & Engineering, National Institute of Technology Jalandhar for carrying out this research work. We thank our colleagues who provided insight and expertise that greatly assisted in the present work.

## REFERENCES

Abdelmounaime, S., & Dong-Chen, H. (2013). New Brodatz-based image databases for grayscale color and multiband texture analysis. *ISRN Machine Vision, 2013*, 1-14.

Ansari, M. D., & Ghrera, S. P. (2016, November 25-27). Feature extraction method for digital images based on intuitionistic fuzzy local binary pattern. In *International Conference on System Modeling and Advancement in Research Trends (SMART)* (pp. 345-349). Moradabad, India.

Ansari, M. D., & Ghrera, S. P. (2018). Intuitionistic fuzzy local binary pattern for features extraction. *International Journal of Information and Communication Technology, 13*(1), 83-98.

Arivazhagan, S., Shebiah, R. N., Nidhyanandhan, S. S., & Ganesan, L. (2010). Fruit recognition using color and texture features. *Journal of Emerging Trends in Computing and Information Sciences*, *1*(2), 90-94.

Clausi, D. A. (2002). An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing, 28*(1), 45-62.

Dalal, N., & Triggs, B. (2005, June 20-25). Histograms of oriented gradients for human detection. In *Computer Society Conference on Computer Vision and Pattern RecognitionCVPR 2005* (pp. 886-893). San Diego, United States.

Dubey, S. R., & Jalal, A. S. (2015). Fruit and vegetable recognition by fusing colour and texture features of the image using machine learning. *International Journal of Applied Pattern Recognition, 2*(2), 160-181.

Hameed, K., Chai, D., & Rassau, A. (2018). A comprehensive review of fruit and vegetable classification techniques. *Image and Vision Computing*, *80*, 24-44.

Jhawar, J. (2016). Orange sorting by applying pattern recognition on colour image. *Procedia Computer Science*, *78*, 691-697.

Lalibertea, A. S., & Rango, A. (2008, August 5-8). Correlation of object-based texture measures at multiple scales in sub-decimeter resolution aerial photography. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (pp.1-6). Calgary, Alberta, Canada.

Moallem, P., Serajoddin, A., & Pourghassem, H. (2017). Computer vision-based apple grading for golden delicious apples based on surface features. *Information Processing in Agriculture, 4*(1), 33-40.

Ninawe, P., & Pandey, S. (2014). A completion on fruit recognition system using k-nearest neighbors algorithm. *International Journal of Advanced Research in Computer Engineering and Technology*, *3*(7), 2352-2356.

Prakasa, E. (2016). Texture feature extraction by using local binary pattern. *INKOM Journal, 9*(2), 45-48.

Prakash, J. S., Vignesh, K. A., Ashok, C., & Adithyan, R. (2012, December 14-15). Multi class Support Vector Machines classifier for machine vision application. In *International Conference onMachine Vision and Image Processing (MVIP)* (pp. 197-199). Taipei, Taiwan.

Pujari, J. D., Yakkundimath, R., & Byadgi, A. S. (2013). Grading and classification of anthracnose fungal disease of fruits based on statistical texture features. *International Journal of Advanced Science and Technology, 52*(1), 121-132.

Reddy, B. R., Suresh, A., Mani, M. R., & Kumar, V. V. (2009). Classification of textures based on features extracted from preprocessing images on random windows. *International journal of advanced Science and technology, 9*, 9-18.

Rocha, A., Hauagge, D. C., Wainer, J., & Goldenstein, S. (2008, October 12-15). Automatic produce classification from images using color, texture and appearance cues. In *XXI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'08)* (pp. 3-10). Campo Grande, Brazil.

Seng, W. C., & Mirisaee, S. H. (2009, August 5-7). A new method for fruits recognition system. In *International Conference on Electrical Engineering and Informatics (ICEEI'09)* (pp. 130-134). Selangor, Malaysia.

Shukla, D., & Desai, A. (2016, December 19-21). Recognition of fruits using hybrid features and machine learning. In *International Conference on Computing, Analytics and Security Trends (CAST)* (pp. 572-577). Pune, India.

Sonka, M., Hlavac, V., & Boyle, R. (2014). Image processing, analysis, and machine vision. *Cengage Learning, 9*(4), 55-58.

Teoh, A., Samad, S. A., & Hussain, A. (2004). Nearest neighborhood classifiers in a bimodal biometric verification system fusion decision scheme. *Journal of Research and Practice in Information Technology, 36*(1), 47-62.

Wang, C., Li, Z., Dey, N., Li, Z., Ashour, A. S., Fong, S. J., & Shi, F. (2018). Histogram of oriented gradient based plantar pressure image feature extraction and classification employing fuzzy support vector machine. *Journal of Medical Imaging and Health Informatics, 8*(4), 842-854.

Zaitoun, N. M., & Aqel, M. J. (2015). Survey on image segmentation techniques. *Procedia Computer Science, 65*, 797-806.

Zawbaa, H. M., Abbass, M., Hazman, M., & Hassenian, A. E. (2014, November 28-30). Automatic fruit image recognition system based on shape and color features. In *International Conference on Advanced Machine Learning Technologies and Applications* (pp. 278-290). Cairo, Egypt.

Zhang, Y., & Wu, L. (2012). Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine. *Sensors*, *12*(9), 12489-12505.