

# Fruit quality evaluation using Machine Learning : A review

Hitanshu

Department of CSE  
HMR Institute of Technology and Management  
Delhi, India  
hitanshujhamb15@gmail.com

Parul kalia

Department of CSE  
HMR Institute of Technology and Management  
Delhi, India  
k.parul97@gmail.com

Akash Garg

Department of CSE  
HMR Institute of Technology and Management  
Delhi, India  
akashgargsdps@gmail.com

Amit Kumar

Department of CSE  
Ambedkar Institute of Advance communication  
Technology and Research  
Delhi, India  
amitr6002@gmail.com

**Abstract**— Automation increases economic growth of the country. By the same time it is also helpful in micro, small, medium and large scale industry. The export market of our country is a leading in terms of exports of fruits. Diseases arises due to contaminated fruits and vegetables has been high in developing countries due to less awareness towards the efficient quality product. The sensory characteristics of fruits are its appearance in terms of color, shape, size etc. This paper presents a overview of different methods, technology and techniques that is implemented to get quality evaluation of fruits.

**Keywords**— KNN, SVM, ANN, Quality Evaluation, Classification, Detection

## I. INTRODUCTION

Image visualization is the most basic method of classifying foodstuff and it is done by human brain. Human brain involves experience to get the best possible result of the query. But the question arises when it comes for freshness of foodstuff. Even an experienced human brain cannot answer this question with perfect accuracy. But it is seen that the better quality demands higher prices “value for money”. Hence, a special quality inspection of fruits is done by investigators and depending upon which a grading to the fruit is done[1]. Fruit of higher grade represents higher quality and definitely higher prices than usual. Thus in field of agriculture and food delivery services a person must know the quality of product which is to be delivered to the customers.

The general method of determination of fruit quality is done visually but to perform such task through machine is challenging. But a machine learns in a similar manner as a human brain does and hence it is required by machine a set of images and some pre-processed inputs that helps in generating a system which can perform a same task which an investigator for inspection of quality of fruits. A detailed summary of fruit quality is not available and hence the objective of this review

paper is to analyze existing techniques for quality detections and different methods involve for making such a machine. Also, the principal components, basic theories and corresponding analysis and processing methods are reported.

## A. Sorting Machines

Sorting machines consists of a mechanical sorter and an attached conveyer belt. This machine done sorting of fruits depending upon dimensions such as weight, size, shape. For example, a product, be it fruit, is dropped on the conveyer belt rotating around a machine which consist of inputs of size shape, etc. and a sorter starts sorting the fruits which does not processes a given threshold[2]. These threshold values are tested by machines for each and every piece of fruit and sometimes a batch of fruits when it comes to quality. Hence this sorting technique is faster and reliable.

## II. FRUITS QUALITY INSPECTION

Quality inspection in Figure 1. of fruits involves machine inputs and classification of fruits depending upon the features they have. Hence this involves processes like pre-processing of image of fruit, segmentation of image, feature extraction of image, classification of fruits depending upon the image and identification done by machine[3].

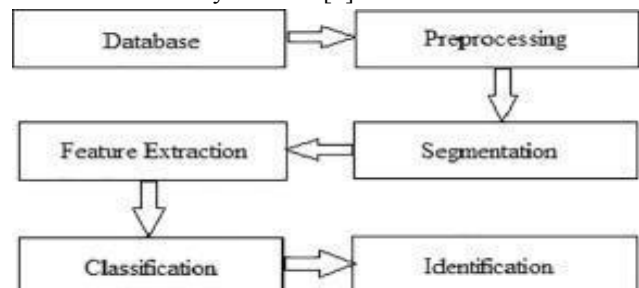


Figure 1: Block diagram of identification in image processing.

### A. Pre-processing

It is basic abstraction of images and using it as inputs and outputs for recognition. It consists of multiple images of a product with different side and angles and providing those image to machine for learning. For example, for a machine to recognize which fruit it is multiple images of the fruit is given as an input to the machine such that it recognize the fruit.

The concept of providing multiple images as dataset is that a machine can perform in a better manner with more amount of data provided. Machine uses pixels of input image and as well as other local processing for the image pre-processing. Better results are shown when dataset of multiple image is used. Hence more the images better the result.

### B. Segmentation

The process of segmentation involves clearing the digital image into multiple segments. This segments consists of different set of pixels as per the color of the image taken/uploaded. This is mainly done by two techniques thresholding and clustering[4]. The thresholding technique consists of replacing color pixels/image pixels of the image with the black pixels and rating the pixel with some number which is in a taken range depending upon the image pixel. Whereas, clustering consists of making clusters of images and segmenting it depending upon the colors of the image. For example, clusters of fruit having similar shapes , colors etc.

### C. Feature Extraction

After segmentation, analysis of image is done which is feature extraction. Feature extraction consists of effective data of image which involves perception made by machine using the image, classifying the object involved in the image and interpretation of the object. For such extraction different factors are defined and depending upon that factors different features are majored. For example, factors such as color, shape etc. plays an decisive role. Like color – red, shape-round will figure out as apple.

### D. Color and shape features

The basic features that is essential for fruit quality majors are its color and shape. For example, color – red, shape-round will figure out as apple. The color features depends upon the image pixels and combination of color in which it is used by machine. Machine uses RGB color in a combination of all three. In similar manner feature extraction through shape is done using different shapes that is given as input to the machine. Such shapes include circular shape , cylindrical shape etc. these shapes are taken in recognition and desired output is shown with the combination of the two images that is shown or taken.

- Texture features

Human brain usually remembers a specific texture for a specific things especially in terms of fruits[5]. But when it comes to machine learning the texture comes in terms of pixels. Bright pixels represents color of the portion is light in nature or shiny whereas dark pixels represent dark color in the image. For example, texture on watermelon is green and black strips.

Texture features have both qualitative and quantitative factors which involves roughness, contrast, directionality etc.

## III. DEFECT DETECTION OF FRUITS

In conventional machines defects are detected when there is any deformation or when it does not pass the threshold requirement of the machine. But visually or by machine it is difficult for to determine the defect and the effects of the defect. It is difficult to identify by machine whether it is a defect or not or the range of defect is under major , minor or medium[6]. For example , a small black patch on banana is common and hence cannot be considered as a defect for banana but same size black patch on orange may result in rotten condition of orange. Hence defect detection of fruits is done in different manner for different fruits. But these defects will result in some specific condition of fruits either eatable or non-eatable(rotten).

## IV. CHALLENGES

Image processing of the fruits requires a large dataset. The quality detection of fruits depends upon this dataset. Thus an error in dataset or an inappropriate input may result in the vital results of the system. For such systems of fruit freshness detection these dataset requires large amount of images for more efficient result and attaining high accuracy. Hence handling such large dataset is a tough task. Apart from this for such systems it is difficult to achieve 100% accuracy hence outcomes may or may not be perfectly accurate but its efficiency can be increased to maximum with maximum number of images that is possible to store and required by system for training.

### A. Optical and visual grading

Some Dedicated algorithms are used for optical and visual grading of such systems. This is done by analysis of fruits done by the system using multiple images of a single fruit in order to ensure consistent performance[7-9]. This analysis is based on external quality which is dependent on color and texture. It also include some other features such as shape dimensions etc. some other techniques that are can be used by the systems are IR cameras and other cameras.

### B. Existing systems and algorithms

Algorithms being an integral part for the evaluation purpose, several algorithm are taken into consideration by the researchers for different purposes and thus for developing required systems integration of algorithms with respect to basic characteristics, using normalization of dataset. Different algorithms like such as clustering and threshold are used for detecting the defects in the system and thus these algorithms are also considered for training the datasets in such a best possible way that avoidance of detection of defects occur. Therefore, the actual result depicted by the system depends upon the dataset provided and to the level it is trained, also maintaining the integrity of the flow so as to maintain the order of thousands. Algorithms such as clustering and threshold algorithms used for detection of detects in the system and hence these algorithms are used to train datasets in best possible manner such that no miss detection takes place. The actual result depicted by the system will depend upon the

dataset provided and to the level at which it is trained. Hence a highly trained set of data will provide a accurate result and high performance.

### C. New methods: deep learning

Growing ways of analysis and with the enhancement in its levels have led to the modern methods such as deep learning. The automatic learning has taken over tuning learning. Considering the functionality of learning algorithms when a good set of “good” fruits (suppose oranges) and another set of defected fruits taken together, the algorithm classifies itself the defected fruit and fresh fruit. Grading of the product classified is depended upon the clusters of classification. Thus a reliable and fast outcomes are generated and therefore provides consistent performance. Thus deep learning nowadays usually used to provide the better ordered, reliable and consistent results.

## V. ANALYSIS

Generally, the identification of products is done manually either by the consumer using self-service system or by the responsible person of the unit in the retail businesses[10]. But the involvement of the human may affect the resulting output. There may occur a case that user may not be able to understand the exact function of the application or may press a wrong button with respect to the condition and state. The emerging technologies has provided a good platform for better GUI designing and achieving high efficiency.

User interface is developed for evaluating the problem of user using the heuristic search technique applied using the expression semantics analysis of the person in a positive or negative expression. Several set of guidelines are followed while developing a user interface. These guidelines are made to reduce the critical errors among designs and upholding the user requirements and hence some of the guidelines are as follows:

- Simple and Natural Dialogue
- Error Prevention
- Be consistent
- Provide Feedback, etc.

This evaluation done by heuristics satisfies some general facts about the guidelines followed by the developers. It shows that an evaluation result is accurate when the guidelines are followed by group or multiple participants[11]. Considering the fact that one participant cannot follow all the guidelines.

### A. Optical and visual grading

Some Dedicated algorithms are used for optical and visual grading of such systems. This is done by analysis of fruits done by the system using multiple images of a single fruit in order to ensure consistent performance. This analysis is based on external quality which is dependent on color and texture. It also include some other features such as shape dimensions etc. some other techniques that are can be used by the systems are IR cameras and other cameras.

### B. Existing systems and algorithms

Algorithms being an integral part for the evaluation purpose, several algorithm are taken into consideration by the researchers for different purposes and thus for developing required systems integration of algorithms with respect to basic characteristics, using normalization of dataset. Different algorithms like such as clustering and threshold are used for detecting the defects in the system and thus these algorithms are also considered for training the datasets in such a best possible way that avoidance of detection of defects occur. Therefore, the actual result depicted by the system depends upon the dataset provided and to the level it is trained, also maintaining the integrity of the flow so as to maintain the order of thousands[12]. Algorithms such as clustering and threshold algorithms used for detection of detects in the system and hence these algorithms are used to train datasets in best possible manner such that no miss detection takes place. The actual result depicted by the system will depend upon the dataset provided and to the level at which it is trained. Hence a highly trained set of data will provide a accurate result and high performance.

### C. New methods: deep learning

Growing ways of analysis and with the enhancement in its levels have led to the modern methods such as deep learning. The automatic learning has taken over tuning learning. Considering the functionality of learning algorithms when a good set of “good” fruits (suppose oranges) and another set of defected fruits taken together, the algorithm classifies itself the defected fruit and fresh fruit. Grading of the product classified is depended upon the clusters of classification. Thus a reliable and fast outcomes are generated and therefore provides consistent performance. Thus deep learning nowadays usually used to provide the better ordered, reliable and consistent results.

### D. Methodology

Generally, methodology is theoretical technique, systematic analysis and technology that can be used to the method applied to a field of study. Methodology here can be defined as the step by step procedure and technique that can be applied to develop a system. The proposed methodology in this paper, step to perform the analyses:

- Selection of fruit image.
- Segmenting the portion of fruit.
- Classification of fruits depending upon the dataset available.
- And the threshold value and associated classification is being taken for the process.

### E. K-Nearest Neighbors Algorithm

The k-nearest Neighbor algorithm can be stated as the type of methodology considered to develop the system for fruit recognition[13]. KNN performs the classification on the fruit data by taking the distance between the values of the features of the unknown fruit with the available feature value of already stored

fruit example and thus provides the nearest result to unknown fruit. KNN classifier classify to fruit mean color value; shape roundness value, area and perimeter values.

## VI. CLASSIFICATION

The classification is described as an essential component to be performed and is defined as an important feature for the food quality evaluation. This contributes to a structure involving the artificial stimulation of human thinking so to enhance the judgment making activity more correct and accurate. Also by using image processing techniques, fruit images can be described through their basic characteristics like color, shape, size, texture and defects. These are used to prepare a training set[14]. Further an appropriate classification algorithms are thus applied to these described training set that evaluates the knowledge base for all unknown cases. At present there are several methods that can be used for the classification process of fruits. Some of them are KNN, ANN and SVM techniques and algorithms that can complete the structure of processing.

KNN follows the Euclidean distance calculation for calculating the nearest neighbors of the desired point. It simply calculate one set point through Euclidean distance and a new set point through the previous point and this cyclic process continues to get the exact result.

Whereas SVM is an algorithm where linear and non-linear data is classified. It uses kernel functions to non-linearly map the data[15]. It calculates a linear optimal hyper plane to find extreme support vector ends.

ANN is a stimulation program of human brain processing done with the help of computer vision. It is a computer program that works exactly as processes an instruction.

## VII. STUDY

### A. Identification and Classification of Mango Fruit

Mango is one of the favorite tropical fruits with an increase in its production every year. Generally, the basic property like color indicates the maturity and presence of defects in fruit. Also the physical appearance affects the value in the markets, thus it is necessary to avoid the defect causing reasons at the time of their handling after harvesting. Thus sorters need to be understand the requirements that must follow for the excellent flow of the fruit trade.

In this Sorting of the objects is usually performed using its physical appearance. Automatic sorting further taken into consideration and developed for the other respective products. Thus the classification of the mangoes depends upon its physical appearance. As discussed it also involves the manual labour and substantially it is dependent on the human vision system. Also classification involves the small componential steps for the evaluation purposes one of them is uniformity.

Uniformity in classification is an essential step so that its certified output is provide for the export purposes . Thus indicates the change in the system of grading processes over the period of time after the quality and other important aspects of the analysis. Change depicts the transfer from traditional grading system to its automated system for classification. These advance are now being applied to the number of other fruits like oranges and peaches. Though the basic characteristics are considered like color, shape and size mostly for all category of the fruit to them but in case of Mango skin texture is founded to be the another important characteristic so as to improve the accuracy of the classification system . The purpose of the study is to use an image processing algorithm that can help in automating the classification of mangoes.

### B. More examples: Banana

While taking Banana for the analysis level color of the peel is taken as the considerable feature for the identification, classification purposes. Change in color of banana during ripening process from green to yellow. Therefore, Color of the peel is considered as a first quality parameter evaluated by traders and consumers.

Number of studies have explained the potential of banana fruit quality analysis using image processing.

However the for the like banana the physiological maturity stages of fruit. They involves the peel factor for the classification the selected fruit .the change in the peel factor helps in indicating the maturity of the fruit.

## VIII. ROLE OF MACHINE LEARNING: SUMMARISED

Several Machine learning methods with the emerging advancement in problem are used and thus this resulted in the evaluation of the factors with the use of attributes, designing of set of data, generally termed as "Training Set".

While discussing the precise points we are able to understand that not only at the ground level but also at the market level the learning methods are used to maintain the basic model of the evaluation in our oriented field. It benefited the evaluation constraints in a significant manner like:

- It provide the retailers an assistant in creating the more efficient way to churn the data to create better quality check.
- The AI technologies are used to determine which corn and which conditions will produce the best decision using suitable algorithms. It will also determine which weather condition will give the highest return.
- Using AI and machine learning is good for the food tech segments.

## IX. CONCLUSION AND FUTURE ASPECTS

This review paper highlights the combination of researches and classification done by researchers. It is an attempt to

compare different methodology that can be applied to the system. The classification of fruits and other products are done using these algorithms. The essential algorithm that is required for such classification is KNN algorithm which finds the nearest neighbor of the image which means image of the product resembling the same product required. Along with this sorting of the product is done due to which correct product is extracted. The basic attributes that are associated with fruits are shapes, size, color and texture. Basic steps to identifying this is pre-processing, segmentation, feature extraction and classification. Hence it is an attempt to know how a fruit freshness can be rectified using such technologies.

## REFERENCES

- [1] P. Mohanaiah, P. Sathyanarayana, L. GuruKumar, "Image Texture Feature Extraction Using GLCM Approach", International Journal of Scientific and Research publications, vol. 3, no. 5, May 2013
- [2] Anand Singh Jalal, Shiv Ram Dubey, "Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns", IEEE Third International Conference on Computer and Communication Technology, pp. 978-0-7695-4872, 2012
- [3] N. Pujitha, C. Swathi, V. Kanchana, "Detection Of External Defects On Mango", International Journal of Applied Engineering Research, vol. 11.7, pp. 4763-4769, 2016.
- [4] R. Manali Satpute, M. Sumati Jagdale, "Automatic Fruit Quality Inspection System" in Inventive Computation Tehnologies (ICICT) International Conference on, IEEE, vol. 1, 2016
- [5] Yogesh, Ashwani Kumar Dubey, "Fruit Defect Detection based on Speeded Up Robust Feature Technique", Reliability InfocomTechnologies and Optimization(Trends and Future Directions) (ICRITO) 2016 5th International Conference on. IEEE, 2016.
- [6] Dameshwari Sahu, Ravindra Manohar Potdar, "Defect Identification and Maturity Detection of Mango Fruits Using Image Analysis", American Journal of Artificial Intelligence., vol. 1, no. 1, pp. 5-14, 2017
- [7] C. Nageswara Rao, S. Sreehari Sastry, K. Mallika, Ha Sie Tiong, K.B. Mahalakshmi, "Co-Occurrence Matrix and Its Statistical Features as an Approach for Identification of Phase Transitions of Mesogens", International Journal of Innovative Research in Science Engineering and Technology(IJRSET), vol. 2, no. 9, September 2013.
- [8] Krishnan Sarath, S. Padmavathi, "Feature ranking procedure for automatic feature extraction", Signal Processing Communication Power and Embedded System (SCOPE) 2016 International Conference on. IEEE, 2016.
- [9] Yassine Ben Salem, Salem Nasri, "Texture classification of woven fabric ased on a GLCM method and using multiclass support vector machine", Systems Signals and Devices SSD'09 6th International Multi-Conference n IEEE, 2009.
- [10] S.A. Khoje, S.K. Bodhe, A. 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, pp. 3251-3256, 2013
- [11] K.N. Ranjit, H.K. Chethan, C. Naveena, "Identification and Classification of Fruit Diseases", International Journal of Engineering Research and Application (IJERA), pp. 11-14, 2016.
- [12] R. Socher, B. Huval, B. Bath, C.D. Manning, A.Y. Ng, "Convolutional-Recursive Deep Learning for 3D Object Classification", Advances in Neural Information Processing Systems, pp. 656-664, 2016.
- [13] Putra, I. W. S. E. (2016). Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101(Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- [14] Zufar, M., & Setiyono, B. (2016). Convolutional neural networks untuk pengenalan wajah secara real-time. Jurnal Sains dan Seni ITS, 5(2).
- [15] Puttemans, S., Baldewijns, G., Croonenborghs, T., Vanrumste, B., & Goedemé, T. (2015, August). Automated walking aid detector based on indoor video recordings. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 5040-5045). IEEE.