Fruit Quality and Grading System

Vimal Kumar
Department of information Technology, Hindustan Institute of Technology

Chetan Chebrolu
Department of information Technology, Hindustan Institute of Technology

Pachaiyappan p
Department of information Technology, Hindustan Institute of Technology

Dasari Vineel
Department of information Technology, Hindustan Institute of Technology

Aravind T
Department of information Technology, Hindustan Institute of Technology

Abstract -Fruit rotting has significant economic effects; it is estimated that decaying elements account for nearly a third of the cost of the fruit. Furthermore, because customers feel that damaged fruits are dangerous to their health, fruit sales will be affected. Reduced levels of amino acids, vitamins, sugar/glucose, and other nutrients unavoidably create public worries about edibility, spurring ideas on ways to prevent or reduce decay. Despite the importance of food status in peoples lives and its contribution to the economy, fruit freshness grading is a time-consuming manual operation. The employment of digital technologies to automate grading is regarded to be the solution to this challenge. The model files for fruits, as well as their categorization and grading systems, are built and evaluated. For this, we gathered a collection of photos from places like Kaggle. The photos are then loaded into the Python program, where the code is integrated. RF (Convolutional and Neural Network) and Machine Learning are used to train and assess these pictures. having to make a personal choice. Grading can save money on marketing costs such as packaging and delivery. Grading is the practice of separating vegetables and fruits into different classes depending on their size, shape, colour, and volume in order to maximize.

I. NEED FOR THE PROJECT

Fruit grading is a crucial stage for growers since it affects both the quality of the fruit and the export market. Human grading and sorting are possible, but it is time-consuming, labour-intensive, mistake-prone, and exhausting. As a result, a sophisticated fruit grading system is necessary. The classification and grading of rotting and fresh fruits are completed. The proposed idea might be a great way for merchants to produce high-quality sales of culinary items, especially short-lived fruits. This is an excellent solution for handling food commodities in markets, stores, and other settings.

II.OBJECTIVES:

To create a method for identifying rotten/fresh fruits and evaluating them.
To signify that RF and Machine Learning are being used.
To signify that RF and Machine Learning are being used.

III.REPORT ORGANISATION:

There are five sections to the report. All of the chapters have been numbered consecutively for easy identification. The first chapter is introductory, giving a concise outline of the project's need and objectives. A literature review is presented in the second chapter. In Chapter 3, the project's design methodology, software design, is discussed. Chapter 4 delves into the findings and their interpretation. The conclusion and future efforts are discussed in Chapter 5.

CHAPTER II

IV.LITERATURE REVIEW

picture using the Kmeans clustering algorithm. The L*a*b* colour model is used in the planned automated strawberry grading system. The greatest fruit diameter is used to determine the size of the fruit, and the colour of the strawberry is retrieved using the Dominant colour technique on the channel. The strawberry size detection inaccuracy is less than 5%, the colour grading accuracy is 88.8%, and the form classification.The approach suggested in [1] can efficiently acquire the form characteristic by drawing lines and then classifying

the strawberry accuracy is above 90%, according to the data. The average time it takes to evaluate one strawberry is less than 3 seconds. The provided date fruit rating method is based on computer vision
Flabbiness, size, form, intensity, and faults are retrieved features in the grading process of date fruit. It categorizes dates into three quality groups (grades 1, 2, and 3) based on the retrieved attributes. Dates with a nice form, a large size, a high intensity, a high flabbiness, and no faults were labelled as being of the highest grade. This study proposes a synthesis of colour and texture. Intensity, colour, shape, and texture are some of the several factors that may be used to grade fruit. Fruit is classified using a minimum distance classifier based on Wavelet transformed sub-bands. Uses the direct colour mapping approach for automated fruit colour grading. The suggested technique calculates a unique set of coefficients for colour space conversion using pre-selected colours. The three-dimensional RGB colour space is broken down into a minimal number of colour indices that are specific to the application. The suggested approach evaluates tomato and date ripeness as well as date surface fault identification. The strategy provided is straightforward yet effective. This novel direct colour mapping idea may be used in a number of colour grading applications that demand quick colour preference selection and tweaking.

Describes a novel method for evaluating pineapples based on colour. Pineapple images were gathered and the backdrop was eliminated. The pineapple image's RGB component was retrieved.
The colour values collected during the sorting step are recorded in a database for training the Neural Network (NN). According to the test findings, the maximum level of accuracy attained for grading pineapples is 75%.
Bernard Gosselin and DevrimUnay used LDC to undertake an apple fruit categorization test. K-NN is a statistical classifier that employs a distance metric to estimate sample similarity (proximity). Within its nearest k neighbours, it allocates data to the most represented category.
According to the exterior surface, AnuradhaGawande [7] suggested a technique to detect the contaminated region from the input photos and classify the diseased patterns as a low, average, medium, high, extremely high, and entirely infected fruits. However, it only processes one image at a time, rather than a batch, which is the system's main drawback.
[7]discussesseveral segmentation algorithms, colour models, and feature extraction strategies for fruit disease diagnosis and grading. The sorting of mango fruits is done in [9] using a Gaussian Mixture Model and Fuzzy logic, using ripeness and size as parameters. For various maturity levels, results range from 88 percent to 92 percent. Colour and texture data are combined with a unique radial basis probabilistic neural network for orange fruit defect classification, and 88 percent accuracy is reached in [10].

[8]Based on the quality ratio, Sahu and Potdar [11] offer an algorithm to determine defect and maturity in mango fruit. The fruit is rotten if the quality ratio value exceeds the threshold value. The fruit is excellent if the quality ratio value is less than the cutoff value. As a result, the suggested algorithm sorts mango fruits based on quality, which is critical for fruit value addition.
[9]Naik and Patel [12] propose a mango grading system based on size and ripeness. A thermal camera is used to forecast the maturity mean intensity method in CIELAB colour space. Mango size is predicted using weight, eccentricity, and area, with an accuracy of 89.00% and a time of 2.3 seconds.
[10]Arakeria and Lakshmana [13] present a computer vision-based tomato fruit grading system that comprises of two phases: hardware development and software development. Without operator involvement, the created technology takes the picture and moves the fruit to the right container. The programme uses image processing techniques to evaluate the fruit for maturity and flaws, resulting in a 96.47 percent accuracy in determining the tomato's quality.

CHAPTER III

V.METHODOLOGY

Our project is primarily concerned with improving the setup's architecture. The hyper parameters including epoch, batch size, and validation split are altered to improve precision.The batch size refers to the number of image sets that will be learned at once. The accuracy improves as the number of photos decreases. The process's iteration count is called an epoch. Validation split is a method of validating data in a randomised order to increase accuracy.

3.1 The samples that were used

For training purposes, a total of 7500+ photos were captured. 5000+ apple photos, 2500+ banana images, and 500+ orange images were taken from the 7500+ training images.

*3.2Filtering and image pre-processing*this component eliminates noise, sharpens, and smoothens the image, as well as does image scaling. RGB pictures are transformed into greyscale images, and the image contrast is boosted to a certain extent. Filtration is a term used to describe certain types of preprocessing processes.

3.2.1 Pre-processing:

Images come in a variety of sizes and forms. They also originate from a variety of places.
We must conduct some pre-processing on any picture data in order to account for all of these variances.

Most "natural pictures" are encoded in RGB, which is the most prevalent encoding standard. Making the photos the same size is also a preliminary step in data pre-processing
.For training, we utilized auto-resizing to transform all of the photos in the dataset to the same resolution.

3.3 Image Segmentation.

The term "segmentation" refers to the process of dividing a picture into several segments. The purpose of picture segmentation is to make an image more understandable and simpler to examine by simplifying and/or changing its representation. The qualities of discontinuity and similarity are used to characterize image segmentation algorithms. Boundarybased approaches are those that are based on discontinuities, whereas region-based methods are those that are based on similarity.

3.4 Extraction of Characteristics

After segmentation, feature extraction is the next stage in the fruit classification process. Colour, size, shape, and texture are the most visible exterior aspects of the fruit. A feature descriptor is a representation of an image or a portion of an image that extracts important information while discarding irrelevant data. It is mostly used to recognize images as well as to object detection.

3.5 Knowledge-based Comparison and Decision Making

The retrieved characteristics from the picture are compared against the preset classification and sorting criteria or rules. On the basis of the retrieved features, the characteristics are compared, and classifications are assigned to the fruits. Random Forest methods were used for knowledge-based comparison and decision-making. The fourth section explains the algorithm. The categorization system comes to a close with this stage.

3.6 DATASET

All of the images used in training and testing came from the fruits 360 dataset, which is freely accessible on Kaggle. The collection comprises 10000 distinct fruit images from seven different classifications. Each fruit kind is represented by a different class. Apple, banana, and orange are the courses that have been picked. The data set was limited in order to keep the study from becoming too long. These constraints are due to the fact that all varieties of fruits belong to the same class. This means that all sorts of apples are classified as apples, and each fruit is similarly classified. Image Net was used to gather images for the data collection. All of the fruits were scaled to 100100 pixels of standard RGB images after the backdrop was removed. 7500+ photos from 3 distinct categories were chosen from the fruits-360 collection.

Fig 3.6 Rotten Banana

apple

Fig 3.2 Rotten

RANDOM FOREST:

Random forest is a supervised learning algorithm which is used for both classification as well as regression. But however, it is mainly used for classification problems. As we know that a forest is made up of trees and more trees means more robust forest. Similarly, random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result.

Working of Random Forest Algorithm

We can understand the working of Random Forest algorithm with the help of following steps −

Step 1− First, start with the selection of random samples from a given dataset.

Step 2− Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree.

Step 3−In this step, voting will be performed for every predicted result.

Step 4− At last, select the most voted prediction result as the final prediction result.

namely, image acquisition, pre-processing, image segmentation, feature extraction and classification.

*3.9 Image Acquisition*Image capture technologies used in food applications include cameras, ultrasound, magnetic resonance imaging (MRI), electrical tomography, and computed tomography (CT). Charged coupled device (CCD) and complementary metal oxide semiconductor (CMOS) image sensors are utilised to create the digital image. The five basic components of a computer vision system are lighting, an image capturing board (digitizer or frame grabber), a camera, and computer hardware. The light systems in the study of fruits and vegetables are designed as front and rear illumination. Front lighting is used to assess surface quality parameters such as colour, texture, and skin flaws. Back lighting is specified to check the border quality aspects such as size and form. Traditional, multispectral, and hyperspectral computer vision systems have a lot of definitions.

3.10 PREPROCESSING

Multiple sounds are present in images captured using diverse procedures, degrading the image's quality. As a result, it is unable to give relevant data for picture processing. Preprocessing improves picture data by overcoming inherent distortions and enlarging image elements that are important for processing, resulting in a more relevant image (degraded form) than the original for a specific application. Pixel preprocessing and local preprocessing are two methods for picture pre-processing for food quality evaluation. Pixel preprocessing is defined as "converting an input picture into an output image in which each output pixel is connected to the input pixel with the associated coordinates." Colour space transformation (CST) is the most widely used pixel pre-processing method for evaluating

3.11 Segmentation

Picture segmentation, which divides a digital image into different sections, is necessary after preprocessing. The main purpose is to separate the backdrop so that the important region may be processed during the object assessment. A correct segmentation is necessary for future image analysis progress, and a poor segmentation will degrade the classifier's performance.Thresholding divides a digital picture into several sections based on image grey levels. The original image is changed into something else. Pixels with a certain grey level belonged to the interest area class, whereas pixels with the same grey level belonged to the background class." The grey level histogram (to acquire appropriate threshold) is obtained from the grayscale picture using the Otsu technique. The Otsu approach offers a number of advantages, including the ability to optimise threshold values and analyse grey level images without prior knowledge of the picture. Although, as the number of clusters grows, this technique has drawbacks, such as the increased processing time necessary for the ideal threshold value. The clustering technique are used instead of thresholding based method to mitigate the computational time. A clustering technique form the cluster based on similar characteristics of pixel and classified into Hierarchical and Partition-based method. The former method is based on the tree structures in which the roots and the branches indicates the whole database and clusters respectively. The latter method uses optimization technique to optimize the cost function. The basic two types of clustering are hard and soft clustering. Hard clustering is a straightforward technique which segments the image based on pixels belonging to identical clusters. Soft clustering is more realistic as noise exact division does not exist in real life. An example of this technique is Fuzzy C-means in which one pixel can belong to more than one clusters.

3.12 Extraction of features Features are estimated after picture segmentation for subsequent analysis. These characteristics are essential in a computer vision system because they provide useful data for visual perception, interpretation, and object categorization. During this step, retrieved characteristics are converted into feature vectors, which are then classified to recognise the input. These feature vectors characterise the form of the item in a unique and accurate manner. The goal of feature extraction is to improve recognition rates by extracting features. These characteristics provide clear data that may be used for quality assessment and analysis in the food sector. Colour, texture, and morphological characteristics are commonly used to assess fruit and vegetable defect and maturity.

3.12.1 Colour characteristics

Colour is one of the characteristics that influences a customer's decision to reject or pick fruits and vegetables. In ripeness, growth, and postharvest processing and handling stages, it is the indirect measurement of quality characteristics such as freshness, desirability and variety, maturity, and safety. It is governed by physical and chemical changes, internal biochemical, and microbial changes. In image retrieval and indexing, the colour feature is the earliest and most extensively utilised visual feature. High efficiency, simplicity of extracting colour information from photographs, size and orientation independence, strong in portraying visual content of images, resilient against backdrop problems, and powerful in differentiating images from one other are only a few of the benefits of the colour feature. For colour examination of fruit and vegetable quality, the RGB colour system, HSI colour space, and CIELab colour space are often employed.Images are captured using RGB colour models, which are based on the primary colours red, green, and blue. This colour model divides a picture into red, green, and blue planes, and determines all colour moments. In a picture, various RGB devices output varying RGB values for the same pixel, hence numerous transformation techniques are employed to normalise these values. RGB is unable to examine the sensory qualities of food goods since it is nonlinear with the visual examination of human eyes. To address this, HSI was suggested and developed as the primary tool for evolving image processing algorithms based on colours that are widely recognised and accepted by people. HSI and RGB, on the other hand, are quite comparable and are unaffected by minor colour variations. As a result, they are not recommended for analysing product colour alteration during processing. The CIELAB colour space was created to represent all colours visible to the human eye as a device dependant model to be used as a reference, where 'L' stands for brightness, and 'a' and 'b' stand for red/green and green/blue balance, respectively. It is perceptually uniform in the sense that human perceptions of colour differences are the same as Euclidean distances in CIELAB space. Because the colour measured by computer vision can be easily compared to the colour derived from the CIELAB colour space, it provides a viable means of assessing the performance in measuring object colour. CIELAB colour space is suggested as the optimum colour space for quantifying meals with curved surfaces.

3.12.2 Morphological features

Fruit and vegetable categorization is typically based on physical characteristics (size and form). Because the size of fruits and vegetables affects their price in the agriculture business, different size groups are assigned to grading fruits and vegetables at different stages of processing. In comparison to inherent irregularity in complex meals, inspecting spherical and quasi-spherical item sizes of fruits and vegetables is quite simple. The predicted area, perimeter, length, breadth, main and minor axes are used to quantify the feature size. These properties are often employed in businesses for automated sorting. The area (scalar amount) determines how many pixels are in the region.Pixels of the region acquire the projected area. Feature extraction is based on the distance between two nearby pixels. The distance between the region's boundaries is measured in perimeter (a scalar number). The area and perimeter of every object, regardless of shape or orientation, are stable and efficient once segmented. The length and width of fruits and vegetables are used to measure their size. Due to the fact that the shape of food items is frequently altered during processing, the orientation at which length and breadth are computed must be restored in a timely manner. The main axis is the object's longest line, determined by the distance between every two border pixels. The minor axis is the longest line traced across the object perpendicular to the major axis. Form is an important visual element for picture content description that is difficult to specify accurately due to the difficulty of measuring shape similarity. The two types of shape descriptors are region-based (based on the object's integral area) and contour-based (boundary segmented using local features). The roundness, aspect ratio, and compactness of a shape are all factors to consider.

3.12.3 Texture features

The texture is an excellent classifier for a wide range of images that are recognized and interpreted using human visual processes. Texture, as determined by a collection of pixels, indicates the distribution of elements and surface appearance, and is valuable in machine vision applications that predict surface roughness, contrast, entropy, and direction, among other things. The texture of the fruit is influenced by its age and sugar content (internal quality of fruits and vegetables). By extracting intensity values between pixels, it may also be used to separate various patterns in photos. Texture may be studied using both quantitative and qualitative methods. Six textural characteristics, namely contrast, coarseness, linelikeness, directionality, roughness, and regularity, were quantified. Contrast, correlation, entropy, and energy are four properties identified via qualitative analysis .Statistical texture, model-based texture, structural texture, and transform-based texture are examples of texture characteristics. Extract matrix based on intensity values of pixels (grey level cooccurrence matrix, grey level pixel run length matrix, and nearby grey level dependency matrix). Fractal model, random field model, an autoregressive model are examples of model-based textures. Lines and edges are created by the intensity of pixels in the structure texture. Spatial domain photos may be used to extract texture based on transforms. Statistical texture is widely utilised due of its low processing cost and great accuracy.The three various types of characteristics employed are colour, texture, mixture of both colour and texture which are then categorised by back propagation neural network. The best accuracy is achieved by combining texture and colour features. To identify the quality of mango production, a technique was suggested and algorithm that incorporates digital fuzzy image processing, content prediction, and statistical analysis. When compared to human expert sorting, this system design and construct an acceptable algorithm for recognising and sorting the mango with more than 80.00 percent accuracy. Adaptive network fuzzy interference system was used to categorise Mozafati dates based on weight and some geometric factors as a decision-making approach. The weight, length, breadth, and thickness of four date parameters are rated using a fuzzy algorithm, yielding a score of 93.50 percent.

3.14 Challenges

Due to their outstanding performance, rising cost reductions, ease of use, and algorithmic resilience, image processing and computer vision systems are scientific mechanisms in agriculture and industry. For quality evaluation of fruits and vegetables, traditional, multispectral, and hyperspectral computer vision systems are now widely employed. Colour, size, shape, texture, and defect are all frequent criteria that classic computer vision systems analyse (TCVS). Multispectral and hyperspectral computer vision systems improve TCVS by providing dynamic tools to a few problems that are difficult to identify with TCVS owing to the dominance of spectral pictures. Uneven distribution of light on the arch surface, strong wavelength selection for diverse applications, stem/calyx recognition, surface evaluation, and long duration are some of the hurdles to overcome in order to improve defect detection accuracy , exhausting of acquisition and processing for spectral image and different defects discrimination, etc. Instead of grading based on certain characteristics such as colour, size, form, or texture, other characteristics could be investigated to enhance the results. Furthermore, rather than applying the same weight value for all characteristics, changing the weight value may increase performance. Terahertz imaging, Raman imaging, and 3D techniques may all be used to assess the quality of fruits and vegetables.

CHAPTER-V

VI.SUMMARY AND CONCLUSION

External characteristics of fruits, such as colour, size, shape, texture, and other defects, are particularly significant for classification and grading. Machine vision, as well as the availability of low-cost hardware and software, are becoming more common as technology improves.Manual fruit sorting and grading has been replaced by automated machine vision systems due to software. The potential of non-destructive automation to generate accurate, quick, objective, and efficient results over manual effort could be the reason for its use. Machine vision will be the future of non-destructive testing, despite several hurdles have to be overcome.We can work on picture categorization for local fruits and veggies in the future. We can also create fruit and vegetable grading algorithms and machine Based on the foregoing methodologies, mobile applications can be developed for farmers and the general public to utilise for horticulture product identification, classification, and grading.

REFERENCES

[1] Xu Liming and Zhao Yanchao, "Automated strawberry grading system based on image processing," Computers and Electronics in Agriculture, vol. 71, no. Supplement 1, pp. S32-S39, April 2010.
[2] Yousef Al Ohali, "Computer vision-based date fruit grading system: Design and implementation," Journal of King Saud University - Computer and Information Sciences, vol. 23, no. 1, pp. 29-39, January 2011
[3] S.Arivazhagan, R.Newlin Shebiah, S.SelvNidhyanandhan, L.Ganesan," Fruit Recognition using Color and Texture Features", Journal of Emerging Trends in Computing and Information Sciences, VOL. 1, NO. 2, pp. 90-94, Oct 2010.
[4] Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis and Implementation of 89C51 Controller Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools & Technology. 2022; 12(2): 34–41p.
[5] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.
[6] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" SurajPunj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756
[7] G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" SurajPunj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:744-749
[8] G.Neelakrishnan, P.Iraianbu, T.Abishek, G.Rajesh, S.Vignesh, "IOT Based Monitoring in Agricultural" International Journal of Innovative Research in Science, Engineering and Technology, March 2020, Volume 9, Issue 3, pp:814-819
[9] G.Neelakrishnan, R.S.Jeevitha, P.Srinisha, S.Kowsalya, S.Dhivya, "Smart Gas Level Monitoring, Booking and Gas Leakage Detector over IOT" International Journal of Innovative Research in Science, Engineering and Technology, March 2020, Volume 9, Issue 3, pp: 825-836
[10] R.Srinivasan, G.Neelakrishnan, D.Vinoth and P.Iraianbu, "Design and Implementation of Novel Three Phase Multilevel Inverter for Smart Grid" International Journal of Multidisciplinary Educational Research, jan 2020, Volume 9, Issue 1(3) pp: 125-135
[11] Dr.C.Nagarajan, G. Neelakrishnan, V.Sundarajan, and D.Vinoth, "Simplified Reactive Power Control for Single-Phase Grid-Connected Photovoltaic Inverters" International Journal of Innovative Research in Science, Engineering and Technology, May 2015; 4(6): 2098-2104
[12] M.Kannan, R.Srinivasan and G.Neelakrishnan, "A Cascaded Multilevel H-Bridge Inverter for Electric Vehicles with Low Harmonic Distortion", International Journal of Advanced Engineering Research and Science, November 2014; 1(6): 48-52.
[13] G.Neelakrishnan, M.Kannan, S.Selvaraju, K.Vijayraj, M.Balaji and D.Kalidass, "Transformer Less Boost DC-DC Converter with Photovoltaic Array", IOSR Journal of Engineering, October 2013