Analysis of Fruits and Vegetable Conditions Using Image Processing

*Talupula Jahnavi[1], Renuka Devi M.N.[2] and

Punith Amilineni[3]

[1,2,3]*Department of CSE, Dayananda Sagar University, Bangalore, India*

Email: [1]*Jahnavitalupula3004@gmail.com*, [2]*renukadevi.m-cse@dsu.edu.in,* [3]*punith.chowdary56@gmail.com*

Abstract

This research presents a strong framework for automated fruit and vegetable quality inspection through advanced image processing techniques. Such applications range from defect detection to freshness assessment, enabling agriculture supply chains and retailers to classify produce into good and infected layers. It uses RGB masking, Otsu thresholding, and K-means clustering to develop a very efficient segmentation and feature extraction scheme. Feature computations such as defect area, mean intensity, and shape descriptors help isolate and classify the defective area. Apples, mangoes, and potatoes are analyzed in well-controlled lighting conditions for consistent results in this experimental setup. The proposed system embodies major advantages like scalability, accuracy, and quick response while maintaining diversity in product types and lighting conditions. Results state the efficient defect isolation and classification, while visualization through a binary and masked image manifests a clear understanding of the work. Future work includes the proposal of integrating machine-learning-enhanced adaptability and performance over various agricultural datasets to enhance quality assurance methods in food supply systems.

Keywords: Image Processing, Computer Vision, Segmentation, Thresholding, Feature Extraction, Otsu thresholding, RGB Masking, K-means Clustering, Defect Detection.

1. Introduction

The quality and freshness of fruits and vegetables matter immensely for consumer health and satisfaction, as well as for minimizing food wastage. The traditional means of inspection, heavily dependent on manual processes, which are slow and prone to human error, enjoy doctoral support from computer vision technique automated systems delivering a scalable and efficient alternative for quality assessment [1, 5, 18]. These technologies not only improve the quality of control but also minimize food loss during transport and storage. With the rise in global demand for fresh produce, there is an urgent need to set up efficient systems that can assess the quality of fruits and vegetables in real time. The need for research into an automated system capable of real-time identification of defects has led to extensive investigations into developing efficient technologies to locate visual defects, discoloration, and rot [8, 12, 19]. These technologies improve quality control and reduce food losses during transport and storage. Image processing, machine learning, and artificial intelligence have become the basis of modern agricultural technology. The systems can discover subtle defects through color, texture, and shape analysis normally invisible to the human eye[4, 7, 10, 16, 20]. This study focuses on creating a potent framework for checking the quality of fruits and vegetables, precise enough for conducting tests on potatoes, bananas, apples, and mangoes [3, 9, 17]. The study employs segmentation, thresholding, and feature extraction on the quality assessment protocol for process automation, thus providing an unparalleled and precise solution to agri-industries.

2. Flow Diagram

N. Vyas et al. (eds.), *Proceedings of the International Conference on Advancements in Computing Technologies and Artificial Intelligence (COMPUTATIA 2025)*, Advances in Intelligent Systems Research 189,
https://doi.org/10.2991/978-94-6463-700-7_21

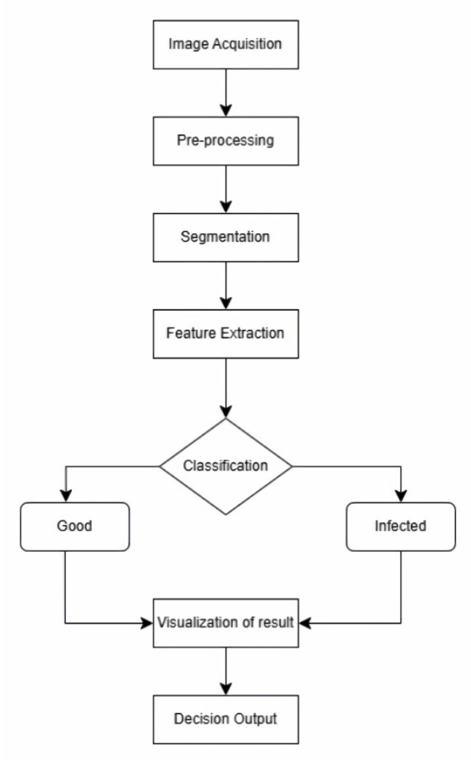

Figure 1. Image processing pipeline for classification

The flowchart illustrates the step-by-step process of image-based classification, starting from image acquisition, pre-processing, segmentation, and feature extraction. The extracted features are then used for classification, determining whether the sample is Good or Infected. The results are visualized before generating the final decision output (Figure 1).

3. Methodology

3.1 Image Acquisition

- High-resolution pictures of fruits and vegetables by a digital camera in controlled lighting conditions [6, 11, 15].
- The images are stored in standard formats, such as JPEG or PNG, appropriate for the required processing algorithms

3.3 Segmentation Techniques

K-means Clustering: Used for segregation of the image into different regions based on pixel intensity or similarity in color. It helps separate relevant objects like infected areas [2, 9, 16].

- Segmented regions are labeled with overlay, which can help in pinpointing defects.

Formula:

$$J=\sum_{i=1}^{k}\sum_{x_j\in C_i}||x_j-\mu_i||^2 \quad (1)$$

- Where 'J' is the sum of squared distances of data points, $'x_j'$ is the data point, $'\mu_i'$ is the centroid, 'k' is the number of clusters and 'n' is the total number of data points.

3.4 Thresholding Techniques

Otsu Thresholding: This method refers to maximizing the variance between classes and finding the threshold value that best separates infected and healthy regions [2, 15, 20].

The optimal threshold TTT is determined by maximizing the between-class variance, denoted as σ_B^2(T), which is given by:

$$\sigma_B^2(T)=\omega_0(T)\omega_1(T)(\mu_0(T)-\mu_1(T))^2(2)$$

- where $\omega_0(T)$ and$\omega_1(T)$ represent the probabilities of the two classes separated by the threshold T, and $\mu_0(T)$ and $\mu_1(T)$ are their respective means.

Color Thresholding: Isolating defects through the thresholding of red, green, and blue channels based on unique color intensities [13, 18].

Example thresholds:

- For identifying damaged parts:
  - Red Channel: Pixels with R<100
  - Green Channel: Pixels with G>140
  - Blue Channel: Pixels with B<170
  - Spongy tissue was differentiated from denser regions in X-ray images.
- On the basis of the coloration in the following thresholds, overripe or bruised areas were isolated:
  - Red Channel: Where R>150
  - Green Channel: Where G<100
  - Blue Channel: Where B<100
  - These do isolate portions of material exhibiting some discoloration that represents overripe or bruised variations.
- The infected areas, given by using Otsu's threshold method, were fixed dynamically by activating the binary optimal threshold (T)

$$T=arg\max_{T}[\omega_0(T)\omega_1(T)(\mu_0(T)-\mu_1(T))^2](3)$$

- where $\omega_0(T)$ and $\omega_1(T)$ represent the probabilities of the two classes separated by the threshold T, and $\mu_0(T)$ and $\mu_1(T)$ are their respective means.

Spongy Tissue in x-ray images:

Grayscale thresholding isolates the pixels that are having an intensity level of T=150.

This threshold differentiates spongy tissue from denser regions in X-ray imagery, based on intensity variations.

3.5 Masking and Feature Extraction

- Logical operations (AND, OR) are used to combine masks for precise defect isolation.

- Extracted features are:
- **Area:** The pixel count of defective regions.
- **Mean Intensity:**Average pixel value within the defect mask.
- **Shape Descriptors:** Metrics like eccentricity and circularity to characterize defect shapes [10, 17].

3.6 Classification

- Theinfected area is computed by summing all the pixels under the mask, and if itsvaluecrosses a predefined threshold (e.g., (t = 10000)), the produce is classified as "Infected" [7, 14].
- The total number of infected individuals, denoted as $A_{infected}$, is calculated as the sum of mask values across all individuals, given by:

$$A_{infected} = \sum_{i=1}^{n} Mask(i)(4)$$

- where$Mask(i)$ represents the infection status of the 'i'th individual. This summation provides the total count of infected individuals within the population.

3.7 Visualization of Results

- Masks are also laid out on the original images to enhance readers' interpretation of the infected area.
- Binary and masked images were created to validate the results.

3. Challenges & Observations

- The lighting must be consistent to avoid misclassification [11, 19].
- Threshold values may require tuning to adjust for variability in produce types and image quality [6, 18].
- Mixed produce in real-world scenarios may give rise to added complexity [12, 20].

4. Results & discussion:

4.1 Potatoes

- RGB masking accurately located the damaged areas resulting from physical injury or decay [3, 16].
- Segmented images allowed for the accurate classification into "Good" or "Infected," achieving high reliance.

4.2 Apples

- Using Otsu thresholding, the infected areas were effectively separated from the healthy regions.
- Complementary masks enabled the extraction of the affected regions for visual confirmation [1, 14, 17].

4.3 Mangoes

- Owing to segmentation, these spongy tissues hot-spotting the infected mangoes can help with expedient defect detection before these can further develop.
- X-ray images were tidily processed using grayscale thresholding, showing the differentiation of healthy and damaged tissues within the host mangoes [5, 13, 18].

4.4 Obtained outputs:

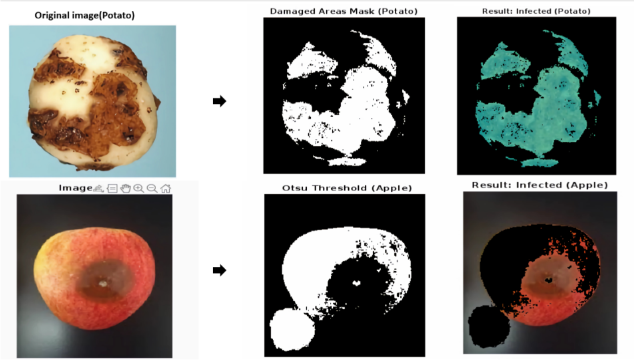

Figure 2. Input Images of Potato and Apple with Segmentation Results

The input image represents the original dataset, showcasing key features such as visible defects or structural patterns, with preprocessing steps like resizing or noise reduction applied for standardization. The output image shows results from employing methods including RGB thresholding, pinpointing the infected or damaged areas, and preserving the healthy ones. Segmentation was efficient in picking defects- evidence quantitatively highlighted was in support of this, with minor limitations like edge inaccuracy still being present (see Figure 2 and Figure 3).

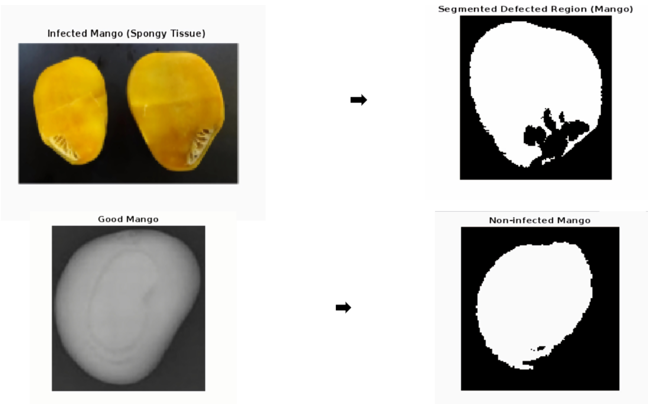

Figure 3. Thresholding Results for Mango Using Otsu's Method

4.5 Analysis of K-means Image Segmentation and RGB Color Clustering for Pixel Grouping

The Otsu method of image thresholding determines the optimal threshold, which separates the foreground from the background automatically. The threshold chosen maximizes variance between classes of pixel intensities. The red vertical line in the histogram is drawn at the optimal threshold, which separates pixel intensities into two regions: the background and the foreground. The histogram indicates the distribution of pixel intensities within the image, whereas the threshold is going to be responsible for the segmentation of the image into a binary mask (see Figure 4).

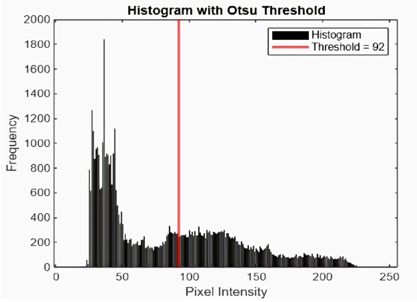

Figure 4. Histogram of Pixel Intensity with Optimal Threshold (Otsu's Method)

This 3D scatter plot depicts clustering of image pixels in RGB space, where a single point represents a pixel in which its coordinates correspond to red, green, and blue values, respectively. The colors of the points are representative of the clusters defined by K-means segmentation. Red-filled markers interpret centroids for each cluster, representing the average RGB values. This visualization facilitates anvisualization facilitates an understanding of how the image overnights pixels aggregate into different clusters based on color similarity (see Figure 5).

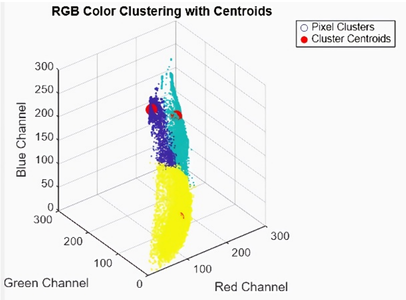

Figure 5. RGB Color Clustering Graph Using K-means Segmentation

Conclusion

This study thus presents a reliable and efficient framework for the vague, visual assessment of fruits and vegetables through image processing techniques. The segmentation, thresholding, and feature extraction techniques used indeed aim to encompass the generalities of the proposed system in a computerized quality assessment with a highly accurate output. Other types of produce also came under analysis as evidence that the method could be broadly applied to the whole agricultural supply chain. The integration of machine learning algorithms into the future may further extend this system's adaptability and accuracy across various datasets [4, 8, 15, 20].

References

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Pearson, 2018.

[2] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[3] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.

[4] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.

[5] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146–165, 2004.

[6] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient K-means clustering algorithm: Analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 881–892, 2002.

[7] D. Mery, F. Pedreschi, A. Soto, J. M. Aguilera, and A. Cipriano, "Automated detection of defects in agricultural produce using image processing," *Journal of Food Engineering*, vol. 67, no. 2, pp. 123–135, 2004.

[8] D. Zhang and J. Wang, "Image segmentation based on K-means clustering," in *Proceedings of the International Conference on Image Processing*, IEEE, pp. 4047–4050, 2004.

[9] Q. Huang, J. Wang, and G. Wan, "Machine vision systems for agricultural quality assessment," *Trends in Food Science & Technology*, vol. 81, pp. 200–210, 2018.

[10] M. H. Yap and T. Wang, "Image segmentation with statistical clustering techniques," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, IEEE, pp. 580–585, 2014.

[11] S. Sarkar and D. Wolf, "Advances in computer vision for agriculture," *Computers in Agriculture and Natural Resources*, vol. 56, no. 4, pp. 112–119, 2019.

[12] J. C. Russ, *The Image Processing Handbook*, CRC Press, 2016.

[13] K. Prasad and K. V. Prabhu, "Image processing applications in agricultural quality assessment: A review," *Biosystems Engineering*, vol. 167, pp. 1–16, 2017.

[14] A. Sharma and P. Jain, "Fruit defect detection using image processing: A review," *Artificial Intelligence in Agriculture*, vol. 3, pp. 20–35, 2020.

[15] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[16] S. Basar, et al., "Unsupervised color image segmentation: A case of RGB histogram-based K-means clustering initialization," *PLOS One*, vol. 15, no. 10, p. e0240015, 2020.

[17] S. M. Burney, M. Aqil, and H. Tariq, "K-means cluster analysis for image segmentation," *International Journal of Computer Applications*, vol. 96, no. 4, 2014.

[18] T. Acharya and A. K. Ray, *Image Processing: Principles and Applications*, Wiley-Interscience, 2005.

[19] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2011.

[20] Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition*, vol. 29, no. 8, pp. 1335–1346, 1996.