COMPUTATIONAL VISION FOR TOMATO CLASSIFICATION USING A DECISION TREE ALGORITHM

Caroline S. da Fonseca[1], Bilton G. Nhantumbo[1], Yuri M. Ferreira[1], Layana A. da Silva[1], Anderson G. Costa[1*]

[1*]Corresponding author. Programa de Pós-Graduação em Engenharia Agrícola e Ambiental/Universidade Federal Rural do Rio de Janeiro/Seropédica - RJ, Brasil.E-mail: acosta@ufrrj.br | ORCID ID: https://orcid.org/0000-0003-0594-8514

KEYWORDS

fruit ripeness, color models, machine learning, fruit acceptability.

ABSTRACT

Computer vision systems combined with machine learning techniques have demonstrated success as alternatives to empirical methods for classification and selection. This study aimed to classify tomatoes based on their colorimetric characteristics, which influence consumer purchasing potential, using the decision tree algorithm. Tomatoes were categorized into two classes based on ripeness: Higher Purchasing Potential (20 fruits) and Lower Purchasing Potential (40 fruits). Images were captured in the RGB color model and converted to HSI and CIELab models. Principal component analysis was employed to evaluate the influence of colorimetric characteristics within each class, and the decision tree algorithm was applied to classify the fruits into the respective categories. Tomatoes in the Higher Purchasing Potential class were primarily influenced by red intensity and chromaticity a and b, while tomatoes in the Lower Purchasing Potential class were influenced by green intensity and hue. The decision tree achieved an accuracy of 83.6% and an F1-score of 90.9%, demonstrating its potential for classifying tomatoes based on colorimetric characteristics linked to consumer preferences.

INTRODUÇÃO

Quality inspection is a crucial step in the production and marketing of tomatoes. Physical attributes such as shape and color are key determinants of consumer acceptability and food safety (Zhang et al., 2023). Agro-industries must adhere to established criteria and specifications to ensure product integrity.

Manual fruit inspection throughout the agro-industrial process is prone to human error, leading to inconsistencies in standardization, incorrect classification, and delays in processing. This is particularly challenging as it requires sustained attention, often resulting in subjective evaluations (Zhang et al., 2018). Improper transportation and storage of tomatoes, especially without separation based on ripeness, exacerbate economic losses across the production chain. Riper tomatoes release ethylene, a climacteric hormone that accelerates ripening across the batch, leading to reduced shelf life and increased susceptibility to fungal and bacterial diseases (Zakriya et al., 2023; Li et al., 2022).

For consumers, ripeness-related attributes such as color, texture, and firmness are critical indicators of quality (Andreuccetti et al., 2005; Brandão et al., 2021). However, reliance on empirical evaluations during marketing contributes to variability in product conformity. Automated selection methods offer a solution to these limitations across agriculture, agroindustry, and retail sectors (Silva et al., 2021). Computer vision systems enhance inspection efficiency, reduce costs, and ensure more consistent standardization, improving the quality control process (Tian et al., 2020). Image acquisition paired with machine learning algorithms allows for precise and efficient analysis, surpassing the limitations of manual inspection (Cubero et al., 2011).

Among machine learning techniques, decision trees (DTs) are particularly notable for their fast data processing, scalability, ability to interpret complex data relationships, and robust classification accuracy (Almeida et al., 2021). Decision trees are intuitive, with a computation process comparable to human reasoning, making them widely used

in classification models (Kotsiantis, 2011). They offer reliable and effective classification while providing simple, interpretable knowledge representations (Podgorelec et al., 2002).

Decision tree algorithms have been successfully applied to various agricultural contexts, such as classifying raisins by size, shape, color, and texture (Raihen & Akter, 2024), identifying data types based on morphological and colorimetric features (Özaltın, 2024), and detecting fruit contamination using digital imaging and chemical gas sensors (Sattar et al., 2024).

Given these advantages, the present study aims to employ the decision tree algorithm to classify tomatoes based on their ripeness condition for consumption using colorimetric characteristics obtained through a computer vision system.

MATERIAL AND METHODS

Tomato fruit characterization and image acquisition

The tomatoes used in this study were sourced from the Central-South region of Rio de Janeiro, Brazil, specifically from the municipality of Paty do Alferes (22°25'10" S latitude, 43°25'21" W longitude; average altitude: 610 m). The fruits belonged to the Italiano cultivar (*Solanum lycopersicum* L.), cultivated in greenhouses and manually harvested 85 days after planting.

According to CEAGESP standards (2003), the tomatoes were classified based on their colorimetric characteristics associated with ripeness: immature (green color), colored (a mix of green and red hues), and mature (more than 90% red color).

For the analysis, 20 fruits were selected per ripeness stage. Since tomatoes with a ripe red appearance are the predominant preference for consumers at the time of purchase (Andreuccetti et al., 2005), fruits classified as mature (ripe red) were grouped into the Major Purchasing Potential (MPP) class. Conversely, tomatoes categorized as immature, or colored were grouped into the Low Purchasing Potential (LPP) class. The characteristics of tomatoes in each class are detailed in Table 1.

TABLE 1. Colorimetric aspects of tomatoes with Low and Major Purchasing Potential (LPP and MPP, respectively).

FRUIT	RIPENESS LEVEL	CLASS
	Immature (Green fruit)	Low purchasing potential (LPP)
	Intermediary (variation between green and red fruit)	
	Mature (red fruit)	Major purchasing potential (MPP)

An experimental setup (Figure 1) was used to capture images of tomatoes. The apparatus consisted of a digital camera capable of capturing images within the visible spectral range (Canon Powershot G9 X RGB) and two 100 W halogen lamps. The camera was mounted on a tripod and positioned 0.24 m above the bench, while the lamps were set at a height of 0.84 m. To minimize shading effects, the bench was covered with a white background.
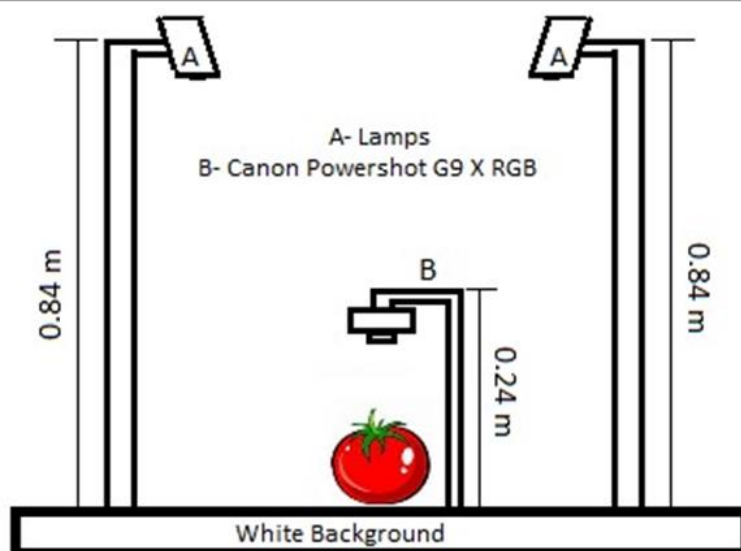
FIGURE 1. Schematic diagram of the image acquisition setup for tomatoes.

Images of each tomato fruit were processed using ImageJ software to extract the average values of the red (R), green (G), and blue (B) bands. These RGB values were then converted into the HSI and CIELab color models following the method of Pedrini & Schwartz (2008). The conversion allowed for the characterization of key colorimetric attributes. Hue (H) represents color in the HSI model, with values ranging from 0° to 360°; luminance (L) corresponds to brightness in the CIELab model, with values ranging from 0 to 100; chromaticity (a and b) stands for color ranges in the CIELab model, where a varies from green (-128) to red (+128) and b from blue (-128) to yellow (+128).

Data analysis

The colorimetric characteristics of tomatoes across the four ripeness stages were subjected to exploratory analysis using PAST 4.03 software. The analysis evaluated key statistical parameters, including maximum, minimum, mean, and median values. Additionally, the coefficient of variation, kurtosis, and skewness were calculated to assess data variability and distribution patterns. These metrics were classified according to the criteria outlined in Table 2.

TABLE 2. Classification of coefficient of variation (CV), skewness (Cs), and kurtosis (Ck) values for colorimetric characteristics of tomatoes at four degrees of ripeness.

Coefficient of variation (CV)	Skewness (Cs)	Kurtosis (Ck)
CV < 12 % (low)	Cs > 0 (to the right)	Ck < 0 (platykurtic)
12 % < CV < 60 % (medium)	Cs < 0 (to the left)	Ck = 0 (mesokurtic)
CV > 60 % (high)	Cs = 0 (symmetrical)	Ck > 0 (leptokurtic)

Each colorimetric characteristic was evaluated for normality using the Shapiro-Wilk test at a significance level of 5%. Colorimetric characteristics that met the criteria for normality were further analyzed using analysis of variance (ANOVA) and the Student's t-test for independent samples, also at a 5% significance level, to compare the means of the

Major Purchasing Potential (MPP) and Low Purchasing Potential (LPP) classes.

Additionally, using PAST 4.03 software, the data were subjected to a multivariate principal component analysis (PCA). This analysis assessed the dispersion of fruits within a two-dimensional plane defined by the most relevant principal components and evaluated the influence of colorimetric characteristics on the MPP and LPP classes.

Development and testing of the Decision Tree (DT)

The decision tree (DT) model was developed in R software (version 4.3.2) to classify tomatoes as belonging to the Major Purchasing Potential (MPP) or Low Purchasing Potential (LPP) classes based on the most relevant colorimetric characteristics. The Part algorithm (Recursive Partitioning and Regression Trees), implementing the CART (Classification and Regression Trees) technique, was used for this purpose.

The Gini index, which maximizes class homogeneity (Tangirala, 2020), was employed to determine the importance of each colorimetric characteristic in data classification. The pruning technique (Gomes Mantovani et al., 2024) was applied to optimize the complexity parameter (cp) of the DT and generate the most accurate model. This was achieved by analyzing the smallest relative error and identifying its stabilization trend in the auxiliary pruning graph.

The dataset, comprising colorimetric characteristics, was randomly divided into training and test sets to ensure representativeness and minimize bias. The training set constituted 75% of the data (15 MPP fruits and 30 LPP fruits), while the test set comprised 25% of the data (5 MPP fruits and 10 LPP fruits). The graphical representation of the trained DT illustrated its structure, including nodes, splits, and resulting classifications.

To evaluate model performance, a confusion matrix was generated, presenting classification hits and errors in the test dataset. Performance metrics, including accuracy (closeness of results to the correct classification, Equation 1), precision (proximity of obtained results, Equation 2), recall (frequency of correctly identified examples of a given class, Equation 3), and F1-score (harmonic mean of precision and recall, Equation 4), were derived from the confusion matrix.

$$Accuracy = \left[\frac{tp + tn}{tp + fp + fn + tn}\right] * 100 \tag{1}$$

$$Precision = \left(\frac{tp}{tp+fp}\right) * 100 \tag{2}$$

$$Recall = \left(\frac{tp}{tp+fn}\right) * 100 \tag{3}$$

$$F1\text{-}score = \left[\frac{(2 * (precision * recall))}{(precision + recall)}\right] * 100 \tag{4}$$

Where:

tp - true positive values regarding the analyzed class;

tn - true negative values regarding the analyzed class;

fp - false positive values regarding the analyzed

class; and

fn - false negative values regarding the analyzed class.

RESULTS AND DISCUSSION

The exploratory analysis (Table 3) revealed that the measures of central tendency (mean and median) were close in 42.9% of the colorimetric characteristics, with differences less than or equal to 0.5. This indicates a data distribution with minimal deviation from the central value.

Regarding the coefficient of variation (CV), the colorimetric characteristics R, G, B, and b showed a variation classified as average. The colorimetric characteristics H and a exhibited a variation classified as high. All colorimetric characteristics had an asymmetry value of zero, indicating a symmetrical distribution of the data.

TABLE 3. Exploratory analysis of the colorimetric characteristics of tomatoes.

Var.	Min	Max	Ave	Med.	Cs	Ck	CV (%)	SW
R	80.0	160.0	121.0	119.0	0.0	-1.3	19.0	0.95*
G	12.0	104.0	53.5	54.5	0.0	-1.4	55.1	0.91*
B	5.0	24.0	11.9	9.0	0.8	-1.0	50.0	0.83*
H	2.0	61.0	27.7	25.5	0.2	-1.5	76.8	0.86*
L	29.3	45.0	34.9	34.5	0.8	0.7	9.5	0.96*
a	-2.9	47.5	24.0	24.4	-0.1	-1.6	77.3	0.86*
b	31.2	53.0	42.4	41.9	0.0	-1.1	13.9	0.96*

Variable (Var), Minimum (Min), average (Ave), median (Md), maximum (Max), coefficient of variation (CV%), skewness (Cs), kurtosis (CK), Shapiro-Wilk (SW), *Significant at 5% by T-test for normal distribution. Red band (R), green band (G), blue band (B), hue (H), luminance (L), and of two color ranges *a* and *b*. (*) Significant normality at 5% by Shapiro-Wilk test.

In terms of kurtosis, only the colorimetric characteristic L exhibited a mesokurtic distribution (value equal to zero). All other colorimetric characteristics demonstrated a platykurtic distribution (values less than zero), indicating flatter distributions compared to a normal curve.

The Shapiro-Wilk test confirmed that all colorimetric characteristics followed a significant normal distribution, allowing for univariate comparisons between the Major Purchasing Potential (MPP) and Low Purchasing Potential (LPP) classes.

The Student's t-test (Table 4) revealed that all colorimetric characteristics, except for L, significantly distinguished between the MPP and LPP classes. Among the RGB color model characteristics, the green (G) component showed the greatest significant difference between classes. When comparing chromaticity parameters from the CIELab color model, chromaticity a (measuring the variation between green and red) exhibited the most significant difference between the classes, making it the most reliable characteristic for identifying tomatoes based on consumer purchasing potential.

TABLE 4. Mean comparison of tomato classes with Major and Low Purchasing Potential (MPP and LPP) by Student's t-test for independent samples.

	MPP average	LPP average	Diference	t-Test
R	147.2	107.83	39.38	10.62*
G	18.45	71.03	52.58	12.16*
B	8.40	13.63	5.23	3.51*
H	3.75	39.68	35.93	10.25*
L	34.03	35.28	1.24	1.38[ns]
A	45.16	13.40	31.76	10.7*
B	49.02	39.15	9.86	9.91*

Red band (R), green band (G), blue band (B), hue (H), luminance (L), and chromaticity of color *a* and *b*. (*) Significant difference between classes at 5% significance level by t-Student test.

Principal component analysis (PCA) enables the simultaneous evaluation of characteristics that discriminate between classes, influencing the acceptability potential of the fruit (Bello et al., 2020). The two-dimensional projection in Figure 2 displays the three ripeness classes of tomato fruits concerning the first two principal components (PC1 and PC2). PC1 accounted for 79.43% of the total data variability and was responsible for the horizontal separation of classes along the x-axis. This component demonstrated its importance in distinguishing between MPP and LPP classes, highlighting its relevance in ripeness classification. PC2 explained 17.73% of the data variability, contributing information about vertical grouping along the y-axis. This component influenced variations associated with the intermediate ripeness class, reflecting a more gradual transition between ripeness categories.
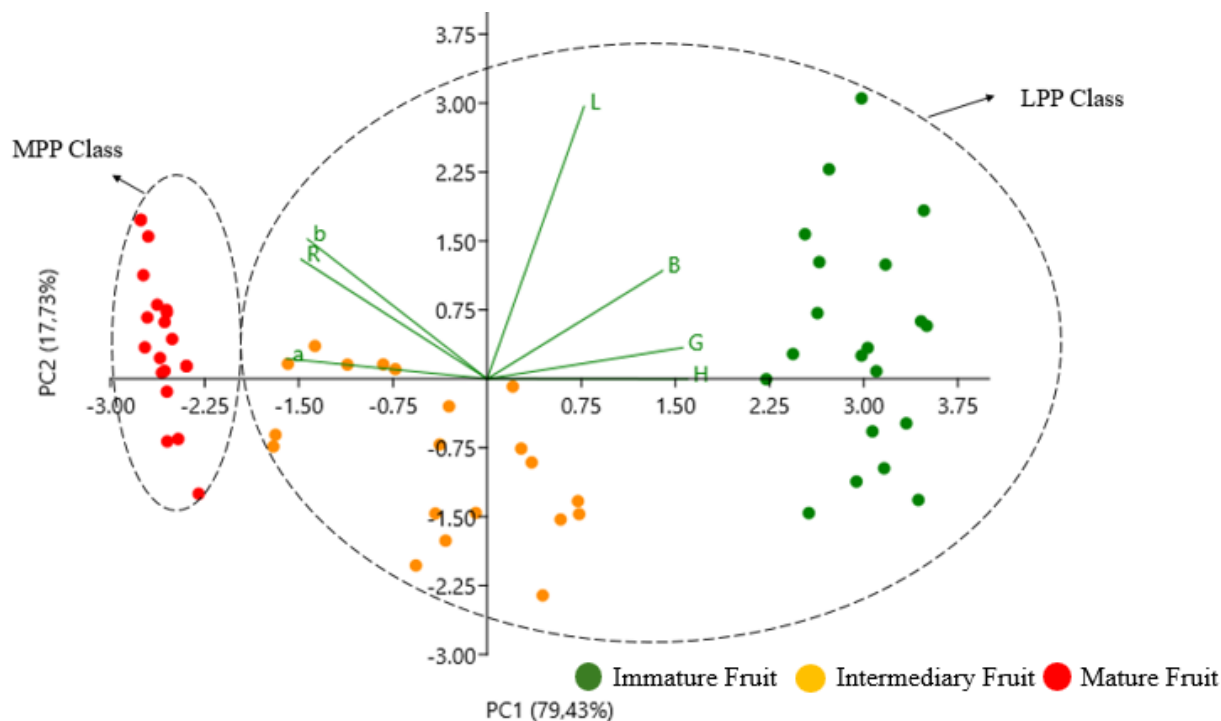

FIGURE 2. Two-dimensional scatterplot for colorimetric traits of MPP and LPP tomatoes from principal component analysis.

PC1 primarily distinguished mature tomatoes (on the left) from immature tomatoes (on the right), driven by specific colorimetric characteristics. The attributes H (hue) and G (green intensity) had the greatest positive influence on PC1, contributing to the dispersion of immature fruits. Conversely, characteristics R (red intensity), a (red chromaticity), and b (yellow chromaticity) were most influential for the MPP class, indicating these factors as key determinants in consumer fruit selection. PC2, complementing PC1, predominantly influenced the intermediate ripeness class. It was strongly impacted by L (luminance) and chromaticity b, suggesting that intermediate fruits displayed higher brightness and more intense yellow tones.

Fruit color remains one of the most significant quality parameters influencing consumer purchasing decisions. As highlighted by Andreuccetti et al. (2005), red and uniform tomato coloration is a decisive factor for consumer acceptance. In this study, the fruits in the MPP class were predominantly characterized by the red color. This was evidenced by the higher intensity of R values compared to the LPP class and the average hue values that align with the red color range in the three-dimensional representation of the HSI model (Pedrini & Schwartz, 2008). Additionally, the chromaticity *a* values from the CIELab model, which were higher and positive for the MPP class, further confirmed the predominance of red over green in the fruits classified in this category.

Furthermore, consumer preference for fresh tomatoes demands that harvesting and post-harvest stages adhere to strict quality control standards to preserve the physical quality of the fruits. This ensures a higher quantity of tomatoes with an acceptable appearance, particularly uniform color and the absence of injuries are suitable for commercialization.

A significant challenge is the lack of uniformity in tomato ripening, which negatively impacts production and economic efficiency, particularly in mechanized harvesting systems (Machado et al., 2018). Prematurely harvested or overly ripe tomatoes have lower market value, increased susceptibility to mechanical damage, and shorter shelf life, resulting in waste throughout the supply chain (Umeohia & Olapade, 2024). Ripening heterogeneity undermines the standardization required to meet consumer market demands, necessitating more intensive manual selection or disposal processes and escalating operational costs (Yang et al., 2023).

In this context, selection methods based on colorimetric characteristics, such as the approach presented in this study, enable accurate and instantaneous classification (Silva et al., 2021). These methods contribute to increased production efficiency, reduced food losses, and the commercialization of higher-quality products. As highlighted by Cubero et al. (2011), computer vision systems improve the inspection quality of fruits and vegetables, with potential applications in classification

systems, the estimation of physicochemical parameters, monitoring fruit processes during storage and evaluating experimental treatments.

In the stages of fruit production and marketing, machine learning algorithms, such as DT, have been increasingly applied for data analysis, pattern recognition, and the development of classification and quality control models (Liu & Xu, 2023).

In this study, the construction of the DT considered the response of the complexity parameter (cp), which evaluates the occurrence of overfitting and limits tree growth (Figure 3). The optimal cp value of 0.019 corresponded to a tree with four nodes, representing the best structure for this case by balancing simplicity and accuracy.
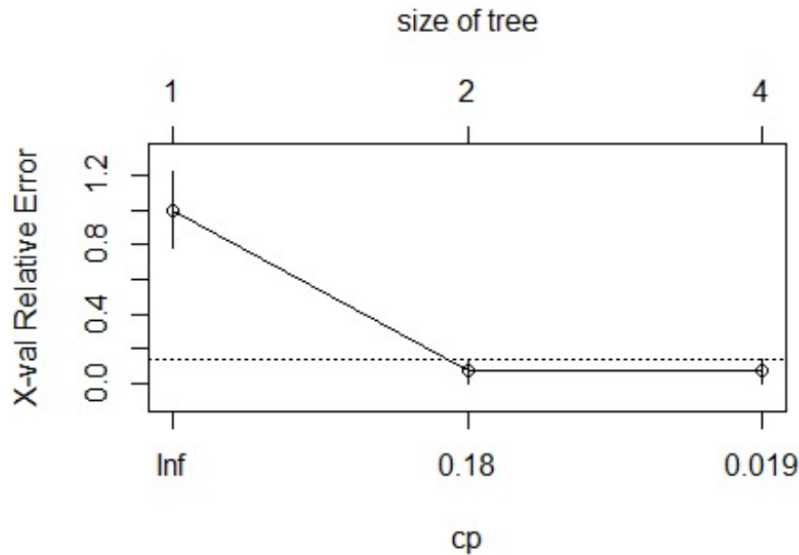

FIGURE 3. Complexity parameter (cp) plot for tomato classification decision tree.

The decision tree generated from the training set (Figure 4) highlighted the importance of the colorimetric characteristics R, G, and B based on the Gini index, which determines the relevance of variables in the tree's composition. The colorimetric characteristic G was allocated at the first node, indicating that 67% of the fruits in the training set could be classified as MPP when G > 71.

At the second node, the characteristic R demonstrated that 7% of the fruits could be classified as LPP when G < 71 and R < 86. Finally, at the third node, the characteristic B differentiated further classifications: 2% of the fruits were classified as MPP when G < 71, R > 86, and B < 19; 31% of the fruits were classified as LPP when G < 71, R > 86, and B > 19.
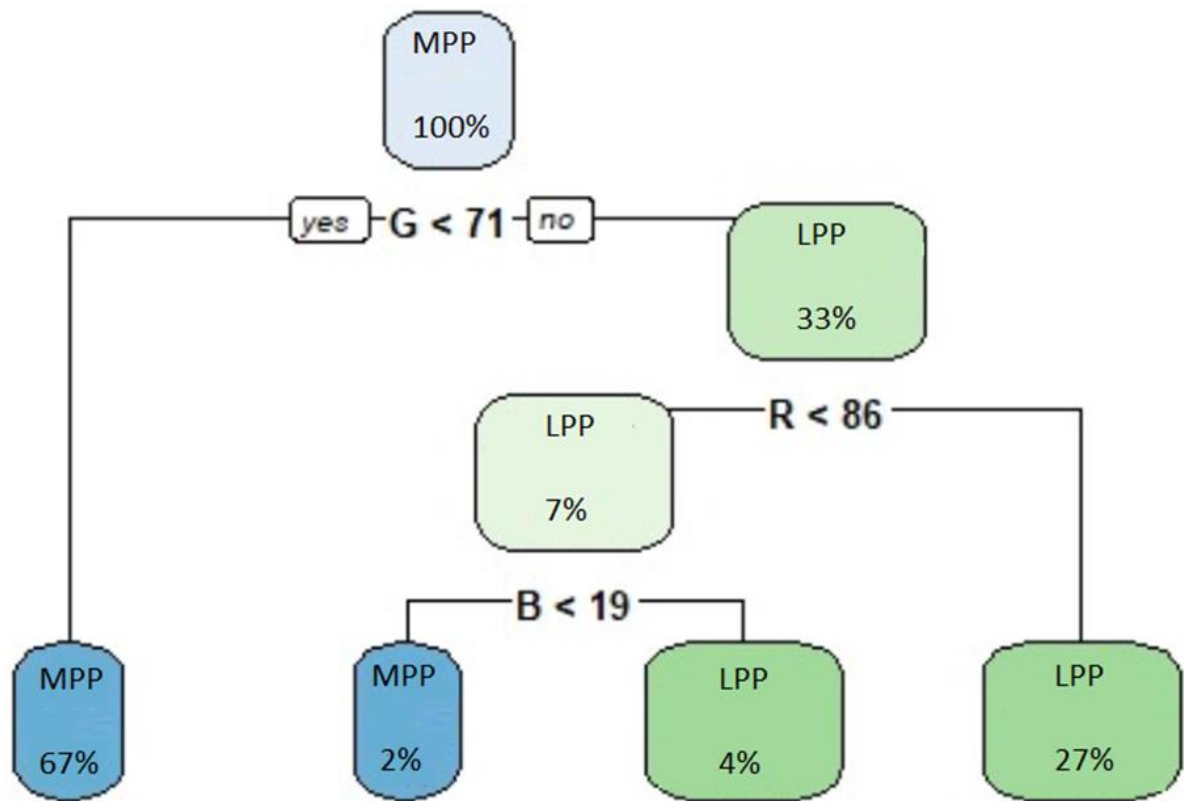

FIGURE 4. Decision tree for classification of MPP and LPP tomatoes.

The results from the confusion matrix (Table 5) for the test sample set demonstrated effective classification performance: LPP Class (10 out of 10 fruits were correctly classified); and MPP Class (3 out of 5 fruits were correctly classified). This corresponds to an overall accuracy of 86.7%. Additional performance metrics indicated an accuracy of 83.3%, reflecting the model's ability to classify samples correctly across both classes. A classification frequency of 100.0% for each class indicates that the model successfully identified instances within both MPP and LPP categories. A global F1-score of 90.9% emphasizes the model's balance between precision and recall. These results validate the potential of the generated decision tree to accurately classify tomato fruits based on consumer preference parameters, highlighting its effectiveness for practical applications in fruit quality assessment.

TABLE 5. Confusion matrix and classification performance metrics generated by the decision tree (Low Purchasing Potential – LPP: 10 and Major Purchasing Potential – MPP: 5).

Confusion matrix		
Forecasts	LPP	MPP
LPP	10	2
MPP	0	3
Performance metrics		
Accuracy	86.7 %	
Precision	83.3 %	
Recall	100.0 %	
F1	90.9 %	

The performance parameters demonstrated that the use of DT was the most efficient method for class identification in this study. This research further confirms the potential of DT algorithms for fruit classification. The accuracy achieved in this study aligns with results from similar applications, such as the detection of diseases in papaya (Habib et al., 2020), the identification of data types (Özaltın, 2024), and the classification of raisins (Raihen & Akter, 2024), all of which reported accuracy levels exceeding 75%.

Automated technologies, including computer vision and artificial intelligence systems, provide innovative solutions to address challenges in fruit classification and selection (Tian et al., 2020). Integrated with automated sorting and harvesting equipment, these systems enable precise and rapid identification and separation of fruits based on visual characteristics like color and texture (Jun et al., 2021; Chakraborty et al., 2023). Such advancements reduce waste, enhance operational efficiency, and yield greater financial returns for producers, displaying their transformative impact on the agricultural sector.

The superiority of automated inspection over manual inspection lies in its ability to deliver greater efficiency, standardization, and reliability in fruit quality assessment processes (Kaur et al., 2018). While manual inspection is prone to challenges that compromise productivity and quality, computerized inspection streamlines the evaluation process, offering a cost-effective return on investment in automation.

The integration of computer vision with DT algorithms not only optimizes fruit classification but also enhances automated agricultural production processes (Tian et al., 2020). By ensuring the selection of high-quality products, these technologies benefit the agricultural industry and simultaneously improve consumer satisfaction and safety.

In conclusion, the application of computer vision and machine learning represents a promising and innovative solution to enhance quality, efficiency, and reliability in tomato production and other agricultural sectors, driving significant advancements and progress in the industry.

ACKNOWLEDGEMENTS

CONCLUSIONS

The colorimetric characteristics R, G, B, H, *a*, and *b* demonstrated the ability to differentiate between the HPP and LPP classes even in a univariate context. Among these, the intensities of the R, G, and B bands were selected to construct the decision tree using the Gini index criterion.

The principal component analysis further highlighted the separation of tomato classes on a two-dimensional plane. Tomatoes in the HPP class were primarily influenced by R (red intensity) and the chromaticity *a* and *b*. In contrast, tomatoes in the LPP class were well-separated, with G (green intensity) and H (hue angle) emerging as the most significant characteristics.

Overall, the combination of computer vision systems with the decision tree algorithm proved to be an effective method for tomato quality classification. Performance metrics including accuracy (0.836), precision (0.833), recall (1.0), and F1-score (0.909) indicate that the decision tree-based learning model has strong potential for classifying tomatoes based on characteristics aligned with consumer preferences.

REFERENCES

Almeida, R. D. C., Neto, W. V., Silva, V. B. D., Carvalho, L. C. B., Lopes, Â. C. D. A. & Gomes, R. L. F. (2021). Decision tree as a tool in the classification of lima bean accessions. *Revista Caatinga,* 34(2), 471-478. https://doi.org/10.1590/1983-21252021v34n223rc

Andreuccetti, C., Ferreira, M. D. & Tavares, M. (2005). Perfil dos compradores de tomate de mesa em supermercados da região de Campinas. *Horticultura Brasileira,* 23(1), 148 - 153. https://doi.org/10.1590/S0102-05362005000100031

Bello, T. B., Costa, A. G., Silva, T. R. D., Paes, J. L. & Oliveira, M. V. M. (2020). Tomato quality based on colorimetric characteristics of digital images. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 24(8), 567-572. https://doi.org/10.1590/1807-1929/agriambi.v24n8p567-572

Brandão, C. P., Oliveira, M. S. P., Santos, J. C. & Oliveira Junior, M. C. M. (2021). Perfil e preferências do consumidor de frutos de pupunha da cidade de Belém, Pará. *Research, Society and Development,* 10(7), e28810716502. https://doi.org/10.33448/rsd-v10i7.16502

Chakraborty, S. K., Subeesh, A., Dubey, K., Jat, D., Chandel, N. S., Potdar, R., Rao, N. R. N. V. G. & Kumar, D. (2023). Development of an optimally designed real-time automatic citrus fruit grading–sorting machine leveraging computer vision-based adaptive deep learning model. *Engineering Applications of Artificial Intelligence*, 120(2023), 105826. https://doi.org/10.1016/j.engappai.2023.105826

Companhia de Entrepostos e Armazéns Gerais de São Paulo - CEAGESP. (2003). *Normas de classificação do tomate: Centro de Qualidade em Horticultura*. https://ceagesp.gov.br/wp-content/uploads/2015/07/tomate.pdf

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. & Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and bioprocess technology,* 4, 487-504. https://doi.org/10.1007/s11947-010-0411-8

Gomes Mantovani, R., Horváth, T., Rossi, A. L., Cerri, R., Barbon Junior, S., Vanschoren, J. & Carvalho, A. C. D. (2024). Better trees: an empirical study on hyperparameter tuning of classification decision tree induction algorithms. *Data Mining and Knowledge Discovery,* 38(3), 1364 - 1416. https://doi.org/10.1007/s10618-024-01002-5

Habib, M. T., Majumder, A., Jakaria, A. Z. M., Akter, M., Uddin, M. S. & Ahmed, F. (2020). Machine vision based papaya disease recognition. *Journal of King Saud University-Computer and Information Sciences,* 32(3), 300 - 309. https://doi.org/10.1016/j.jksuci.2018.06.006

Jun, J., Kim, J., Seol, J., Kim, J. & Son, H. I. (2021). Towards an efficient tomato harvesting robot: 3D perception, manipulation, and end-effector. *IEEE Access,* 9, 17631 - 17640. https://doi.org/10.1109/ACCESS.2021.3052240

Kaur, G., Uisan, K., Ong, K. L. & Lin, C. S. K. (2018). Recent trends in green and sustainable chemistry & waste valorisation: rethinking plastics in a circular economy. *Current opinion in green and sustainable chemistry,* 9, 30 - 39. https://doi.org/10.1016/j.cogsc.2017.11.003

Kotsiantis, S. (2011). Decision trees: a recent overview. *Artificial Intelligence Review*, 39, 261-283. https://doi.org/10.1007/s10462-011-9272-4

Li, S., Wu, P., Yu, X., Cao, J., Chen, X., Gao, L., Chen, K. & Grierson, D. (2022). Contrasting roles of ethylene response factors in pathogen response and ripening in fleshy fruit. *Cells,* 11(16), 2484. https://doi.org/10.3390/cells11162484

Liu, H. & Xu, Z. (2023). Machine vision and machine learning for plant phenotyping and precision agriculture. *Frontiers in Plant Science*, 14, e1331918. https://doi.org/10.3389/fpls.2023.1331918

Machado, T. A., Fernandes, H. C., Megguer, C. A., Santos, N. T. & Santos, F. L. (2018). Quantitative and qualitative loss of tomato fruits during mechanized harvest. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 22(11), 799-803. https://doi.org/10.1590/1807-1929/agriambi.v22n11p799-803

Özaltın, Ö. (2024). Date Fruit Classification by Using Image Features Based on Machine Learning Algorithms. *Research in Agricultural Sciences,* 55(1), 26-35. https://doi.org/10.5152/AUAF.2024.23171

Pedrini, H. & Schwartz, W. R. (2008). *Análise de imagens digitais: princípios, algoritmos e aplicações*. Thomson Learning.

Podgorelec, V., Kokol, P., Stiglic, B. & Rozman, I. (2002). Decision trees: an overview and their use in medicine. *Journal of Medical Systems*, 26(5): 445-463. https://doi.org/10.1023/A:1016409317640

Raihen, M. N. & Akter, S. (2024). Prediction modeling using deep learning for the classification of grape-type dried fruits. *International Journal of Mathematics and Computer in Engineering,* 2(1), 1–12. https://doi.org/10.2478/ijmce-2024-0001

Sattar, A., Ridoy, M. A. M., Saha, A. K., Babu, H. M. H. & Huda, M. N. (2024). A Comprehensive Approach to Detecting Chemical Adulteration in Fruits Using Computer Vision, Deep Learning, and Chemical Sensors. *Intelligent Systems with Applications,* 23, e200402. https://doi.org/10.1016/j.iswa.2024.200402

Silva, T. R. D., Costa, A. G., Paes, J. L., Oliveira, M. V. M. & Pinto, F. A. C. (2021). Comparação de parâmetros de qualidade obtidos por métodos destrutivos e óptico para classificação de tomates. *Revista Ciência Agronômica*, 52(3), e20207475. https://doi.org/10.5935/1806-6690.20210066

Tangirala, S. (2020). Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm. *International Journal of Advanced Computer Science and Applications*, 11(2), 612 - 619. https://doi.org/10.14569/IJACSA.2020.01 10277

Tian, H., Wang, T., Liu, Y., Qiao, X. & Li, Y. (2020). Computer vision technology in agricultural automation - A review. *Information Processing in Agriculture,* 7(1): 1 - 19. https://doi.org/10.1016/j.inp a.2019.09.006

Umeohia, U. E. & Olapade, A. A. (2024). Quality attributes, physiology, and Postharvest Technologies of Tomatoes (*Lycopersicum esculentum*)–A review. *AJFST*, 12(2): 42 - 64. https://doi.org/10.126 91/ajfst-12-2-1

Yang, Z., Amin, A., Zhang, Y., Wang, X., Chen, G. & Abdelhamid, M. A. (2023). Design of a Tomato Sorting Device Based on the Multisine-FSR Composite Measurement. *Agronomy,* 13(7), 778. http s://doi.org/10.3390/agronomy13071778

Zakriya, M., Hussain, A., Mahdi, A. A., Yasmeen, F., Kausar, T., Rehman, A., Yaqub, S., Fatima, P., Noreen, S., Kabir, K., Nisar, R., Gorsi, F. I., Fatima, H. & Korma, A. S. (2023). Effect of different types of ethylene scavengers used in different combinations, on the post-harvest quality and phytochemicals retention of tomatoes (*Solanum lycopersicum L.*). *Chemical and Biological Technologies in Agriculture,* 10(90), 1-15. https://doi.org/10.1186/s40538-023-00465-w

Zhang, B., Gu, B., Tian, G., Zhou, J., Huang, J. & Xiong Y (2018). Challenges and solutions of optical-based nondestructive quality inspection for robotic fruit and vegetable grading systems: A technical review. *Trends in Food Science & Technology,* 81, 213-231. https://doi.org/10.1016/j.tifs.2018.09.018

Zhang, J., Liu, S., Zhu, X., Chang, Y., Wang, C., Ma, N., Wang, J., Zhang, X., Lyu, J. & Xie, J. (2023). A comprehensive evaluation of tomato fruit quality and identification of volatile compounds. *Plants*, 12(16), 2947. https://doi.org/10.3390/plants12162947