Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.2, April 2026, pp. 1404~1415
e-ISSN: 2548-6861

Fruit Ripeness Analysis Using Colour Transition Matrix and Eigenvalues

Septiana Sulastri [1]*, Yufis Azhar [2]**
* Department of Informatics Engineering, Universitas Muhammadiyah Malang
sulastriseptiana24@webmail.umm.ac.id [1], yufis@umm.ac.id [2]

Article Info

Article history:

Received 2026-01-29
Revised 2026-03-16
Accepted 2026-04-10

Keyword:

Banana Ripeness,
CDTM,
Eigenvalue,
Image Processing.

ABSTRACT

Fruit ripeness classification plays an important role in maintaining fruit quality, reducing post-harvest losses, and ensuring consistent grading in agricultural supply chains. Bananas, particularly the Cavendish variety, are widely traded globally and require reliable methods for determining ripeness levels during distribution and storage. However, conventional visual inspection methods remain subjective and often lead to inconsistent evaluations. This study proposes an interpretable mathematical approach for fruit ripeness analysis based on the Colour Dominant–Transition Matrix (CDTM) combined with eigenvalue analysis. The dataset consists of 3,495 real images of Cavendish bananas obtained from an open-access dataset. Image preprocessing includes object segmentation, colour normalization, and noise reduction to ensure illumination consistency. Dominant colour features are extracted using K-Means clustering in the CIELAB colour space, and the CDTM is constructed to model spatial colour transition probabilities. The second-largest eigenvalue (λ_2) is used as a quantitative indicator of colour stability and homogeneity. Experimental results show that optimally ripe bananas produce the lowest λ_2 value (0.1045), indicating high colour uniformity, while unripe bananas exhibit the highest λ_2 value (0.7972). Using λ_2 as a single feature, an SVM classifier achieved an accuracy of 64%. Although the classification performance remains moderate, the proposed CDTM-based eigenvalue feature provides an interpretable and computationally efficient indicator suitable for non-destructive fruit ripeness monitoring systems.

I. Introduction

The development of digital image processing technology has contributed significantly to improving the efficiency and accuracy of quality assessment in agricultural products, including fresh fruits. Image-based quality inspection systems enable the physical evaluation of fruit to be performed automatically, rapidly, and consistently in a non-destructive manner [1]. This condition becomes increasingly relevant in response to the growing demand for high-quality fruit products within modern supply chains and smart agriculture practices.

However, fruit quality assessment methods that still rely on human visual inspection suffer from several limitations. This process tends to be subjective and is strongly influenced by lighting conditions, variations in colour perception among inspectors, and human fatigue [2]. These factors often lead to inconsistencies in quality evaluation and reduced assessment accuracy, particularly in large-scale post-harvest operations. Consequently, objective and reliable image-based approaches for fruit quality assessment are essential to support efficient and sustainable inspection systems.

Fruit quality changes are generally characterized by variations in physical attributes, particularly surface colour. Colour is one of the earliest and most easily observable indicators for determining ripeness stages, freshness levels, and potential decay. In various fruit commodities such as bananas, mangoes, papayas, and pears, colour changes occur gradually following natural physiological processes, progressing from unripe to ripe and eventually decayed stages [3]. As a result, colour analysis plays a crucial role in image-based fruit quality assessment systems. Numerous studies have explored colour-based features, including mean and variance of colour channels, colour histograms, and transformations into different colour spaces such as RGB, HSV, and LAB [4].

In addition to colour features, the combination of colour and texture descriptors has been widely adopted to improve classification performance. Colour features in the LAB colour space combined with texture descriptors have been employed for fruit quality classification using artificial neural networks (ANN) [5]. Texture-based features derived from Gray Level Co-occurrence Matrices (GLCM) have also been utilized for mango fruit quality classification [6], while combined colour and texture features have been applied for guava quality grading [7].

With recent technological advances, deep learning-based approaches have become increasingly popular for fruit quality assessment. Convolutional neural network architectures such as DenseNet-121 have demonstrated high accuracy in fruit ripeness classification tasks [8]. Despite their strong performance, most deep learning models operate as black-box systems and provide limited interpretability regarding the underlying dynamics of colour changes on fruit surfaces. These methods typically yield final classification decisions without explicitly explaining the gradual colour transition processes associated with fruit ripening and decay.

Fruit physiology studies indicate that ripening and decay processes occur progressively and spread spatially to surrounding tissues rather than occurring simultaneously across the entire fruit surface [9]. As a consequence, surface colour changes exhibit spatial non-uniformity, which cannot be adequately captured using global colour statistics such as histograms or mean colour values alone.

Several studies have focused on dominant colour extraction using clustering-based and segmentation-based techniques. Dominant colour extraction methods based on perceptually relevant colour features have been proposed using K-means clustering in the CIELAB colour space combined with region adjacency graph (RAG) analysis [10]. Other approaches have utilized clustering strategies and colour space transformations to identify visually salient colours across images with wide hue variations [11]. However, these methods generally extract dominant colours in a static manner and do not explicitly model spatial relationships or colour transition dynamics between neighboring regions.

In contrast, fruit quality changes are inherently gradual and spatial in nature. Biochemical reactions during ripening and decay propagate progressively through adjacent tissues, resulting in localized colour changes that expand over time [12]. Therefore, analyzing colour transitions between neighboring pixels becomes essential for capturing spatial propagation patterns that are not represented by static colour features.

Research on matrix-based representations in colour image analysis has demonstrated their effectiveness in modeling spatial relationships among pixels. Graph-based image representations utilizing Laplacian matrices and eigenvalue analysis have been applied to colour image segmentation tasks [13], while comprehensive reviews have highlighted the importance of colour and texture features in image quality assessment [14]. Nevertheless, studies that explicitly integrate colour transition matrices with eigenvalue analysis as quantitative indicators of fruit quality remain limited.

Mathematically, surface colour changes in fruit can be modeled as a spatial transition process among dominant colours. These transitions can be represented using a colour transition matrix formulated as a stochastic matrix, where each element denotes the probability of transitioning from one dominant colour to another. According to Markov chain theory, a stochastic transition matrix possesses a dominant eigenvalue equal to one, while the second-largest eigenvalue (λ_2) reflects the convergence behavior and stability of the colour distribution [15]. A λ_2 value approaching zero indicates a stable and uniform colour distribution, whereas larger λ_2 values signify spatial irregularity, which is commonly associated with early decay processes.

Based on the literature review, several research gaps can be identified: (1) most existing studies rely on static colour features and do not explicitly model spatial colour transitions; (2) deep learning-based approaches lack interpretability with respect to colour change dynamics; and (3) few studies have employed Dominant Colour Transition Matrices combined with eigenvalue analysis as quantitative indicators of fruit quality.

This study focuses on the development of an image-based fruit quality assessment method through the analysis of the Dominant Colour Transition Matrix and eigenvalue features, particularly the second-largest eigenvalue (λ_2), which theoretically represents the stability level of colour distribution. A λ_2 value approaching zero indicates a stable and homogeneous colour distribution (for example, evenly ripened fruit conditions), whereas a larger λ_2 value indicates spatial non-uniformity or significant colour changes that characterize the early stages of decay.

Accordingly, this study does not merely classify fruits into specific categories (such as fresh versus decayed), but also analyzes the gradual stages of fruit quality changes. This characteristic makes the proposed approach more suitable for dynamic fruit quality monitoring applications within fruit supply chains. The proposed method is expected to produce quantitative indicators that are more interpretable, easy to implement, and do not require high computational resources as demanded by deep learning-based methods.

The main contributions of this study include: (i) the proposal of a CDTM-based mathematical model to represent spatial colour transitions on fruit surfaces, (ii) the application of the second-largest eigenvalue (λ_2) as a quantitative indicator of colour stability and potential fruit deterioration, and (iii) the validation of the proposed model using a Cavendish banana dataset to demonstrate the relationship between λ_2 values and fruit freshness levels across different ripening stages.

The results of this study are expected to provide both academic and practical benefits. From an academic perspective, this research enriches the field of digital image processing and colour analysis by introducing a mathematical model that is more interpretable compared to deep learning approaches. From a practical perspective, the proposed method can serve as a foundation for developing efficient, low-cost, and easily implementable non-destructive fruit quality inspection systems for modern agricultural supply chains.

II. METHODOLOGY

This study employs a quantitative experimental research design to develop and evaluate a novel feature extraction method, namely the Colour Dominant–Transition Matrix (CDTM) combined with eigenvalue analysis, as an objective indicator of fruit quality.

The proposed methodology consists of three main stages: (1) image acquisition and preprocessing, (2) feature extraction using the Colour Dominant–Transition Matrix (CDTM) and eigenvalue analysis, and (3) fruit quality classification and evaluation. The overall workflow of the proposed method is illustrated in Fig. 1.

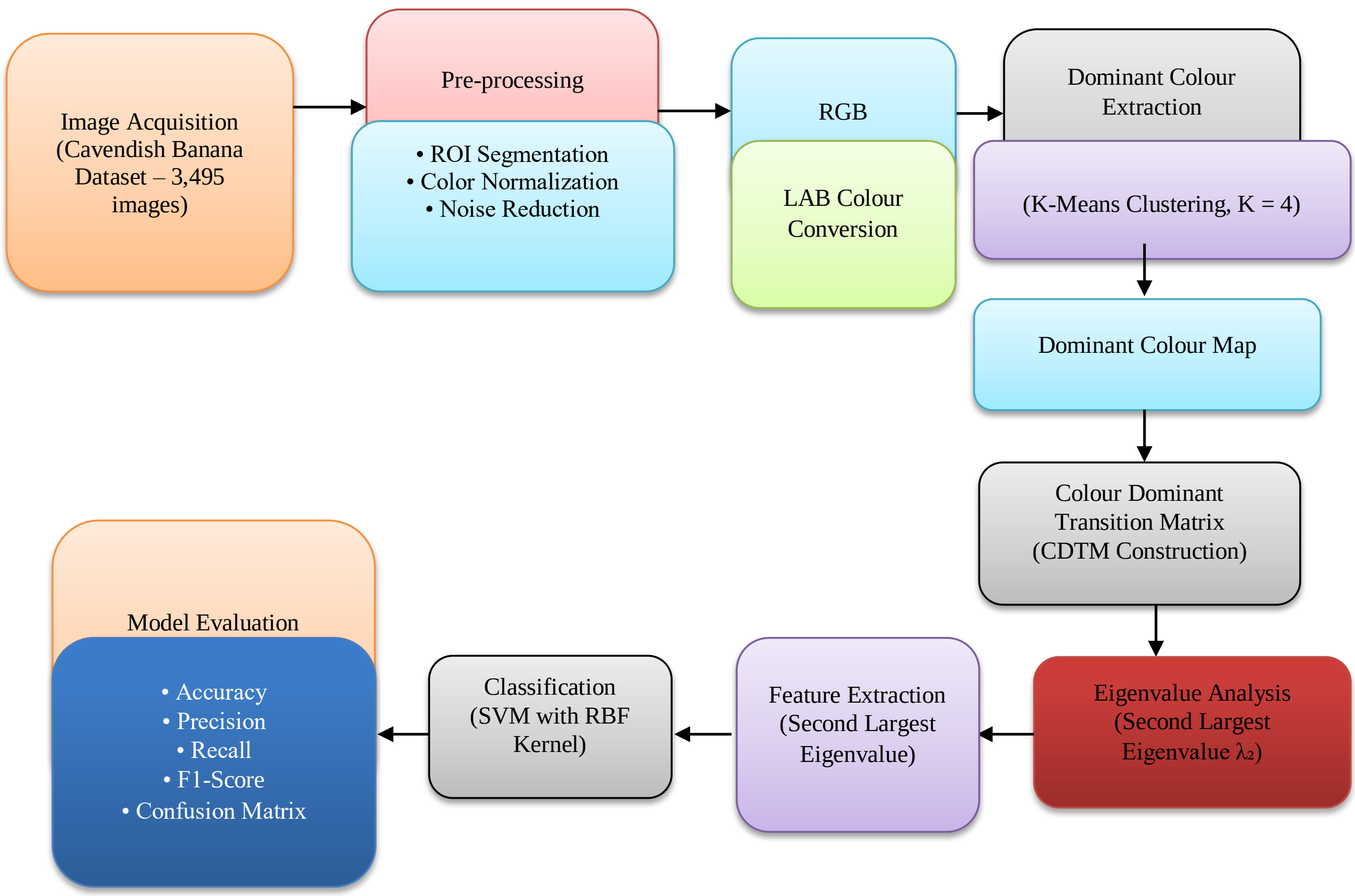

Figure 1. Flowchart of Fruit Quality Assessment Method Based on CDTM and Eigenvalue

A. Image Acquisition

The dataset used in this study consists of 3,495 real images of Cavendish bananas obtained from the Banana Ripeness Images Dataset available on the Kaggle platform. The images represent four ripeness categories: unripe, optimally ripe, overripe, and rotten. The images were captured under relatively consistent imaging conditions, with the fruit placed in front of a plain background to minimize visual noise and facilitate the segmentation process. Such controlled acquisition conditions help maintain consistency in colour representation while still preserving natural variations in fruit appearance.

Although the dataset also provides 161,280 synthetically generated images, this study focuses exclusively on real images to ensure that the analysed colour characteristics accurately reflect real fruit surface conditions. The use of real images helps preserve natural variations in colour distribution, lighting intensity, and surface texture, which are important for evaluating the robustness of the proposed CDTM-based feature extraction method. This dataset has previously been utilized for banana ripeness classification using CNN-based models [16], allowing objective comparison and reproducibility of the proposed method.

In this study, only real images were used, as the analysis aims to evaluate the intrinsic colour characteristics of fruit surfaces without the influence of artifacts introduced by synthetic image generation. All real images were used according to their original labels provided by the dataset source to maintain ground truth consistency.

The images in the dataset were captured under relatively controlled acquisition conditions. The bananas were positioned in front of a plain background to reduce background interference. The camera was placed at a fixed distance from the object to maintain consistent image scale across samples. Lighting conditions were kept relatively stable to minimize excessive illumination variations, allowing the captured images to represent natural colour transitions during the ripening process.

1) Pre-processing

Image preprocessing was conducted to ensure consistency and reproducibility prior to feature extraction, thereby minimizing irrelevant variability. First, object segmentation was performed to separate the fruit region from the background using colour-based masking in the HSV colour space, which has been widely applied for colour-based image segmentation tasks [17]. The HSV thresholds were empirically determined to isolate banana surface pixels while suppressing low-saturation background regions. Morphological opening and closing operations were subsequently applied to remove small artifacts and refine the Region of Interest (ROI).

To reduce illumination variability, colour normalization was applied to minimize lighting intensity differences and environmental illumination effects, ensuring consistent colour representation across image samples [18].

Finally, noise reduction was performed using a median filter with a 3×3 kernel to eliminate salt-and-pepper noise without degrading the spatial structure of the image. This ensures that pixel-level colour transitions accurately reflect intrinsic fruit surface characteristics rather than background interference or acquisition noise.

Specifically, luminance normalization was performed on the L channel of the CIELAB colour space using min–max scaling to standardize brightness values while preserving chromatic information in the a and b channels. The normalization process follows the min–max scaling formulation in Eq. 1:

$$L' = \frac{L - L_{\min}}{L_{\max} - L_{\min}} \tag{1}$$

Where L represents the original luminance value, and $L_{\max}$ *and* $L_{\min}$ denote the minimum and maximum luminance values within the image.

B. Feature Extraction Based on Colour Dominant Transition Matrix (CDTM)

This stage aims to model the spatial transition probabilities among dominant colours on the fruit surface.

1) Dominant Colour Determination

Colour transformation was first applied to convert the ROI images from the RGB colour space to the LAB colour space, allowing perceptual colour differences to be represented more objectively[19]. Subsequently, K-means clustering was applied to the *a* and *b* channels to group pixels into a limited number of dominant colour clusters *($K = 4$)*. The result is a dominant colour map $C(x, y)$, in which each pixel is assigned a colour cluster label $k \in \{1, 2, \ldots, K\}$.

K-Means clustering was employed to group pixel colours into dominant clusters. The number of clusters was set to **K = 4**, corresponding to the main colour groups commonly observed during banana ripening stages: green (unripe), yellow (ripe), brown (early decay), and dark brown to black (severely decayed). This configuration enables the clustering process to capture the primary chromatic variations associated with ripeness progression while maintaining a compact representation for the Colour Dominant Transition Matrix.

2) Construction of the Colour Dominant–Transition Matrix (CDTM)

Colour transitions were computed based on spatial relationships between neighboring pixels using an 8-neighborhood connectivity scheme [20]. Transitions were calculated in the horizontal (0°) and vertical (90°) orientations.

The Colour Dominant–Transition Matrix M is defined as a $K \times K$, where each element $M_{i,j}$ epresents the probability that a pixel with dominant colour i is followed by a neighboring pixel with dominant colour j. The matrix elements are computed using Eq. (2):

$$M_{i,j} = \frac{N(i,j)}{\sum_{t=1}^{K} N(i,t)} \qquad (2)$$

Where $N(i, j)$ denotes the number of pixel pairs transitioning from dominant colour i to dominant colour j. The matrix M is stochastic, meaning that the sum of each row equals one, representing the total transition probability from each dominant colour [20]. The definitions of the main symbols used in this study are summarized in Table I.

TABLE I
DEFINITION OF MAIN SYMBOLS

Simbol	Units	Deskripsi Kuantitas
K	—	Number of dominant colour clusters (e.g., $K = 4$)
M	—	Colour Dominant–Transition Matrix (CDTM), size $K \times K$
$M_{i,j}$	—	Transition probability from dominant colour i to j
$N(i,j)$	—	Number of pixel pairs transitioning from colour i to j
t		Summation index in Σ
λ	—	Eigenvalue of matrix M
λ_1	—	Largest eigenvalue (always equal to 1)
λ_2	—	Second-largest eigenvalue (quality feature)

C. Eigenvalue Analysis of the CDTM

The colour transition matrix M was further analyzed to obtain its eigenvalues λ(lambda). he eigenvalues were computed by solving the characteristic equation shown in Eq. (3):

$$\det(M - \lambda I) = 0 \qquad (3)$$

Where $\boldsymbol{I}$ denote the identity matrix of the same order as $\boldsymbol{M}$. Based on the properties of stochastic matrices, the largest eigenvalue (stochastic matrix), nilai eigen terbesar λ1 =1. For fruit quality analysis, this study focuses on the second-largest eigenvalue λ2, defined as:

$$\lambda 2 = max\,\{\lambda i : \lambda i \neq 1\} \qquad (4)$$

The value of λ2 represents the convergence rate of the colour distribution or the level of spatial stability on the fruit surface. A λ2 value close to zero indicates a stable and homogeneous colour distribution, typically associated with evenly ripened fruit, whereas larger λ2 values indicate spatial colour irregularities or localized changes, which are characteristic of fruit undergoing decay.

1) *Quality Feature Extraction*

The primary quality feature proposed in this study is the second-largest eigenvalue (λ_2). This value λ_2 uantifies the convergence rate of the transition matrix toward its steady-state vector and reflects the stability of the surface colour distribution. A lower λ_2 value indicates a stable and uniform colour distribution, while higher values indicate increased spatial non-uniformity.

D. Fruit Quality Classification and Evaluation

The extracted second-largest eigenvalue ($\lambda_{2)}$ from each fruit sample was used as a single input feature for classifying fruit quality into three categories: Fresh, Optimally Ripe, and Decayed/Damaged. Classification was performed using a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel, which is effective for separating low-dimensional, non-linear data distributions.

The dataset was divided into 80% for training and 20% for testing to evaluate the generalization capability of the model. In addition, 5-fold cross-validation was applied to the training data to ensure model stability and prevent overfitting. Key SVM parameters, including the cost parameter (C) and gamma, were empirically tuned to achieve optimal performance.

Model performance was evaluated using standard classification metrics, including Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. Accuracy measures the overall proportion of correct predictions, Precision and Recall assess the balance between false positives and false negatives, while the F1-Score provides a harmonic measure of Precision and Recall. The Confusion Matrix was used to analyze classification errors across each fruit quality category.

E. Hardware and Software Environment

The proposed method was implemented using a Python-based Google Colab environment. Computations were performed on a CPU runtime with an Intel Xeon processor at 2.20 GHz and 13 GB of memory. Image processing tasks were conducted using the OpenCV library, while classification and model evaluation were performed using the scikit-learn library.

III. RESULT AND DISCUSSION

A. Dataset Description

The dataset used in the experimental stage consists of 3,495 real images of Cavendish bananas, categorized into four ripeness levels. Class A represents unripe fruit dominated by green colouration, Class B corresponds to optimally ripe fruit with uniformly distributed yellow colour, Class C represents fruit in the early decay stage characterized by the appearance of brown spots on the peel surface, and Class D represents severely decayed or damaged fruit with dark brown to black colouration. Although the original dataset also provides 161,280 synthetic images, this study exclusively utilizes real images to ensure that the analyzed colour characteristics accurately reflect authentic visual conditions of the fruit surface. The relatively large number of real images used in this study allows the proposed method to be evaluated across diverse visual variations of banana ripeness, thereby improving the reliability of the experimental results.

The dataset was proportionally divided into training, validation, and testing subsets, as summarized in Table II, to maintain class balance and prevent data leakage during model training and evaluation.

TABLE II
CAVENDISH BANANA DATASET DISTRIBUTION

Class	Train	Validation	Test
A	857	286	286
B	489	163	163
C	335	112	112
D	415	139	138

B. Pre-Processing Results

The preprocessing stage aims to minimize irrelevant variability prior to feature extraction. Fig. 2(a) shows the original banana image, while Fig. 2(b) illustrates the segmentation result in the HSV colour space, which effectively separates the fruit object from the background.

Subsequently, colour normalization and noise reduction using a median filter were applied, as shown in Fig. 2(c).

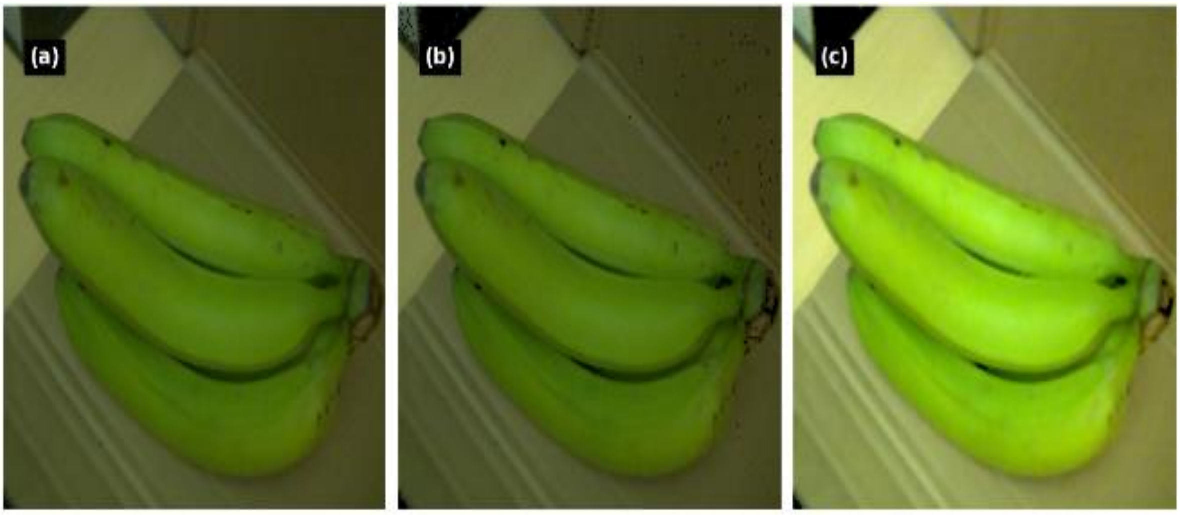

Figure 2. Banana Image Pre-Processing Results

The segmentation result shown in Fig. 2(b) demonstrates that the background was effectively removed, leaving only the fruit region as the Region of Interest (ROI). Subsequently, the normalization process illustrated in Fig. 2(c) improves colour intensity uniformity across pixels, reduces uneven illumination effects, and preserves surface texture details. Therefore, the preprocessed images are suitable for dominant colour feature extraction.

C. CDTM Feature Extraction Results and Eigenvalue Calculation

After the preprocessing stage, the images were converted into the CIELAB colour space and quantized using K-Means clustering with the number of clusters set to *K=4K = 4K=4*. Figure 3 illustrates the resulting dominant colour map, where each colour represents a cluster of pixels with similar chromatic characteristics. The selection of *K=4K = 4K=4* is based on the observation that banana surfaces are typically dominated by four main colour groups corresponding to different ripeness stages, namely green (unripe), yellow (optimal ripe), brown (early deterioration), and dark brown to black (severely rotten).

The dominant colour representation simplifies the original RGB image while preserving essential chromatic information related to fruit ripeness. Each pixel is assigned a dominant colour label, enabling a structured spatial analysis of colour distribution. This representation facilitates the construction of the Colour Dominant-Transition Matrix (CDTM), which captures spatial transitions between dominant colours across neighboring pixels.

The selection of $K = 4$ allows the clustering process to represent the major colour groups observed in banana ripening stages, thereby simplifying the colour representation while preserving meaningful chromatic information relevant to fruit maturity. An example of the dominant colour representation generated using K-Means clustering is illustrated in Figure 3.

Figure 3. Dominant Colour Map Using K-Means

The quantized images display segmented colour regions that represent gradations of fruit ripeness. Green colour tends to dominate unripe bananas, while yellow to brown colours appear in more mature fruits. This colour representation provides more structured visual information compared to the original RGB images, as each pixel is now assigned a limited dominant colour label. Consequently, this representation facilitates the analysis of texture and spatial colour patterns.

Subsequently, the relationships among dominant colours are analyzed using the Colour Dominant Transition Matrix (CDTM), as shown in Fig. 4. Each matrix element *Mi,j* represents the probability of a transition from dominant colour *i* menuju to dominant colour *j* based on spatial relationships between neighboring pixels in the horizontal and vertical directions. Diagonal elements indicate the stability of a particular colour (self-transitions), whereas non-diagonal elements represent the frequency of transitions between different dominant colours.

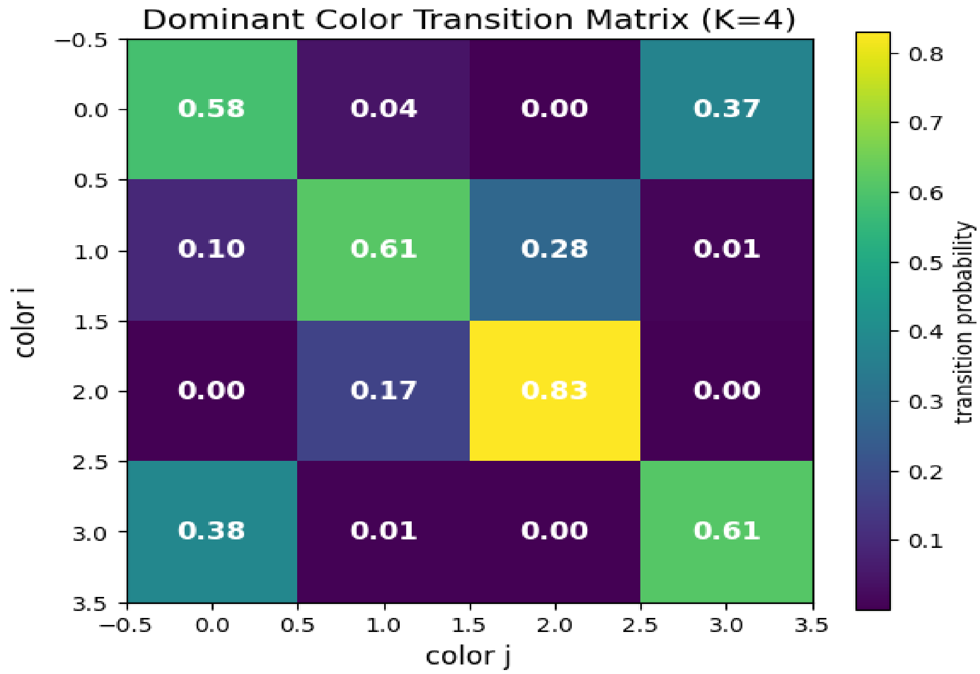

Figure 4. Colour Dominant Transition Matrix (CDTM

Based on the results shown in Fig. 4, the highest probability values are observed along the diagonal elements, particularly at (M2,2=0.83), indicating that the corresponding dominant colour exhibits a stable and homogeneous distribution over the fruit surface. In contrast, transitions between different colours, such as from colour 0 to colour 3 (M0,3=0.37) indicate relatively significant colour variation, which commonly occurs in transition areas between green and yellow regions of the banana peel.

These transition patterns form the basis for calculating the second largest eigenvalue (λ_2) in the subsequent stage, which is used as a quantitative indicator of colour variation and visual quality of the fruit surface. A smaller λ_2, value indicates a more stable and uniform colour distribution, which generally corresponds to optimal ripeness and fresh fruit condition.

After constructing the CDTM for each ripeness class, the second largest eigenvalue (λ_2) is computed as an indicator of the stability of colour distribution on the fruit surface. The value of λ_2 represents the convergence rate of the colour distribution toward a steady state. A smaller λ_2 value indicates a more uniform and stable colour distribution, whereas a larger λ_2 value reflects increased colour irregularity, which is typically associated with unripe or unevenly matured fruit.

Table III presents the computed λ_2 values for the four ripeness classes of Cavendish bananas.

TABLE III
SECOND LARGEST EIGENVALUE (Λ_2) FOR EACH RIPENESS CLASS

Class	λ_2
A	0.7972
B	0.1045
C	0.4680
D	0.2423

Based on Table III, Class A exhibits the highest λ_2 value (0,7972), indicating the most unstable and heterogeneous colour distribution among all classes.

To further illustrate the relationship between the second largest eigenvalue (λ_2) and banana ripeness levels, a graphical visualization is presented in Fig. 5.

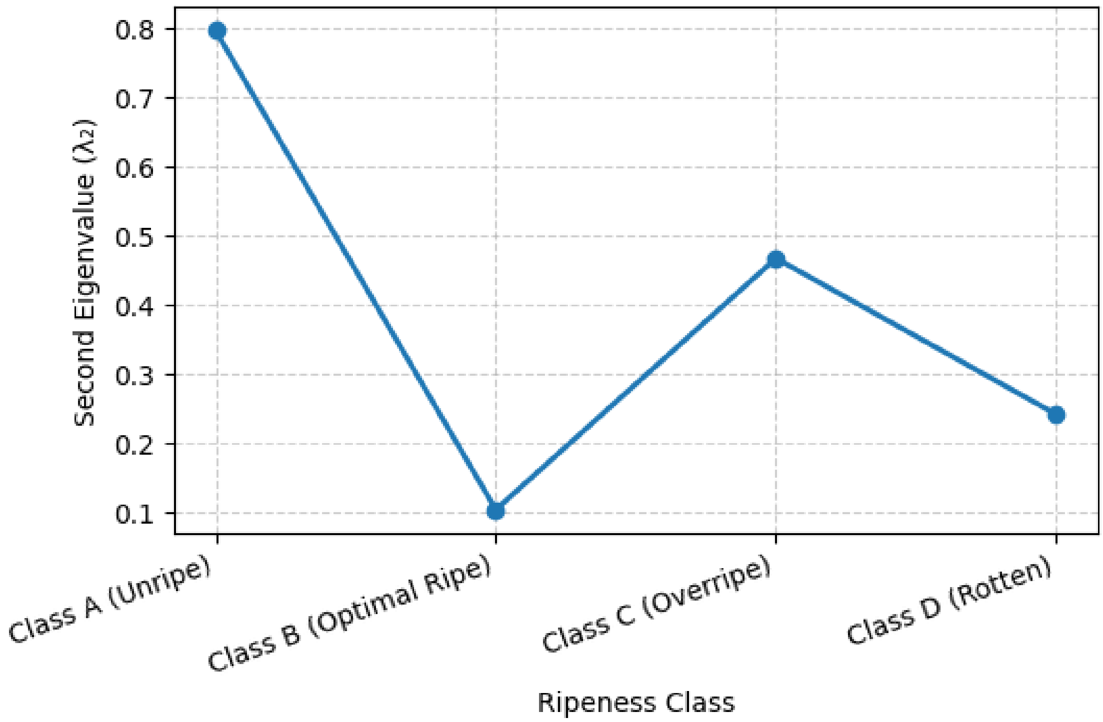

Figure 5. Relationship Between the Second Largest Eigenvalue *(λ_2)* and Banana Ripeness Levels

Fig. 5 illustrates the relationship between the second largest eigenvalue (λ_2) nd the ripeness level of bananas. It can be observed that the highest λ_2 value occurs in *Class A (Unripe)*, indicating the highest degree of colour instability due to the dominance of green colour with large spectral variation.

In contrast, *Class B (Optimal Ripe)* exhibits the lowest λ_2 value (0.1045), indicating the most stable and homogeneous colour distribution. This suggests that at the optimal ripeness stage, the fruit surface displays a more uniform yellow colour with minimal variation.

Meanwhile, *Class C (Overripe)* and *Class D (Rotten)* show an increase in λ_2 values, reaching 0.4680 and 0.2423, respectively. This phenomenon is caused by the appearance of brown spots and dark regions on the banana peel, which increase colour variability across the fruit surface.

Therefore, the λ_2 value can be used as a quantitative indicator of fruit ripeness and visual quality, where lower values indicate optimally ripe and fresh fruit, while higher values reflect colour irregularities associated with unripe or rotten conditions.

D. SVM-Based Classification Results

After obtaining the second largest eigenvalue (λ_2) for all samples, the classification process was conducted using a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel. The model was evaluated using 20% of the dataset as test data. Based on the evaluation results, the model achieved an overall accuracy of 64%.

More detailed evaluation metrics, including precision, recall, and F1-score for each class, are presented in Table IV.

TABLE IV
SVM EVALUATION RESULT (PRECISION, RECALL, F1-SCORE)

Class	**Precision**	**Recall**	**F1-Score**	**Support**
A	0.67	0.91	0.77	286
B	0.54	0.23	0.32	163
C	0.55	0.50	0.53	112
D	0.64	0.67	0.65	138
Accuracy			**0.64**	699

The model achieved an accuracy of 64%, indicating that the λ_2 feature derived from the CDTM is capable of classifying banana ripeness levels with moderate accuracy.

The model achieved an accuracy of 64%, indicating that the λ_2 feature derived from the CDTM is capable of classifying banana ripeness levels with moderate accuracy. This moderate accuracy suggests that while the λ_2 feature captures important colour transition characteristics, additional visual descriptors may be required to fully represent the complexity of banana ripening patterns.

To further analyze misclassification patterns, the confusion matrix of the SVM classification results is shown in Fig. 6.

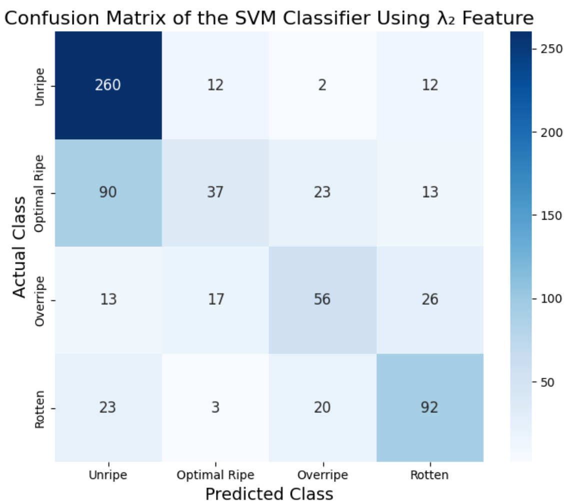

Figure 6. Confusion matrix of the SVM classification results

The confusion matrix in Fig. 6 shows that Class A achieves the highest prediction performance with a recall of 0.91, while Classes B and C exhibit lower performance due to more complex colour variations and overlap between adjacent ripeness levels. Overall, the model demonstrates good discrimination for extreme conditions (unripe and rotten), but weaker performance for transitional conditions (optimal ripe and overripe). Model performance could be improved by incorporating texture features or additional nonlinear features.

Figure 6 presents the confusion matrix of the SVM model trained using the λ_2 feature. he prediction patterns in the matrix illustrate the model's ability to distinguish between ripeness classes based on the stability of colour distribution.

Class A (Unripe) shows a high number of correct predictions, with 260 out of 286 samples correctly classified. This result indicates that unripe bananas exhibit distinctive colour patterns characterized by unstable and heterogeneous green colour distributions, resulting in relatively high λ_2 values that are easily separable from other classes.

For Class B (Optimal Ripe), the model demonstrates relatively unstable performance. Out of 163 samples, only 37 are correctly classified. Most misclassifications occur because although many Class B samples exhibit homogeneous yellow colouration, some show minor variations that cause λ_2 values to overlap with those of Class C or D. As a result, the model often misinterprets optimally ripe bananas as early rotten or transitioning fruits.

In Class C (Overripe), the model correctly identifies 56 out of 112 samples. This behavior can be explained by the diverse colour transition patterns caused by brown spots, which are not as visually dominant as the dark regions observed in Class D, leading to confusion with Class B.

For Class D (Rotten), the model performs relatively well, correctly classifying 92 out of 138 samples. The dominance of dark colours on the fruit surface produces characteristic λ_2 patterns that can be recognized by the classifier.

Overall, the confusion matrix indicates that the λ_2 feature effectively separates Classes A and D, but is less effective in distinguishing Classes B and C due to overlapping λ_2 values. This explains the overall accuracy of 64% and the relatively low precision and recall values for Class B (Optimal Ripe).

Therefore, the inclusion of additional colour or texture features is required to improve class separability, particularly for ripeness levels with similar visual characteristics.

E. Evaluation and Interpretation of Results

The second largest eigenvalue (λ_2) exhibits a consistent pattern corresponding to the visual characteristics of each banana ripeness level. Class B (Optimal Ripe) shows the lowest λ_2 values , indicating a stable and homogeneous yellow colour distribution, whereas Class A (Unripe) exhibits higher λ2\lambda_2λ2 values due to heterogeneous green colour dominance. This pattern aligns with previous findings that identify colour variation as a primary indicator in the fruit ripening process [21]

In the classification stage, λ_2 is used as a single input feature for the Support Vector Machine (SVM) model. The evaluation results show an overall accuracy of 64%. The highest recall is achieved in Class A (0.91), indicating that the model is highly effective in identifying unripe bananas. However, the lowest recall is observed in Class B (0.23), suggesting that the model struggles to correctly identify bananas in the optimal ripeness stage. This imbalance indicates that the discriminative capability of λ_2 varies across different ripeness levels.

The confusion matrix further reveals that λ_2 effectively separates extreme classes, particularly unripe (Class A) and rotten bananas (Class D). However, its performance decreases in transitional ripeness stages, namely optimal ripe (Class B) and overripe (Class C). This limitation is mainly caused by similarities in colour transition patterns between these classes, resulting in overlapping λ_2 values.

From a visual perspective, Class C samples typically contain brown spots that increase colour variability across the peel surface. Meanwhile, some samples in Class B also exhibit slight colour variations due to lighting conditions, peel texture, or early ripening spots. These similarities cause the model to frequently misclassify Class B samples as Class C or Class D. Therefore, relying solely on a single eigenvalue feature is insufficient to capture the complex visual patterns that occur during intermediate ripening stages.

These findings are consistent with previous studies reporting that single colour-based descriptors are generally limited in distinguishing closely related ripeness levels unless they are combined with additional features such as texture, spatial distribution, or multi-channel colour information [22]

Compared to deep learning approaches such as the CNN-based method proposed by Chuquimarca *et al.* [16], the CDTM-based approach is computationally simpler and more interpretable. CNN models are capable of automatically extracting complex hierarchical features from images, which often leads to higher classification accuracy. In contrast, the λ_2-based CDTM approach relies on handcrafted features derived from colour transition patterns, which may limit its ability to represent more complex visual variations.

Nevertheless, the proposed method offers several advantages, including lower computational cost, easier implementation, and better interpretability of the extracted features. These characteristics make the approach suitable for lightweight applications or systems with limited computational resources.

Overall, the results demonstrate that CDTM and the λ_2 feature provide reliable indicators for identifying extreme fruit conditions, particularly highly unripe and severely rotten bananas. However, additional features such as texture descriptors, multiple eigenvalue components, or colour channel combinations should be incorporated to improve classification performance in transitional ripeness stages. Future work should also include validation using larger datasets, varying lighting conditions, and different fruit varieties to enhance the robustness and generalizability of the proposed approach.

Statistical validation between ripeness classes can be explored in future work using hypothesis testing methods such as ANOVA or Kruskal–Wallis tests to further validate the discriminative capability of the extracted features.

IV. Conclusion

Based on the analysis conducted, it can be concluded that the preprocessing stages, dominant colour extraction, and the construction of the Colour Dominant Transition Matrix (CDTM) are able to effectively represent the visual characteristics of Cavendish banana surfaces at different ripeness levels. The segmentation, colour normalization, and noise reduction processes improve image quality, allowing the fruit region to be accurately identified. The colour quantization results obtained using K-Means clustering with four dominant clusters demonstrate consistent differences in colour distribution across ripeness classes.

The second largest eigenvalue (λ_2) derived from the CDTM exhibits significant variation among the ripeness classes. Class B (optimal ripeness) shows the lowest λ_2 values, indicating the most homogeneous colour distribution, while Class A (unripe) presents the highest λ_2 value due to the heterogeneous dominance of green colour. These findings indicate that λ_2 can serve as a quantitative indicator for assessing colour homogeneity and fruit ripeness levels.

can serve as a quantitative indicator for assessing colour homogeneity and fruit ripeness levels λ_2 feature was used as a single input for the SVM model and achieved an overall accuracy of 64%. The model was able to distinguish extreme classes, such as unripe (Class A) and overripe/rotten (Class D), relatively well; however, its performance decreased for transitional classes (Classes B and C) that exhibit more similar colour patterns. This result suggests that although λ_2 is effective as a visual indicator, a single feature is not sufficient to robustly separate complex ripeness classes.

Overall, the CDTM method combined with λ_2 analysis is proven to be simple, efficient, and interpretable, and shows potential for application in non-destructive fruit quality inspection systems. Further development is required by incorporating additional texture or colour features to improve classification accuracy, particularly for transitional ripeness conditions.

The proposed CDTM-based approach has potential practical applications in automated fruit grading systems used in agricultural supply chains. Due to its relatively low computational complexity and interpretable features, the method can be integrated into lightweight image-based monitoring systems for post-harvest quality assessment, enabling rapid and non-destructive evaluation of fruit ripeness. Such systems may support farmers, distributors, and retailers in improving sorting efficiency and maintaining product quality.

REFERENCES

[1] J. Blasco, N. Aleixos, and E. Moltó, "Machine Vision System for Automatic Quality Grading of Fruit," *Biosyst Eng*, vol. 85, no. 4, pp. 415–423, Aug. 2003, doi: 10.1016/S1537-5110(03)00088-6.

[2] D. Lorente, N. Aleixos, J. Gómez-Sanchis, S. Cubero, O. L. García-Navarrete, and J. Blasco, "Recent Advances and Applications of Hyperspectral Imaging for Fruit and Vegetable Quality Assessment," May 2012. doi: 10.1007/s11947-011-0725-1.

[3] S. Hira and S. Lande, "Detection of fruit ripeness using image processing," *Int J Health Sci (Qassim)*, pp. 3874–3886, Jul. 2022, doi: 10.53730/ijhs.v6nS6.10146.

[4] L. E. Chuquimarca, B. X. Vintimilla, and S. A. Velastin, "A review of external quality inspection for fruit grading using CNN models," *Artificial Intelligence in Agriculture*, vol. 14, pp. 1–20, Dec. 2024, doi: 10.1016/J.AIIA.2024.10.002.

[5] D. D. Andayani, A. N. A. Rafii, K. Mahdar, W. A. Ahmad, and M. Mustamin, "Papaya Fruit Quality Classification Based on Lab Colour and Texture Features Using Artificial Neural Networks (ANN)," *INTEK: Jurnal Penelitian*, vol. 10, no. 1, pp. 15–21, Jun. 2023, doi: 10.31963/intek.v9i2.4246.

[6] W. Wijaya Widiyanto and E. Purwanto, "Classification Of Mango Fruit Quality Based On Texture Characteristics Of Glcm (Gray Level Co-Occurrence Matrices) With Algorithm K-NN (K-Nearest Neighbors)," vol. 20, no. 1, pp. 31–40, 2019, [Online]. Available: http://jurnalnasional.ump.ac.id/index.php/Techno

[7] T. Y. Prahudaya and A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji Menggunakan Knn Berdasarkan Fitur Warna Dan Tekstur," *Jurnal Teknosains*, vol. 6, no. 2, p. 113, Aug. 2017, doi: 10.22146/teknosains.26972.

[8] S. Choirunisa and S. -, "Klasifikasi Tingkat Kematangan Citra Buah Tomat Menggunakan Densenet-121," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.6652.

[9] R. K. Yadav and R. K. Goyal, Post Harvest Technology of Horticultural Crops. Jodhpur, India: Agrobios (India), 2012.

[10] Y. Chang and N. Mukai, "Colour Feature Based Dominant Colour Extraction," *IEEE Access*, vol. 10, pp. 93055–93061, 2022, doi: 10.1109/ACCESS.2022.3202632.

[11] Y. Kynash and M. Semeniv, "New Approach to Dominant and Prominent Colour Extraction in Images with a Wide Range of Hues," *Technologies (Basel)*, vol. 13, no. 6, Jun. 2025, doi: 10.3390/technologies13060230.

[12] B. M. Nicolaï *et al.*, "Spatio-temporal dynamics of the metabolome of climacteric fruit during ripening and post-harvest storage," Oct. 31, 2023, *Oxford University Press*. doi: 10.1093/jxb/erad230.

[13] C. Theoharatos, "Colour image segmentation using Laplacian eigenmaps," *J Electron Imaging*, vol. 18, Apr. 2009, doi: 10.1117/1.3122369.

[14] M. Gao, C. Song, Q. Zhang, X. Zhang, Y. Li, and F. Yuan, "Research Progress on Colour Image Quality Assessment," Sep. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/jimaging11090307.

[15] S. Lafon and A. B. Lee, "Diffusion Maps and Coarse-Graining: A Unified Framework for Dimensionality Reduction, Graph Partitioning and Data Set Parameterization."

[16] L. E. Chuquimarca, B. X. Vintimilla, and S. A. Velastin, "Banana Ripeness Level Classification Using a Simple CNN ModeTrained with Real and Synthetic Datasets," in *Proceedings of the International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Science and Technology Publications, Lda, 2023, pp. 536–543. doi: 10.5220/0011654600003417.

[17] F. Garcia-Lamont, J. Cervantes, A. López, and L. Rodriguez, "Segmentation of images by colour features: A survey," *Neurocomputing*, vol. 292, pp. 1–27, May 2018, doi: 10.1016/J.NEUCOM.2018.01.091.

[18] P. Dhiman, A. Kaur, V. R. Balasaraswathi, Y. Gulzar, A. A. Alwan, and Y. Hamid, "Image Acquisition, Preprocessing and Classification of Citrus Fruit Diseases: A Systematic Literature Review," *Sustainability (Switzerland)*, vol. 15, no. 12, Jun. 2023, doi: 10.3390/su15129643.

[19] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 Colour-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations," *Col Res Appl*, vol. 30, pp. 21–30, 2005, doi: 10.1002/col.

[20] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," IEEE Trans. Syst. Man Cybern., 1973.

[21] D. A. Levin, Y. Peres, and E. L. Wilmer, "Markov Chains and Mixing Times, second edition."

[22] M. Rizzo, M. Marcuzzo, A. Zangari, A. Gasparetto, and A. Albarelli, "Fruit ripeness classification: A survey," *Artificial Intelligence in Agriculture*, vol. 7, pp. 44–57, Mar. 2023, doi: 10.1016/J.AIIA.2023.02.004.

[23] C. S. Nandi, B. Tudu, and C. Koley, "A Machine Vision-Based Maturity Prediction System for Sorting of Harvested Mangoes," *IEEE Trans Instrum Meas*, vol. 63, no. 7, pp. 1722–1730, 2014, doi: 10.1109/TIM.2014.2299527.