Banana Cultivar Classification using Scale Invariant Shape Analysis

Kwankamon Dittakan, Nawanol Theera-Ampornpunt, Waraphon Witthayarat,
Sararat Hinnoy, Supawit Klaiwan, and Thunyatorn Pratheep
College of Computing
Prince of Songkla University, Phuket Campus
Phuket, Thailand
{kwankamon.d,nawanol.t}@phuket.psu.ac.th, {S5630211027,S5630211054,S5630211030,S5730211068}@email.psu.ac.th

***Abstract*—This paper presents scale-invariant shape analysis with respect to banana cultivar detection. We consider three cultivars: Cavendish, Lady Finger, and Pisang Awak. We present the appropriate image preprocessing methods and compare different feature selection algorithms, numbers of features, as well as various machine learning models as the classifier.**

We found that the best feature selection method is chi square, and the optimal number of features is 100. Differences between prediction accuracy of machine learning models are small, but overall, Bayesian network performs the best, with overall AUC of 0.933 and overall accuracy of 84%.

***Index Terms*—object recognition, banana cultivar, shape analysis**

I. Introduction

Banana is an herbaceous plant in the genus Masa in the family Musacesa which can be grown in many countries across the world. Banana fruit is produced all year, maling it an important economic crops that can be use for domestic consumption as well as exports. Banana is appropriate for everyone because the it is rich in nutrients, vitamins, and minerals such as carbohydrates, vitamin A, vitamin B6, vitamin B12, etc. Many types of banana are grown commercially in Thailand, such as Cavendish banana, Lady Finger banana, and Pisang Awak banana. At present, classification of banana cultivar is being done by human. The use of machine for the classification can increase the effectiveness, reduce the cost, and reduce mistakes. The techniques and resulting models can be utilized by factories that process many banana cultivars, and make up an important piece that enables automatic checkout at supermarkets.

In this work, we consider digital images as raw input. Because it is inefficient to ensure that all images are taken with exactly the same distance and focal length, we rely on the scale invariant feature transform (SIFT) technique which removes the dependency on the correct distance. For this problem, color and texture do not offer much discriminating power, as they are similar across banana cultivars. Therefore, we focus on shape as the sole feature.

We present proper techniques in the preprocessing step, which involves image scaling, image segmentation, edge detection, feature extraction, and feature selection. For feature selection, we give a comparison of three methods, as well as a comparison of the number of features, with respect to prediction accuracy. Finally, we compare seven machine learning models as the classifier, so that the optimal model can be chosen.

The rest of the paper is organized as follows. Section II presents related work. Section III presents the design of our banana cultivar classification framework. Section IV describes our dataset and presents evaluation methods and results. Finally, Section V concludes the paper.

II. Previous Work

Image classification has been applied in many domains. In agriculture, it has been used for evaluating quality of produce and separating different kinds of produce, as it can reduce costs and errors.

Spectroscopy is widely used for detecting and classifying agriculture crops. Pholpho et al. used visible spectroscopy to classify non-bruised and bruised longan fruit[1]. Three models were compared: PCA, partial least square discriminant analysis, and soft independent modeling of class analogy. Yang et al. used spectral analysis to classify between mature fruit, near-mature fruit, near-young fruit, young fruit, and leaf of blueberry [2]. Many other works utilized spectroscopy to analyze quality [3], [4], [5], type [6], [7], detect disease [8] of agriculture crops in both raw and processed forms. However, spectroscopy requires specialized equipments, which are more costly than regular cameras.

Regular digital images are also widely used for crop recognition and quailty classification. The main potential features are color, shape, and texture. Marchal et al. used computer vision techniques to classify olive oil based on the amount of impurity [9]. Thamachot used image processing to classify grades, species, and types of splendid squid [10]. External features of the squid are used as input features. The machine learning models used include neural network and regression analysis. Arivazhagan et al. studied object recognition with 15 types of fruit, namely plum, Agata potato, Asterix potato, cashew, onion, orange, Tahiti lime, kiwi, Fuji apple, Granny Smith apple, watermelon, honeydew melon, nectarine, Williams pear, and Diamond peach [11]. Input features used are color and texture. Image segmentation is done using the

S channel in the HSV color space. The results show that recognition rate is significantly higher when both color and texture are used.

III. Design

The overview diagram of the banana cultivar classification framework is shown in Figure 1. The steps are divided into two phases: training phase and prediction phase. In the training phase, raw image goes through three steps of preprocessing to extract the important features. The three steps are segmentation, representation, and feature selection. The features are then fed as input to learning algorithm to build a classification model. In the prediction phase, the image first goes through the same three preprocessing steps as in the training phase. The classification model then predicts a class using the extracted features as input.

A. Preprocessing

The raw digital images are represented using three values between 0–255, corresponding to red, green, and blue channels, for each pixel, and the number of pixels is in the millions. Using these raw data to train classification models directly is impractical, because of computational complexity, not to mention that a large amount of training data is required to prevent overfitting. In this work, we would like to use the shape (i.e., outline) of the banana as the input to the model. To this end, we use the following preprocessing steps: segmentation, representation, and feature selection.

1) Segmentation: The goal of the segmentation step is to separate the banana from irrelevant parts of the image, such as shadows and background.

The segmentation is separating data that is interested out of the background or out of the data that does not want which also is consideration the luminance of image for gray scale image and consideration the different of color for color image which includes edge of image and feature of texture too. The performance of this method (Segmentation) is based on the intensity selection of starting pixel of each group which can be determined by user. This process has been calculated automatically from histogram of image by using the maximum consideration of histogram (Peak value) became to the intensity value of starting pixel of group. Normally, the intensity value will be had more than one value per group. Thus it is done by merging for integrating the group that gas the value of statistical similarly. The image enhancement must be done before segmentation process because image must be improved to better for analyzing the appearance with edge detection.

Due to the banana images have the different environment such as luminance and shadow of image, so should be selected the best method for obtaining the best result for analyzing which the RGB image is still not appropriate for processing. Thus must be converted into another color system that is appropriate. YCbCr is the one of color system that can use well. The presentation of YCbCr color system will use the luminance signal and color different signal which is divided into 3 types as follows: Y is collected the luminance, Cb is collected the blue color that has been cut the luminance and Cr is collected the red color that has been cut the luminance.

Once the image is in the YCbCr color format, we compare how well each channel lets us separate the object from the background. The segmentation is done using thresholding. We find that the Cb channel works best for this purpose, as the resulting binary image most closely resembles the shape of the banana.

After we have the binary image, holes within the object are filled, and edge detection is performed to extract the outline of the object. There are many edge detection methods such as Canny, Sobel, Prewitt, etc. For this work, we use Canny's edge detection algorithm. An sample image of a banana after applying Canny edge detector is shown in Figure 3.

2) Representation: In the representation step, the outline of the object from the previous step is converted to a common representation. Two common representation formats are 1) representation of the object's outline and 2) representation of the pixels representing the object. We use 1) in this work, since it is more compact, and the pixels representing the object do not carry any meaningful information in addition to the object's outline.

The object's outline is represented in relative to the centroid of the object. The three cultivars of banana are symmetric about the y axis, but not the x axis. Thus, to find the centroid, we first find the midpoint along the x axis, which becomes the x coordinate of the centroid. Then, at that x coordinate, we draw a vertical line within the object's outline. The midpoint along the y axis gives us the y coordinate of the centroid.

Once we have the centroid, the object's outline is represented using polar coodinates. In the polar coordinate system, a point is referred to using the distance from the reference point and the angle from the reference direction. The centroid serve as our reference point, and the angle follows the usual convention, with the direction pointing to the right as 0 degree and continuing counterclockwise. We use 360 features to represent the object's outline, with each coming from the distance from the reference point along a particular angle. Since many points may fall into the same bin of angle, the average distance over all these points is used. For example, the first feature is the average distance from the centroid to all points within degree 0–1. The second feature correponds to degree 1–2, and so on. Figure 4 illustrates this process.

3) Feature selection and data discretization: After the first two steps, each banana object is now represented by 360 features, each describing the distance from the centroid to the outline of the banana at a particular angle. While the number of features is already in the manageable range for most machine learning models, reducing the size of the feature set further will speed up the training process and reduces the risk of overfitting especially when the number of samples is small.

In this work, we compare three standard metrics for feature selection using neural network as the machine learning model: chi square, gain ratio, and information gain. The features are ranked according to their degree of discrimination or

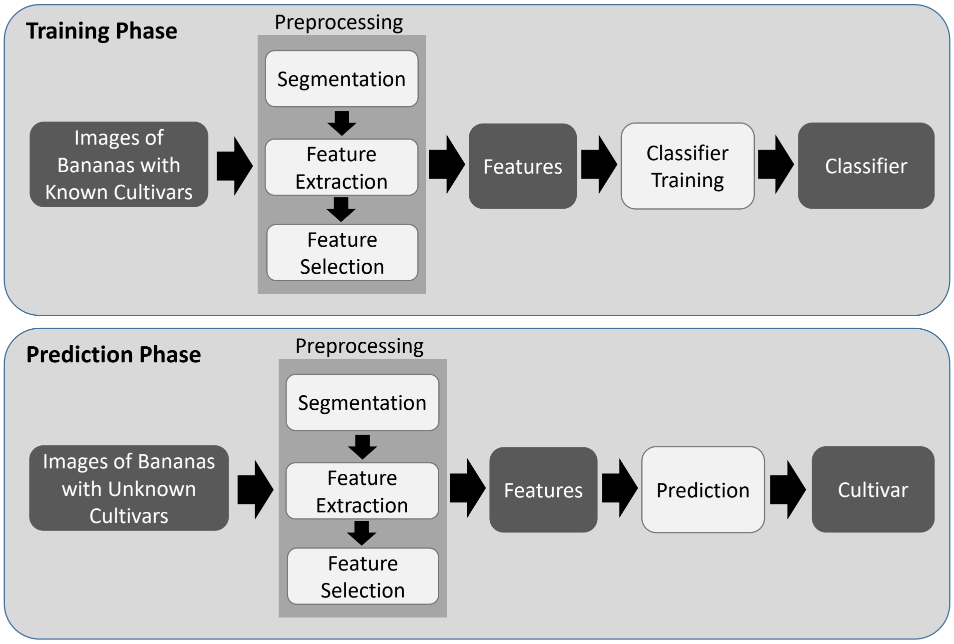

Fig. 1. Overall diagram showing different steps of the image analysis.

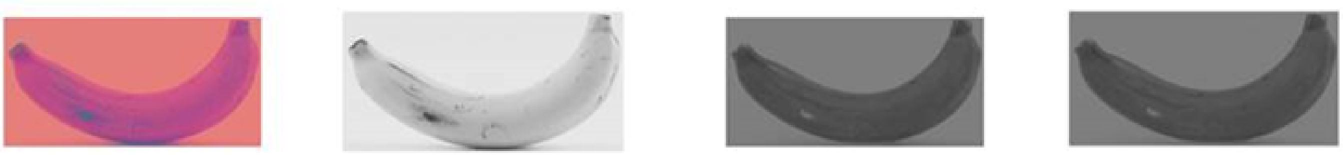

Fig. 2. Sample image of a banana in the YCbCr color system.

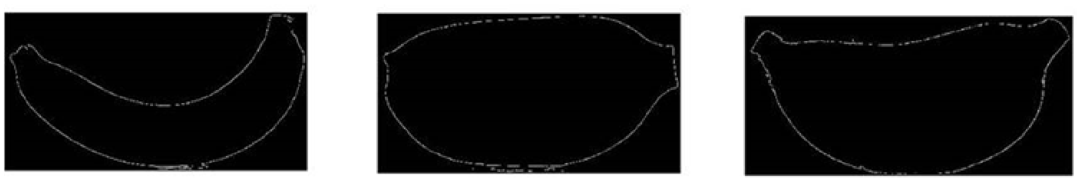

Fig. 3. Output of Canny edge detector for the three cultivars of banana. White dots represent edges, which are the object's outline.

dependence with the classes. The top K features are then used as input features to the model training step, where K is specified by the user. We compare the resulting neural network's prediction performance to choose the best metric for feature selection, as well as the optimal K.

B. Machine Learning Models

After the preprocessing step, the resulting features are used as input to train classification models. In this work, we compare seven different machine learning models in order to choose the optimal one. The models compared are decision tree, naive Bayes, averaged one-dependence estimators (AODE), Bayesian network, support vector machine (SVM), logistic regression, and neural network.

IV. Evaluation

In this section, we evaluate our methods by measuring the classification accuracy of the different configurations and models.

A. Dataset

We first collected the three cultivars of banana: Cavendish, Lady Finger, and Pisang Awak, from markets in Phuket, Thailand. We then carefully took the photos of the bananas, in order to reduce shadows and make the object as clear as possible. In order to show the variance, the bananas were divided into two datasets. The number of bananas for each cultivar in each dataset is shown in Table I.

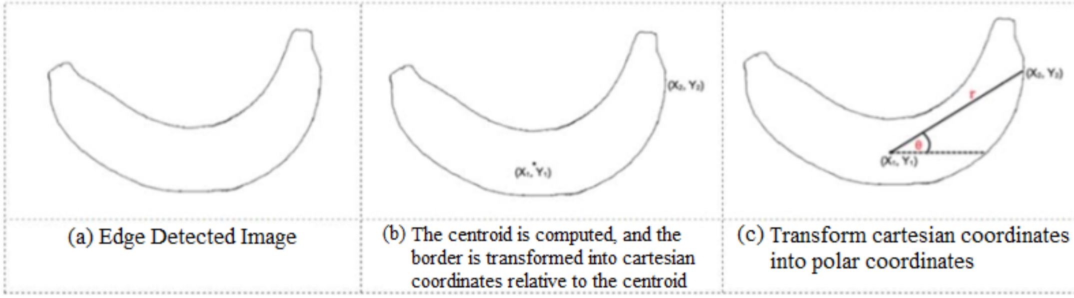

Fig. 4. Representing object's outline using polar coordinates.

TABLE I
NUMBER OF SAMPLES

Cultivar	Dataset 1	Dataset 2
Cavendish	36	39
Lady Finger	35	38
Pisang Awak	34	39
Total	**105**	**126**

B. Preprocessing

The raw digital images go through many preprocessing steps in order to create features for the classification models. The image at each step is shown in Figure 5. First, the image is resized to a common size so that scale of the object is not used in the classification process. The RGB image is then converted to the YCbCr color format. Only the Cb channel is used for the segmentation step. The grayscale image is then thresholded to produce a binary image. Holes within the objects are filled, and Canny's edge detection is applied to extract the outline of the object.

The outline of the object is converted to numerical features using polar coordinate as the representation. First, we find the centroid of the image. Then, for each of the 360 degrees, we compute the average distance between the centroid and the edge within that angle. This average distance becomes the potential input features of classification models.

C. Feature selection method

In this section, we compare three feature selection methods: Chi-Square, gain ratio, and information gain. Here, we fix the number of features of the reduced feature set to 100, and the classifier to neural network. We use 10-fold cross validation as the evaluation method.

The results are shown in Table II. As the AUC encompasses the overall accuracy under different decision boundaries, it serves as the best metric of prediction accuracy for our purpose. In dataset 1, the AUC is similar for all three methods. However, in dataset 2, gain ratio performs significantly worse. Overall, chi square is the best feature selection method overall. Thus, this is what we will use for further evaluation.

D. Number of features

To choose the optimal number of features, denoted K, we fix the feature selection method to chi square and the classifier to neural network, and vary K from 30 to 100 in increments of 10. For each value of K, we perform 10-fold cross validation to obtain prediction accuracy.

The results are shown in Table III. While there is some variance, the overall trends show that higher K result in higher prediction accuracy. Since the time needed to train the classifier is still small, we decided to use $K = 100$ for the next steps.

E. Classifier

In this section, we compare various machine learning classifiers so that the optimal one can be chosen. Classifiers included in the comparison are decision tree, naive Bayes, average one-dependence estimators (AODE), Bayesian network, support vector machine (SVM), logistic regression, and neural network. All classifiers are implemented in Weka machine learning software. 10-fold cross validation is used to obtain prediction accuracy.

The results are shown in Table IV. Overall, Bayesian network, naive Bayes, and AODE are the top three performing classifiers, while others are slightly worse. Bayesian network's overall accuracy is 84%, indicating that scale invariant shape analysis is effective for banana cultivar classification.

V. CONCLUSION

This paper presented banana cultivar detection using scale invariant shape analysis. We presented the appropriate preprocessing methods, and compared different feature selection methods, number of features, and machine learning models as classifier. In the image segmentation step during preprocessing, we find that using the Cb channel in the YCbCr color system works best. Canny edge detector is used to extract the outline of the object, and the outline is represented using polar coordinate. We evaluate our methods using two datasets, with 135 and 156 images of bananas, respectively. For feature selection, we found that chi square performs best. The optimal number of features is found to be 100. We found that Bayesian network is the best classifier, achieving overall AUC of 0.933 and overall accuracy of 84%, followed closely by naive Bayes and AODE.

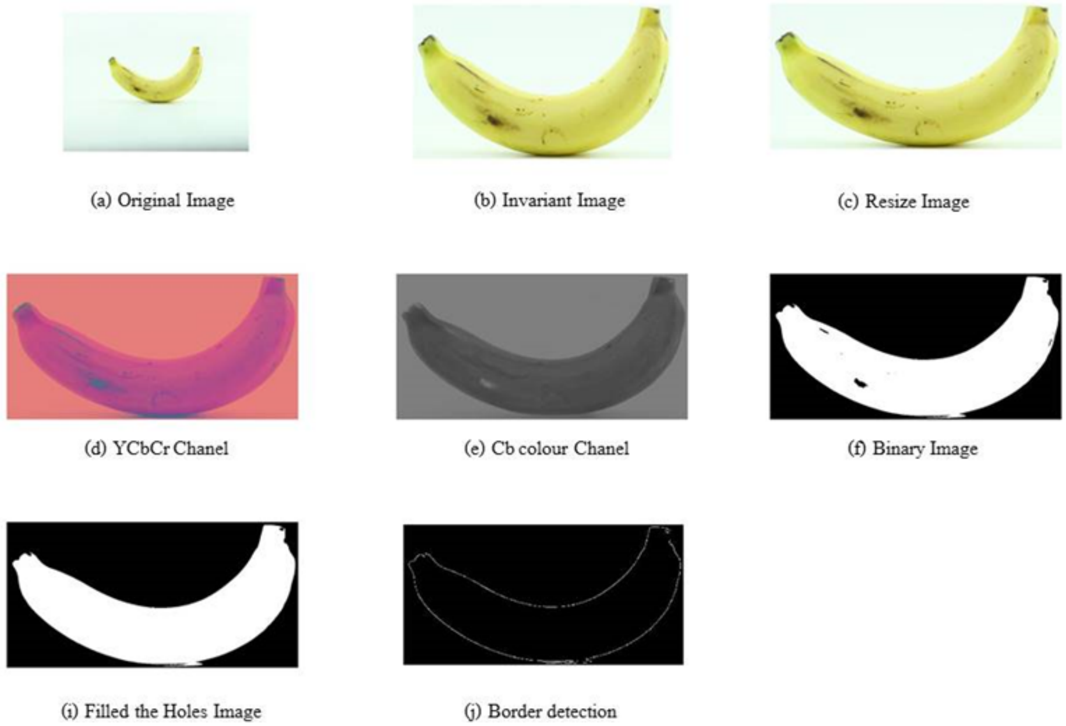

Fig. 5. Sample image at each preprocessing step.

TABLE II
RESULTS OF THREE FEATURE SELECTION METHODS. AC, AUC, SN, SP, PR, AND FM REFER TO ACCURACY, AREA UNDER THE ROC CURVE, SENSITIVITY, SPECIFICITY, PRECISION, AND F-MEASURE, RESPECTIVELY.

Method	**Dataset 1**						**Dataset 2**					
	AUC	AC	SN	SP	PR	FM	AUC	AC	SN	SP	PR	FM
Chi Square	0.887	0.778	0.778	0.890	0.777	0.777	0.949	0.853	0.853	0.859	0.731	0.716
Gain Ratio	0.897	0.785	0.785	0.894	0.787	0.785	0.852	0.853	0.853	0.922	0.853	0.852
Information Gain	0.874	0.748	0.748	0.876	0.752	0.750	0.924	0.821	0.821	0.907	0.817	0.817

TABLE III
CHI-SQUARE STATISTICS FOR EACH VALUE OF THE NUMBER OF FEATURES. AC, AUC, SN, SP, PR, AND FM REFER TO ACCURACY, AREA UNDER THE ROC CURVE, SENSITIVITY, SPECIFICITY, PRECISION, AND F-MEASURE, RESPECTIVELY.

Number of features (K)	**Dataset 1**						**Dataset 2**					
	AUC	AC	SN	SP	PR	FM	AUC	AC	SN	SP	PR	FM
30	0.876	0.793	0.973	0.818	0.784	0.793	0.908	0.801	0.801	0.899	0.806	0.803
40	0.881	0.748	0.748	0.876	0.747	0.747	0.930	0.853	0.853	0.923	0.853	0.853
50	0.880	0.733	0.733	0.868	0.727	0.729	0.933	0.833	0.833	0.913	0.835	0.834
60	0.871	0.726	0.726	0.825	0.725	0.725	0.925	0.840	0.840	0.915	0.840	0.839
70	0.862	0.704	0.704	0.853	0.965	0.967	0.939	0.840	0.840	0.915	0.841	0.840
80	0.863	0.733	0.733	0.868	0.734	0.733	0.937	0.840	0.840	0.916	0.841	0.840
90	0.913	0.748	0.748	0.888	0.748	0.748	0.945	0.856	0.856	0.928	0.867	0.865
100	0.887	0.778	0.778	0.890	0.777	0.777	0.949	0.853	0.853	0.922	0.853	0.852

TABLE IV
CLASSIFICATION ACCURACY FOR VARIOUS LEARNING METHODS. NUMBER OF FEATURES IS FIXED AT 100.

Number of features (K)	**Dataset 1**						**Dataset 2**					
	AUC	AC	SN	SP	PR	FM	AUC	AC	SN	SP	PR	FM
Decision Tree	0.899	0.844	0.844	0.924	0.856	0.846	0.830	0.724	0.724	0.859	0.731	0.717
Naive Bayes	0.937	0.852	0.852	0.927	0.852	0.852	0.924	0.821	0.821	0.904	0.820	0.814
AODE	0.935	0.844	0.844	0.927	0.849	0.845	0.941	0.808	0.808	0.896	0.815	0.793
Bayesian Network	0.940	0.852	0.852	0.927	0.852	0.852	0.925	0.827	0.827	0.908	0.826	0.821
SVM	0.813	0.711	0.711	0.857	0.707	0.709	0.900	0.853	0.853	0.921	0.854	0.852
Logistic Regression	0.783	0.726	0.726	0.865	0.723	0.724	0.924	0.840	0.849	0.917	0.837	0.835
Neural Network	0.877	0.778	0.778	0.890	0.777	0.777	0.949	0.853	0.853	0.922	0.853	0.852

REFERENCES

[1] T. Pholpho, S. Pathaveerat, and P. Sirisomboon, "Classification of longan fruit bruising using visible spectroscopy," *Journal of Food Engineering*, vol. 104, no. 1, pp. 169–172, 2011.

[2] C. Yang, W. S. Lee, and J. G. Williamson, "Classification of blueberry fruit and leaves based on spectral signatures," *Biosystems engineering*, vol. 113, no. 4, pp. 351–362, 2012.

[3] D. Cozzolino, W. Cynkar, N. Shah, and P. Smith, "Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality," *Food Research International*, vol. 44, no. 7, pp. 1888–1896, 2011.

[4] B. M. Nicolaï, T. Defraeye, B. De Ketelaere, E. Herremans, M. L. Hertog, W. Saeys, A. Torricelli, T. Vandendriessche, and P. Verboven, "Nondestructive measurement of fruit and vegetable quality," *Annual review of food science and technology*, vol. 5, pp. 285–312, 2014.

[5] M. R. C. Inácio, M. d. F. V. de Moura, and K. M. G. de Lima, "Classification and determination of total protein in milk powder using near infrared reflectance spectrometry and the successive projections algorithm for variable selection," *Vibrational Spectroscopy*, vol. 57, no. 2, pp. 342–345, 2011.

[6] M. Daszykowski, J. Orzel, M. Wrobel, H. Czarnik-Matusewicz, and B. Walczak, "Improvement of classification using robust soft classification rules for near-infrared reflectance spectral data," *Chemometrics and Intelligent Laboratory Systems*, vol. 109, no. 1, pp. 86–93, 2011.

[7] K. E. Kane and W. S. Lee, "Spectral sensing of different citrus varieties for precision agriculture," in *2006 ASAE Annual Meeting*. American Society of Agricultural and Biological Engineers, 2006, p. 1.

[8] C. Jones, J. Jones, and W. Lee, "Diagnosis of bacterial spot of tomato using spectral signatures," *Computers and Electronics in Agriculture*, vol. 74, no. 2, pp. 329–335, 2010.

[9] P. C. Marchal, D. M. Gila, J. G. García, and J. G. Ortega, "Expert system based on computer vision to estimate the content of impurities in olive oil samples," *Journal of Food Engineering*, vol. 119, no. 2, pp. 220–228, 2013.

[10] N. Thammachot, S. Chaiprapat, and K. Waiyakan, "Development of an image processing system in splendid squid grading," in *The 9th International Conference on Computing and InformationTechnology (IC2IT2013)*. Springer, 2013, pp. 175–183.

[11] S. Arivazhagan, R. N. Shebiah, S. S. Nidhyanandhan, and L. Ganesan, "Fruit recognition using color and texture features," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 1, no. 2, pp. 90–94, 2010.