Indonesian Journal of Electrical Engineering and Computer Science
Vol. 24, No. 3, December 2021, pp. 1523~1533
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v24.i3.pp1523-1533

Comparison of color-based feature extraction methods in banana leaf diseases classification using SVM and K-NN

Nur Sholehah Binti Mat Said, Hizmawati Madzin, Siti Khadijah Ali, Ng Seng Beng
Department of Multimedia, Faculty of Computer Science and Information Technology,
Universiti Putra Malaysia, Serdang, Malaysia

Article Info

Article history:

Received Jun 11, 2021
Revised Oct 14, 2021
Accepted Oct 27, 2021

Keywords:

Banana leaf diseases
Classification
Color feature extraction
Image processing
k-Nearest neighbors
Support vector machine

ABSTRACT

In Malaysia, banana is a top fruit production which contribute to the economy growth in agriculture field. Hence, it is significant to have a quality production of banana and important to detect the plant diseases at the early stage. There are many types of banana leaf diseases such as Banana Mosaic, Black Sigatoka and Yellow Sigatoka. These three diseases are related to color changes at banana. This research paper is an experiment based and need to identify the best color feature extraction method to classify banana leaf diseases. Total of 48 banana leaf images that are used in this research paper. Four types of color feature extraction methods which are color histogram, color moment, hue, saturation, and value (HSV) histogram and color auto correlogram are experimented to determine the best method for banana leaf diseases classification. While for the classifiers, support vector machine (SVM) and k-Nearest neighbors (k-NN) are used to evaluate the performance and accuracy of each color feature extraction methods. There are also preliminary experiments to identify accurate parameters to use during classification for both classifiers. Our experimental result express that HSV histogram is the best method to classify banana leaf diseases with 83.33% of accuracy and SVM classifier perform better compared to k-NN.

Corresponding Author:

Hizmawati Madzin
Department of Multimedia, Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
43400 Serdang, Selangor Darul Ehsan, Selangor, Malaysia
Email: hizmawati@upm.edu.my

1. INTRODUCTION

Banana ranks as the leading fruit in world agricultural production and trade. The average global banana production increases up to 116 million tons in 2017-2019 which approximate value of 31 billion USD. In Malaysia, banana is the most produced fruit which the production is around 350,000 tons per year [1]. According to Malaysia fruit crops statistic, the total hectares of banana planting in Malaysia are 30 thousand hectares including in East Malaysia. For Peninsular Malaysia there are total of 11 states that involved in banana planting and the top three states that dominate the planting area are Johor leading with 10,396.1 hectare, Pahang 4,654.5 hectare and Perak 2,860.4 hectare. Ministry of Agriculture, Malaysia also state that banana is one of the fruits that highly produced each year. Based on this statistic, this proves that banana is one of the important fruits in Malaysia for its agricultural industry and economy [2]. This statistic shows the high demand of banana globally and in Malaysia.

However, the drawback of banana planting is there are so many banana diseases that will affect the production such as fruit and leaf diseases. Banana leaf diseases such as Yellow Sigatoka, Black Sigatoka and

Banana Mosaic can cause loss to our farmers and agricultural industry. Generally, the disease will be appearing in the banana leaf which shown the changes of the color and texture of the leaf as shown in Figure 1. The traditional approach to detect the disease is based on human eye observation by the expertise. However, this method is time and money consuming, and this method is convenient only for small farms, but not suitable for large farm.

The aim of this paper is to identify banana leaf disease based on suitable color feature extraction method. In order to achieve this aim, there are three objectives that need to be achieved, which are; i) to study color feature extraction methods, ii) identify suitable color feature extraction method for banana leaf disease, and iii) evaluate the performance of identified color feature extraction method using two classifiers namely support vector machine (SVM) and k-Nearest neighbors (kNN). This paper will focused on four color features extraction methods namely color histogram, color moment, hue, saturation, and value (HSV) histogram and color auto correlogram. Total of 48 images which consist of 16 images of Banana Mosaic, Black Sigatoka and Yellow Sigatoka diseases will be used as image dataset. The evaluation of the performance each of color feature extraction and the classification accuracy will be using Waikato environment for knowledge analysis (WEKA) application. WEKA application is a set of machine learning algorithms that can be applied to a data set directly. This paper will be illustrated into several sections. First section is Introduction, follow with second section; literature Review. In section three is the proposed methodology, section four will be the results and finally is the conclusion.

Figure 1. Black Sigatoka disease

2. LITERATURE REVIEW

Classification in banana plant using digital image processing has been widely used by researchers and expertise. This process is significant to identify and classify many banana diseases especially on banana leaf [1]. There are many stages to classify banana leaf disease using image processing namely image acquisition, pre-processing image, feature extraction and classification. In this literature review,there will be two sub-sections which the first sub-section briefly explain about the types of banana leaf diseases and the last sub-section explains on how these diseases will be classify using image processing techniques.

2.1. Types of banana diseases

In Malaysia, banana is the second most grown fruit crop. Overall banana production has decreased due to the increasing threat of banana disease, high labour costs and marketing issues [3]. The development of computer technology detection system can support the farmers in the identification of diseases at initial stage and provide useful information for its control. There are several types of banana plant diseases such as Anthracnose, Black Sigatoka (black leaf streak), Cigar end rot, Yellow Sigatoka, Panama diseases, Rhizome rot, Banana Mosaic and many more. However, this research paper emphasizes on banana leaf diseases namely Black Sigatoka (black leaf streak), Yellow Sigatoka and Banana Mosaic.

Black Sigatoka is caused by Mycosphaerella Fijiensis which also known as black leaf streak. It is a leaf spot disease that will change the color of original healthy banana leaf [4]. The symptoms are red/brown/black with yellow border appear on banana leaf as shown in Figure 1. Black Sigatoka causes greater yield loss than Yellow Sigatoka, significant reductions in leaf area and premature ripening in exported fruits [5].

Figure 2 shows the image of Yellow Sigatoka, the diseases that are caused by Mycospharella Musicola and this disease also other leaf spot diseases that will change the color of the original healthy banana leaf [6]. Banana Mosaic or also known as cucumber mosaic virus (CMV) that are caused by a virus and the symptom visible on the leaf are discontinuous green between light green and dark green as shown in

Figure 3. This virus can result in growth defects, reduced suckering and misshapen fruit. Moreover, Banana Mosaic can course fruit rejection which lead to economic losses [7].

Banana Mosaic or also known as cucumber mosaic virus (CMV) that are caused by a virus and the symptom visible on the leaf are discontinuous green between light green and dark green as shown in Figure 3. This virus can result in growth defects, reduced suckering and misshapen fruit. Moreover, Banana Mosaic can couse fruit rejection which lead to economic losses [8].

Figure 2. Yellow Sigatoka disease

Figure 3. Banana Mosaic disease

2.2. Classification in agriculture

Classification in agriculture using digital image processing has been widely used by researchers and expertise. Pandurng and Lomte [8] describe four examples of classification in agriculture namely, i) crop management, ii) identification of nutrient deficiencies and plant content, iii) crop and land estimation and object tracking and iv) fruits quality inspection, sorting and grading.

This research paper emphasizes on fruits quality inspection and grading including fruits and leaf diseases classification that used image processing and machine learning to achieve the result. Dubey and Jalal [7] proposed apple fruit diseases identification using images using K-Means clustering technique for segmentation and multi-class support vector machine (SVM) to classify the images into different classes of diseases. For feature extraction the researchers use global color histogram (GCH) and color coherence vector (CCV). Result shows that CCV performs better than CGH. However, the drawbacks of K-means clustering technique is if they choose an inappropriate K in the beginning, it will affect the result of the experiment.

Tigadi and Sharma [9] proposed a software solution for automatic banana plant disease detection to identify final percentage of infection by using deep learning artificial neural network technique. They use color feature extraction method namely histogram of template (HOT) to extract the features for both test and training data. According to Jeyalakshmi and Radha [10], proposed a software is good and can replace the manual system of plant disease detection.

There is also research of guava leaf disease classification. Texture feature extraction methods namely scale invariant feature transform (SIFT), space extrema detection, keypoint localization, orientation assignment and keypoint descriptor are used in this research. SVM and k-NN are used as the classifiers. The result shows that SVM is slightly superior to k-NN in the classification. The drawback of this work is it is high computational process since so many methods have been used [11].

3. PROPOSED METHODOLOGY

Generally in plant disease classification, there are four stages to be completed which are i) image acquisition, ii) pre-processing image, iii) feature extraction and iv) classification. Every stages have several steps to be followed. This section explains the implementation detail of every stages. This section also describes the methods used in this research project such as the color-based feature extraction methods and types of classifiers.

3.1. Image acquisition

Image acquisition is to retrieve images from some source. It is initial stage in the image processing process as without an image, no processing is possible. In this research project, the dataset used 48 images from 3 different types of banana leaf diseases which are Banana Mosaic, Black Sigatoka and Yellow Sigatoka. These images are stored in .jpg format.

3.2. Image pre-processing

Every banana leaf image needs to be resized into 256x256 pixels. The purposed of resizing is to reduce high computational processing and to help improve the storage efficiency. It is significant to resize the image before further processing to obtain accurate results during classification.

3.3. Color feature extraction in banana leaf images

There are many features can be extracted from banana leaf such as color, texture, morphology and structure in order to identify banana leaf disease [12]. However, in this research paper we are focusing using color feature extraction method. Figure 4 shows the flowchart of extracting color-based features. Color is a useful descriptor in low level feature because of its relatively robust to background compilation and an independent image in the aspect of size and orientation [13]. In this section, four types of color feature extraction methods will be described and later will be compare the performance of these methods in classifying three banana leaf diseases. The color feature extraction methods are color histogram, color moment, color auto correlogram and HSV histogram [14].

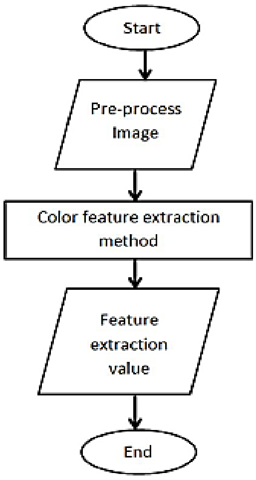

Figure 4. Flowchart for future extraction process

3.3.1. Color histogram

Color histogram is a type of bar graph, where each of the bar represent the color from color space that being used [15]. Color histogram can be built from any kind of color space such as RGB, HSV and others color space. Statistically, color histogram is a way to present approximate joint probability of the values of its color space that being used by splitting the range of data into equally sized bins.

Some commonly used parameters for color histogram are mean, standard deviation, skewness and kurtosis. Color descriptors for color histogram can be obtained globally by extracting the information on its mean values [16]. Therefore, this research paper used mean values as the parameter for color histogram to extract features of image dataset.

The mean (μ) for the range of intensity value is computed as (1);

$$\mu = \sum_{i=LB}^{UB} r_i \, p(r_i) \quad (1)$$

where r_i is the intensity values and p is the size of the image. However, the drawback of color histogram for classification is that the representation is dependent of the color of the object being studied, ignoring its shape and texture [17], [18].

3.3.2. Color moment

Color moments characterize the color distribution in an image that can interpret by its probability distribution. Color moments are scaling and rotation invariant. In this research paper, first three-color moments namely mean value, standard deviation and skewness are used for feature extraction [19]. This method extracts both shape and color information which suitable use under changing lighting conditions. However, color moments cannot handle occlusion very successfully [20]. In (2) shows the calculation of mean value in color moments.

$$E_i = \sum_{j=1}^{N} \frac{1}{N} p_{ij} \quad (2)$$

where p_{ij} is the value of the j-th pixel of the image at the i-th color channel and N is the number of pixels in the image.

The second color moment is standard deviation and can be obtain by calculating the square root of the variance of the color distribution. Standard deviation is defined by the (3).

$$\sigma_i = \sqrt{(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)x^2)} \quad (3)$$

Where E_i is the mean value, for the i-th color channel of the image which represent the R, G and B channel.

The third color moment is skewness that is used to measure the asymmetric distribution of colors in an image. It can be concluded that in color moment there are total nine values as for each color R, G and B space components. The formula of skewness is defined in the (4).

$$s_i = \sqrt[3]{(\frac{1}{N}\sum_{j=1}^{N}(p_{ij} - E_i)^3)} \quad (4)$$

Where E_i is the mean value, for the i-th color channel of the image which represent the R, G and B channel.

3.3.3. Color auto correlogram

Color correlogram explains how spatial correlations of pairs of colors change with distance while color auto correlogram is a subset of color correlogram, only captures the spatial correlation between identical colors and consist of rows of the form(i,j) [20]. RGB color space is used and color quantization for each image is produced which result to sixty-four colors because of four levels for each color space channel. Four levels for R, G and B, (4x4x4=64). The values from all sixty-four colors will be stored as feature extraction values. The correlogram of the image I is defined for level g_i,g_j at a distance d using the (5).

$$\gamma_{g_i,g_j}^{d}(I) \equiv Pr_{p1\in I_{g_i},p2\in I}\left[p_2 \in I_{g_j}\,||p_1 - p_2 = d|\right] \quad (5)$$

In (5) gives the probability of any pixel p_1 of level g_i, and pixel p_2 at a distance d in certain direction from level g_i. In (6) defines how auto correlogram captures the spatial correlation of identical levels.

$$\propto_{g}^{d}(I) = \gamma_{g_i,g_j}^{d}(I) \quad (6)$$

3.3.4. HSV Histogram

HSV histogram is generated using similar approach to the RGB color space. The features in HSV histogram can be described as i) hue scale which divided into eight groups, ii) saturation scale which divided into two groups, and iii) the intensity scale which divided into four groups [21]. Therefore, a total of 64 cells to represent a 64-component HSV color histogram. For the combination values of H, S, and V, the corresponding histogram component is determined. The respective histogram component is updated by one for each pixel having the corresponding color combination.

4. CLASSIFIER FOR BANANA LEAF DISEASE CLASSIFICATION

In machine learning, classification refers to process of predicting the class label based on given example of input data. The main goal is to identify which class the new data will fall into. As for this research project the issue is to classify the banana leaf image into which class of banana leaf disease [22]. Classifier is used as an algorithm to map banana leaf image to a specific class of banana leaf disease. Classifier needs to train banana leaf images to understand how the given input variables are related to the class as shown in Figure 5. There are many types of classifiers such as k-Nearest neighbors (k-NN), decision tree, Naive Bayes (NB), artificial neural networks (ANN) and support vector machine (SVM). However, in this research paper, only two classifiers will be used and compare the accuracy performance of k-NN and SVM. This is because these two classifiers are the most suitable for plant disease classification [23]. The accuracy measurement of a classifier is given as the percentage of total correct predictions divided by the total number of input data.

4.1. Support vector machine (SVM)

SVM is widely used for classification and pattern recognition due efficient and simple classifier algorithm. SVM transfers the original features set to a high or infinite-dimensional space by using the kernel function. There are many types of kernels can be used in SVM model such as linear, polynomial, radial basis functions (RBF) [24]. As for this research project, preliminary experiment will be conducted to identify the suitable seed and fold to achieve highest accuracy for banana leaf disease classification. In (7) shows SVM scoring function. Scoring function are used to compute the score for an input vector x as in a trained (SVM) has a scoring function to compute a score for a new input.

$$\sum_{i=1}^{m} \propto_i y^{(i)} K\left(x^{(i)}, x\right) + b \tag{7}$$

In (7) operates every data point in the training set (i = 1 through m). x is the input vector while K is the kernel function to operates on two vectors. b is the scalar value [25].

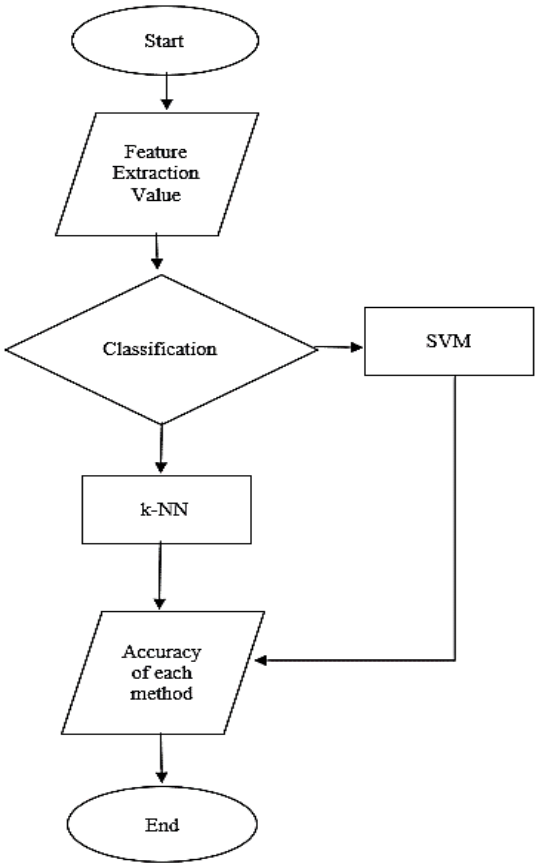

Figure 5. Flowchart for classification process

4.2. k-Nearest neighbors (k-NN)

k-NN is often used in classification, image recognition, and decision-making models. Initially the algorithm transforms data points into feature vectors then find distance between data points of x and y using mathematical formula such as Euclidean distance as shown in (8). It then finds the probability of these points being similar to the test data and classifies it based on the highest probabilities [26].

$$\sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \tag{8}$$

4.3. Research framework

The aim of this research project is to classify disease on banana leaf image with three categories of banana leaf disease namely Black Sigatoka, Yellow Sigatoka and Banana Mosaic. Figure 6 shows the framework of banana leaf disease classification. It is starts with image acquisition, where put the images into a database. The images will be resized to 256x256 and need to convert to HSV format. The next step is feature extraction process of the banana leaf images. This stage there will be performance comparison between four color feature extraction methods which are Color Histogram, Color Moment, Color auto correlogram and HSV Histogram. Later is to classify the images into banana leaf disease. The classification involved two classifiers namely SVM and k-NN. Then the best classifier will be chose based on the highest accuracy achieved between the classifiers.

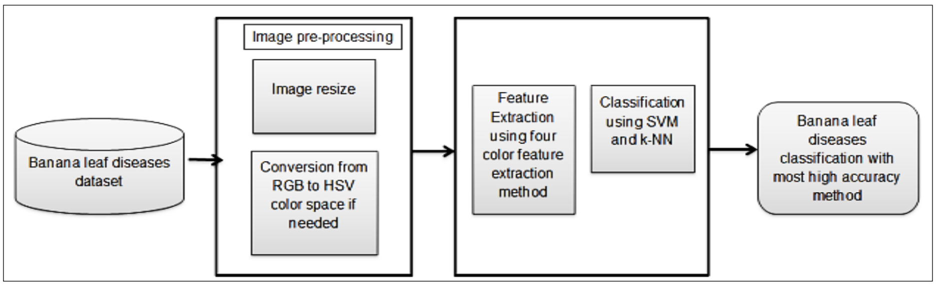

Figure 6. Banana leaf disease classification framework

4.4. Experimental Setup

There are several experiments that need to be executed in order to identify the suitable color feature method for feature extraction and the classifier to classify the banana leaf disease. There are three experiments involved i) identify the best parameters of seed and fold for SVM classifier, ii) identify the accurate *k* fold value for k-NN classifier, and iii) identify the best color-based feature extraction method to classify banana leaf disease.

For the first experiment, there are two parameters need to be determined when using SVM classifier which are seed and fold values. These values are significant to access high accuracy in classifying banana leaf disease. Followed by second experiment which to determine *k* and fold values for k-NN method. Both classifiers later will be compared in order to identify which classifier is suitable to classify banana leaf disease. The input data for both experiments is 76 feature vectors extracted from four color-based feature extraction. The third experiment is to compare the classification accuracy values of each color-based feature extraction methods namely color histogram, color moment, color auto correlogram and HSV Histogram methods using both SVM and k-NN classifiers.

5. EXPERIMENTAL RESULTS

The description of research framework and experimentation setup have been explained in previous section. This section will be emphasized on the results of the experiments of i) identify the best parameters of seed and fold for SVM classifier ii) identify the accurate *k* fold value for k-NN classifier iii) identify the best color-based feature extraction method to classify banana leaf disease.

5.1. Seed and fold parameter for SVM

In this section explains the process of selection the suitable values of seed and fold values for SVM classifier in WEKA tool. Seed value is used to fix the random numbers that are being generated to split the input data. However, if the seed value is fixed to some specific value, the same split every time execution which is useful to make the scores reproducible. The fold value is to check for over-fitting of input data. Fold value determines the subsets of the available input data to be trained using the hyperparameters of predictive models.

Initially the experiment is to look for the seed values. Therefore, the experiment is to compare the seed values between 1 to 10 and used fixed fold value of 10. Table 1 shows the best seed values to classify banana leaf diseases with high accuracy rate are seed 2, 6 and 7.

Next is the experiment to determine the best fold value for the seed values of 2, 6 and 7. Table 2 shows the list of fold of SVM classifier with high accuracy results. From the table it shows that the accuracy results for the combination of seed values of 2, 6 and 7 fold values of 3, 4, and 8 are 68.75%. Table 2 shows that combination of seed 7 and fold 8 achieve highest value of accuracy which is 70.83%. Therefore, to classify banana leaf disease using SVM the suitable seed and fold is seed 7 and fold 8.

Table 1. Comparison of seed values accuracy rate

Seed	Fold	Accuracy
1	10	56.25%
2	10	58.33%
3	10	54.17%
4	10	56.25%
5	10	54.17%
6	10	58.33%
7	10	58.33%
8	10	52.08%
9	10	56.25%
10	10	56.25%

Table 2. Best seed and fold values for classification of banana leaf diseases

Seed	Fold	Accuracy
2	3	68.75%
2	8	68.75%
6	4	68.75%
6	8	68.75%
7	8	70.83%

5.2. K value and fold parameters for k-NN

This section explains the second experiment results to identify the best *k* value and fold parameter for banana leaf diseases classification using k-NN method. Initially the experiment is to identify the best *k* value with fix fold value of 10. Table 3 shows *k* value of 7 obtain highest accuracy rate of 68.75% compared to other *k* values.

Table 3. Best k values of k-NN for classification of banana leaf diseases

Seed	Fold	Accuracy
1	10	62.50%
2	10	47.92%
3	10	60.42%
4	10	56.25%
5	10	60.42%
6	10	54.17%
7	10	68.75%
8	10	64.58%
9	10	62.50%
10	10	64.58%

Then next is to identify the best fold value for k-NN. Table 4 shows the best combination of *k* and fold parameters for banana leaf disease classification. From the table it shows that fold values of 2 and 8 obtain high accuracy rate of 70.833%. Therefore, *k* value for k-NN=7 and fold parameters of 2 and 8 suitable to classify banana leaf diseases.

Table 4. Best *k* values of k-NN and fold parameters for classification of banana leaf diseases

k	Fold	Accuracy
7	1	-
7	2	70.83%
7	3	66.67%
7	4	68.75%
7	5	66.67%
7	6	64.58%
7	7	68.75%
7	8	70.83%
7	9	60.42%
7	10	68.75%

5.3. Comparison of color-based feature extraction methods using SVM and k-NN classifiers

This section shows the comparison of accuracy rate of four-color feature extraction methods namely color histogram, color moment, color auto correlogram and HSV histogram. Table 5 shows the accuracy rate of SVM and k-NN classifiers for color feature extraction method. The accuracy rate for color auto correlogram using SVM is 52.08% while for k-NN classifier is 12% lower which is 45.83%. For color histogram method, the accuracy rate for both SVM and k-NN classifiers are 70.83%. For color moment method, accuracy rate of SVM classifier is 5% higher which value of 77.08% compared to k-NN classifier which is 72.91%. The final color feature extraction method, which is HSV histogram, obtained 83.33% of accuracy rate via SVM classifier which is the highest accuracy rate compared to all methods. Therefore, the best color-based feature extraction method for banana leaf disease classification is by using HSV histogram and the best classifier is SVM with seed value of 7 and fold value of 8.

Figure 7 shows the percentage of correctly classified the three banana leaf disease using HSV histogram method. HSV histogram has the ability to separate chromatic and achromatic components which make it closer to human conceptual understanding compared to RGB color space [27], [28]. From the figure it shows that Banana Mosaic disease obtain highest rate of correctly classified which is 93.75% using SVM classifier. Then followed by Black Sigatoka disease which is 81.25% and finally Yellow Sigatoka disease which is 75%.

Table 5. Accuracy rate for color feature extraction methods using SVM and k-NN classifier

Feature Extraction	Accuracy SVM	Accuracy k-NN
Color auto correlogram	52.08%	45.83%
Color histogram	70.83%	70.83%
Color moment	77.08%	72.92%
HSV histogram	83.33%	77.08%

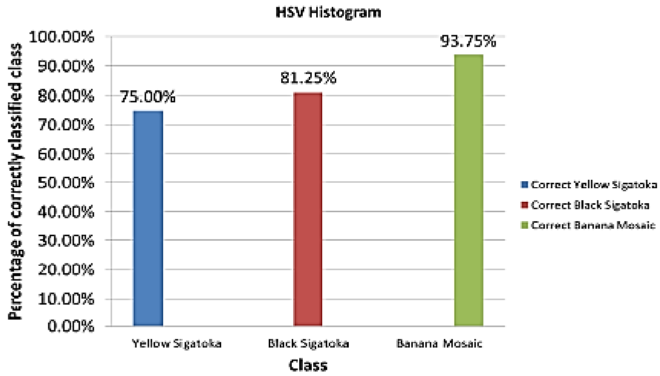

Figure 7. Percentage of correctly classified the banana leaf disease using HSV Histogram color-based feature extraction method

6. CONCLUSIONS

This research paper discusses on the performance of four-color feature extraction methods namely color histogram, color moment, HSV histogram and color auto correlogram in order to classify the banana leaf diseases. There are three banana leaf diseases to classify which are Black Sigatoka, Yellow Sigatoka and Banana Mosaic. Three experiments have been explained in this paper which are: i) Identify the best parameters of seed and fold for SVM classifier, which the results are seed value = 7 and fold value = 8; ii) Identify the accurate *k* and fold value for k-NN classifier, which the results are *k* value = 7 and fold value = 8; and iii) identify the best color-based feature extraction method to classify banana leaf disease. The finding shows that HSV histogram is the best color-based feature extraction for banana leaf diseases classification with accuracy rate of 83.33% using SVM classifier. For future work for this research paper, there will be features such as shape and texture to increase accuracy rate for banana leaf diseases classification. Furthermore, to apply deep learning approach to classify the banana leaf classification. However, to do that, more dataset needed to train and modelling the features for classification purposes.

REFERENCES

[1] S. Sairam, R. Selvarajan, S. Handanahalli, and S. Venkataraman, "Towards understanding the structure of the capsid of Banana Bunchy Top Virus," *BioRxiv*, 2020, doi: 10.1101/2020.02.12.945212.
[2] M. G. Selvaraj *et al*., "Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 110-124, 2020, doi: 10.1016/j.isprsjprs.2020.08.025.
[3] N. M. M. Nik Rozana, C. Suntharalingam, and M. F. Othman, "Competitiveness of Malaysia's Fruits in the Global Market: Revealed Comparative Advantage Analysis," *Malaysian Journal of Mathematical Sciences*, pp. 143-157, 2017.
[4] V. Tyagi, "Color Feature" in *Content-Based Image Retrieval*, Singapore: Springer, 2017, pp. 133-159, doi: 10.1007/978-981-10-6759-4_7.
[5] S. Deenan, S. Janakiraman, and S. Nagachandrabose, "Image Segmentation Algorithms for Banana Leaf Disease Diagnosis," *Journal of The Institution of Engineers (India): Series C*, vol. 101, no. 5, pp. 807-820, 2020, doi: 10.1007/s40032-020-00592-5.
[6] J. Chaki and N. Dey, "Introduction to Image Color Feature," in *Image Color Feature Extraction Techniques*, Singapore: Springer, 2021, pp. 1-28, doi: 10.1007/978-981-15-5761-3_1.
[7] S. R. Dubey and A. S. Jalal, 2012 "Adapted Approach for Fruit Disease Identification using Images," *Image Processing Concepts Methodologies Tools & Applications*, 2012, pp. 1395-1409, doi: 10.4018/ijcvip.2012070104.
[8] J. A. Pandurng and S. S. Lomte, "Digital Image Processing Applications in Agriculture: A Survey," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 5, no. 3, pp. 622-624, 2015.

[9] B. Tigadi and B. Sharma, “Banana Plant Disease Detection and Grading using Image Processing”, *International Journal of Engineering Science and Computing,* vol. 6, no. 6, pp. 6512-6516, 2016, doi: 10.4010/2016.1565.
[10] S. Jeyalakshmi and R. Radha “A Review on Diagnosis of Nutrient Deficiency Symptoms in Plant Leaf Image Using Digital Image Processing,” *ICTACT Journal on Image and Video Processing*, vol. 7, no. 4, pp. 1515-1524, 2017, doi: 10.21917/ijivp.2017.0216.
[11] M. Thilagavathi and S. Abirami, “Application of image processing in diagnosing guava leaf diseases,” *International Journal of Scientific Research and Management (IJSRM),* vol. 5, no. 7, pp. 5927-5933, 2017, doi: 10.18535/ijsrm/v5i7.19.
[12] G. Mandloi, “A Survey on Feature Extraction Techniques for Color Images,” *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 4615-4620, 2014.
[13] H. Singh and D. Agrawal, “Result Analysis and Comparison of Hybrid Method based on Local Binary Pattern (LBP) and Color Moment (CM) for Efficient Image Retrieval,” *International Journal of Computer* Applications, vol. 159, no. 5, pp. 14-19, 2017, doi: 10.5120/ijca2017912515.
[14] S. R. Singh and S. Kohli, “Enhanced CBIR using Color Moments HSV Histogram Color Auto Correlogram and Gabor Texture,” *International Journal of Computer Systems*, vol. 2, no. 5, pp. 161-165, 2015.
[15] M. Bhange and H. A. Hingoliwala, “Smart Farming: Pomegranate Disease Detection Using Image Processing,” *Procedia Computer Science,* vol. 58, pp. 280-288, 2015, doi: 10.1016/j.procs.2015.08.022.
[16] E. Prasetyo, R. D. Adityo, N. Suciati, and C. Fatichah, “Mango leaf image segmentation on HSV and YCbCr color spaces using Otsu thresholding,” *2017 3rd International Conference on Science and Technology - Computer (ICST)*, 2017, pp. 99-103, doi: 10.1109/ICSTC.2017.8011860.
[17] V. Vapnik and R. Izmailov, “Knowledge transfer in SVM and Neural Networks,” *Annals of Mathematics and Artificial Intelligence*, vol. 81, no. 1, pp. 3-19, 2017, doi: 10.1007/s10472-017-9538-x.
[18] L-Y. Hu, M-W. Huang, S-W. Ke, and C-F. Tsai, “The distance function effect on k-nearest neighbor classification for medical datasets” *SpringerPlus*, vol. 5, no. 1, 2016, doi:10.1186/s40064-016-2941-7.
[19] F. Garcia-Lamont, J. Cervantes, A. López, and L. Rodriguez, “Segmentation of images by color features: A survey,” *Neurocomputing*, vol. 292, no. 1-27, 2018, doi: 10.1016/j.neucom.2018.01.091.
[20] N. Varish *et al*., “Image Retrieval Scheme Using Quantized Bins of Color Image Components and Adaptive Tetrolet Transform,” in *IEEE Access*, vol. 8, pp. 117639-117665, 2020, doi: 10.1109/ACCESS.2020.3003911.
[21] U. Erkut, F. Bostancıoğlu, M. Erten, A. M. Özbayoğlu, and E. Solak, “HSV Color Histogram Based Image Retrieval with Background Elimination,” *2019 1st International Informatics and Software Engineering Conference (UBMYK)*, 2019, pp. 1-5, doi: 10.1109/UBMYK48245.2019.8965513.
[22] R. Sharma, S. S. Kamble, A. Gunasekaran, V. Kumar, and A. Kumar, “A systematic literature review on machine learning applications for sustainable agriculture supply chain performance,” *Computers & Operations Research*, vol. 119, 2020, doi: 10.1016/j.cor.2020.104926.
[23] Y. Zhang, G. Wang, F-L. Chung, and S. Wang, “Support vector machines with the known feature-evolution priors,” *Knowledge-Based Systems*, vol. 223, 2021, doi: 10.1016/j.knosys.2021.107048.
[24] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, “Tomatoes classification using K-NN based on GLCM and HSV color space,” *2017 International Conference on Innovative and Creative Information Technology (ICITech)*, 2017, pp. 1-6, doi: 10.1109/INNOCIT.2017.8319133.
[25] S. S. Jasim and A. A. M Al-Taei, “A Comparison between SVM and K-NN for classification of Plant Diseases,” *Diyala Journal for Pure Science*, vol. 14, no. 2, pp. 94-105, 2018, doi: 10.24237/djps.1402.383B.
[26] A. Dadrasnia, M. M. Usman, R. Omar, S. Ismail, and R. Abdullah, “Potential use of Bacillus genus to control of bananas diseases: Approaches toward high yield production and sustainable management,” *Journal of King Saud University-Science*, vol. 32, no. 4, pp. 2336-2342, 2020, doi: 10.1016/j.jksus.2020.03.011.
[27] V. Ahamedemujtaba, A. K. Cherian, P. M Namitha, V. I. Louis, and S. Beena, “Detection and biophysical characterization studies of cucumber mosaic virus causing infectious chlorosis disease of banana,” *Journal of Pharmacognosy and Phytochemistry*, vol. 8, no. 1, pp. 2606-2611, 2019.
[28] D. Som *et al*., “A review on biology and study of major viral diseases in banana” *The Pharma Innovation Journal*, vol. 7, no. 12, pp. 218-222, 2018.

BIOGRAPHIES OF AUTHORS

Nur Sholehah binti Mat Said is a graduate student from Universiti Putra Malaysia in Bachelor of Computer Science (Multimedia). She was born in Alor Setar, Kedah. She was really actives during her studies in the university and participated in many leadership programs. Currently, she is involving in education and videography field.

Hizmawati Madzin was born in Selangor, Malaysia in 1980. Received her bachelor degree in computer science major in software engineering from Universiti Putra Malaysia in 2003. She pursues her master's in software engineering (MSE) in 2007 and PhD degree in multimedia information retrieval in 2013. Both master and PhD degrees from University of Malaya, Kuala Lumpur. Currently is a senior lecturer of Multimedia at Universiti Putra Malaysia. Her work focuses specifically on the medical imaging and image processing in agriculture and now moving towards to deep learning and artificial intelligent. She also involved in interactive education technology projects such as augmented reality and virtual reality. She has won several gold awards in innovation teaching and learning exhibitions. She also has published more than 15 journals in indexed journal. She loves to do community activities with her students to help the unfortunate community and earn industry grant to run programs with community.

Siti Khadijah Ali was born in Johor, Malaysia in 1986. She received the B.Sc degree in applied mathematical sciences from Universiti Sains Malaysia, Malaysia in 2008, Then, she received her M.Sc degrees in computer graphics from the Universiti Putra Malaysia, Malaysia, in 2012 and the Ph.D. degree in automatic control and system engineering from University of Sheffield, Sheffield, United Kingdom, in 2018. From 2008 to 2017, she was a tutor in Universiti Putra Malaysia, Malaysia. Since 2018, she has been a Senior Lecturer with the Multimedia Department, Faculty Computer Science and Information Technology, Universiti Putra Malaysia. She is the author of more than 10 articles and 3 inventions. She also a principal investigator for several internal grants by Universiti Putra Malaysia, also a grant in national level by Higher Ministry of Malaysia Education (Fundamental Research Grant Scheme 2019). Her research interests include physics-based animation/simulation, control system application (specifically in modelling the control system for an exoskeleton) and computer graphics. She was a recipient of the International Conference on Applied Engineering (ICAE 2018) for the best paper award.

Ng Seng Beng, received his B.Sc., M.Sc. and PhD in Computer Science from Universiti Putra Malaysia (UPM), in 2004, 2007 and 2015 respectively. He is currently a senior lecture in the Faculty of Computer Science and Information Technology, UPM and a member of the Computer Graphics, Vision and Visualisation (CGV2) research group. His research interest is point cloud processing, image-based modelling, and geometrical reverse engineering.