International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Segmentation and Classification of Arecanut Bunches before harvesting

Dhanesha R.[1]
DOS in Computer Science,
Davangere University, Shivagangothri Campus,
Davangere, Karnataka, India.
e-mail: dhanesh025@davangereuniversity.ac.in

Umesha D K.[2]
Department of Computer Science and Engineering,
Shivasharane Akkamahadevi Governament Womens Polytechnic, Shiralakoppa,
Shivamogga, Karnataka, India.
e-mail: umeshadk@gmail.com

Anitha A C.[3]**, Shrinivasa Naika C L.**[4]
DOS in Computer |Science and Engineering,
UBDT Colloege of Engineering,
Davangere, Karnataka, India.
e-mail: anithakrishna2008@gmail.com[3], naika2k6@gmail.com[4]

Sudhakar H R.[5]
DOS in Computer Science,
Davangere University,Shivagangothri campus,
Davangere, Karnataka, India.
e-mail: sudhak05@gmail.com

Abstract—In the agriculture sector, arecanuts are an extremely valuable crop. The price of an arecanut depends on its stage of ripeness. As a result of a lack of expertise in judging the maturity level of arecanut bunches before harvest, farmers often lose profit. Precision agricultural techniques based on image processing and computer vision have recently assisted farmers in determining crop maturity quality. Precision agricultural techniques based on image processing and computer vision have recently assisted farmers in determining crop maturity quality. Therefore, accuracy in segmenting arecanut bunches is vital for automated maturity level identification. In proposed work S-channel, Cr-channel and Pr-channel of HSV, YCbCr and YPbPr respectively color models are used to segment arecanut bunches. Three color features (i.e., mean of an arecanut bunch image on red, green, and blue bands), and two texture features (i.e, correlation, and entropy) were used in classification procedure. A random forest classifier was employed to classify maturity levels of arecanut bunch. This experiment uses a dataset of 1017 images of arecanut bunches to assess the segmentation performance of each color model. As a result of the experiment, it has been concluded that the S-channel of the HSV color model was effective in segmenting arecanut bunches from input images. The proposed methodology effectively classifies arecanut bunch maturity levels with an accuracy of 87.80%.

Keywords-Arecanut bunches; S-channel of HSV; Cr-channel of YCbCr; Pr-channel of YPbPr; Segmentation; classifier; Random-forest.

I. INTRODUCTION

In terms of economic development, agriculture plays an important role in India [1]. There are 4.73 lakh hectares (HA) of arecanut cultivation in India with a production of 7.06 lakh tons [2]. In India, arecanuts are mostly grown in the states of Karnataka and Kerala. In addition to agriculture, arecanuts play a significant role in Indian religion, culture, and economics. In terms of usage, it comes in fourth place after nicotine, alcohol, and caffeine in the world. Additionally, arecanuts are used to make wine, soft drinks, soaps and ayurvedic medicine, as well as for chewing with betel leaves [2].

Depending on the type of arecanut plant, it can grow up to 30 to 40 feet tall. Its stems have a round bark that measures 15 to 20 centimeters in diameter. Depending on the variety of the plant, these measurements vary. Figure 1, shows parts of arecanut plant. Spadix contains both male and female flowers, which are the starting point of arecanut bunches. The female flowers mature into arecanuts bunch which would take 6 months to reach maturity level and male flowers matures as

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

waste inflorescence. Each plant would have 3-5 Spadix which sprouts at different time and reach maturity level at different time of the year. Maturity level of arecanuts can be classified into two levels such as, immature and matured arecanuts. The matured arecanuts are categorized in to Hasa, Bette, Mine and Gotu depending upon the quality of maturity post harvesting process [3]. Harvesting and processing of arecanut is complete labor intensive work. It includes minimum two expertises with required skills, one with experience of climbing the tree for plucking bunch and another one person to hold gunny bag or long rope to catch the plucked bunch from the tree. Farmer has to invest huge amount towards expertise labors, While harvesting, By cracking one spadix, farmers would test the maturity level of a nut from a bunch. If it passes the test, labors would cut the bunch from all the nearby plants sprouted at similar time. The presence of distinct arecanut plantations and nutritional changes noticed by the plants, it is a monumental undertaking to keep track of each bunch's maturity status in a plantation of one hectare. Hence, the probability of harvesting unripened or over ripened bunches is high, due to which the harvest fetch less profit in the market as analysed in Arecanut market Analysis.

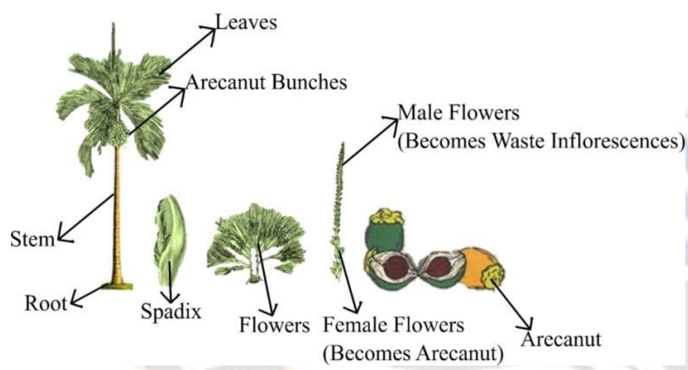


Figure 1**:** Arecanut_parts

Arecanut is a tropical plantation crop cultivated primarily for its kernel. This kernel is obtained from its fruit [4]. The Figure 2, (a), (b) & (c) shows the raw fruits, raw kernels and the processed arecanut.

(a) (b) (c)

Figure 2: Arecanut (a)Raw Fruit (b) Raw Kernel (c) Processed arecanut

Figure 3, represents the analysis of Arecanut in Market. The Figure 3 (a), (b), (c) and (d) were plotted by calculating day average price of different categorized arecanuts depending upon maturity level in the duration of four years. The market prices for each variety collected for the duration (01 Jan 2017 to 31 December 2020) from MEMCOS [5] web site. It is one of the important markets of arecanut in Karnataka which provides market and storage facilities to arecanut farmers and traders. It can be observed from Figure 3, that as the maturity level increases the cost decreases. Hasa fetched good price compare to all other varieties of Arecanut. As harvesting time increases maturity level of arecanut changes from Hasa to Gotu. Arecanuts market price is determined by its maturity level. Arecanut costs decrease as maturity levels increase. Prior to harvest, farmers should concentrate on maturity level (Hasa) in order to maximize crop profitability.

In present scenario harvesting of arecanut bunch is done by labors (gone gowdru) manually and for non-expertise it is very complicated job. Nowadays, it is difficult to identify arecanut bunch maturity levels because of a lack of expert labor. Two problems may cause when labor goes to harvest arecanut bunch to arecanut plantation. First, if labor harvests immature arecanut bunch then no use of Arecanuts and second, if failed to notice matured arecanut bunch it will become over-matured arecanut bunch. In both the cases it is loss for Farmer. So, Automation of identifying maturity level before harvesting process is needed to make Farmer profitable.

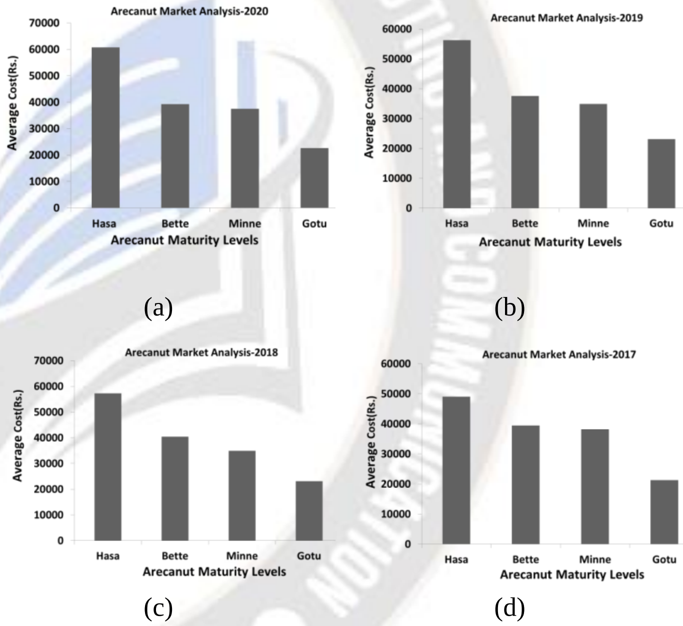


Figure 3**:** Arecanut Market Analysis

To maximize profitability of arecanut crop in market: First, farmer need to concentrate on identifying and harvesting matured arecanut bunches in-time. Second, fruit industries have found and benefited from the use of machine vision and image processing techniques. Third, Using availability of arecanut plant climbing machine to harvest arecanut Bunches. Maturity level of Arecanuts can be classified into two levels such as, immature and matured Arecanuts. Matured Arecanuts are categorized in to Hasa, Bette, Mine and Gotu depending upon the quality of maturity level after harvesting and processing of Arecanut. The un-matured Arecanuts are not considered in arecanut market price analysis because in market no price for it. Gotu belongs to over-matured Arecanut, while Hasa is matured Arecanut.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

A. Related Work

Computer vision and image processing are rapidly becoming more common in precision agriculture. For fruit ripening classification, a variety of approaches based on shape, size, colors, and texture have been devised. Different ways for segmenting fruits to retrieve the features needed for the classification process were also available. Van Huy Pham and Byung Ryong Lee [6], Proposed a method for detecting deficiency in orange fruits was presented. Defects in orange fruits were detected using a segmentation technique. Graph-based and k-means clustering approaches were employed to classify defects in orange fruits.

X.E. Pantazi, et.al [7] proposed an automated approach for detecting leaf disease in numerous crop species by evaluating image features and deploying a classifier. The feature extraction was done with Local Binary Pattern (LBP), and the illness leaf classification was conducted with a support vector machine (SVM) classifier. Yanan Li, Zhiguo Cao, et.al [8], presented a technique for detecting in-field cotton using image segmentation. Unsupervised region creation and supervised semantic labeling prediction were utilized in conjunction with region-based segmentation. Each region was segmented using histogram-based color and texture features.

NST Saia, Ravindra Patil, et.al [9], developed a content-based image retrieval system which used two distinct feature vector methods. For gray scale, RGB, and YCbCr color images, the SVD (singular value decomposition) feature of increasingly truncated DCT (Discrete cosine transform) images and the DWT (Discrete wavelet transform) decomposed image were computed. The relevant images from the dataset were extracted based on color feature. Yaqoob Majeed, Jing Zhang, et.al [10], proposed a segmentation method to segment Apple trunk and branch. The images were stored in cloud from Kinect V2 sensor and using deep learning-based semantic method the images are segmented. To remove the background trees Depth and RGB features are extracted from cloud data.

By using the fundamental K nearest neighbour (KNN) model, Daneshwari Ashok Noola et.al. [11] concentrated on designing and developing the enhanced-K nearest neighbour (EKNN) model. EKNN is used to differentiate between illness classes. High-quality fine and coarse features are produced to gather discriminative, boundary, pattern, and structurally linked information, which is then utilized in the classification process. Gradient-based characteristics of excellent quality are provided by the classification procedure. Uoc Quang Ngo et.al. [12], devised a technique for precisely calculating the leaf area of cucumber plants using digital image processing. The suggested ways extract the cucumber plant's skeleton from RGB images and correctly estimate the leaf area of cucumber plants.

Megha.P.Arakeria, Lakshmana [13], To assess tomato quality, a computer vision method with 2 phases was developed. The hardware was created in the initial phase to collect tomato images and transport tomato fruit to suitable containers. The software was developed in the second phase to detect defects and ripeness in tomato fruit using image processing techniques. The RGB color model was used to determine ripeness. Yang Yu and Sergio A. Velastin, et.al [14], proposed a quick and efficient approach for achieving automated apple grading . For each apple sample, four images were taken (top, bottom and two sides). To differentiate the apple deformities, stem, and calyx, the grey value of each apple was acquired. K-means clustering was utilized to find the defective zone in an apple. Santi Kumari Behera, et.al [15], applied KNN, SVM and Naive Bayes classifier to classify the papaya fruits based on maturity status. Features such as LBP, Histogram of Oriented Gradients (HOG) and Gray Level Co-occurrence Matrix (GLCM) were extracted from papaya fruit image to fed as input to classifier.

Suresha M, et.al [16], presented a method to classify the diseased arecanut using texture features. The LBP , Haar Wavelets, GLCM and Gabor features of texture were used to determine the diseased arecanut. The HSI(Hue Saturation Intensity) and YCbCr color models were used. The LBP, Haar Wavelets, GLCM and Gabor methods were applied on HSI and YCbCr color model to extract the texture features. The KNN(K-nearest neighbor) classifier was used to classify the diseased and un-diseased arecanut.

Ajit Danti, et.al [17], presented a method for separating raw arecanuts into two categories. Red and green colors were used to determine upper and lower limits for the classification of raw arecanuts. Dhanesha R., et.al [18], introduced a brand-new method for segmenting arecanut bunches using the active contour method. The segmentation methodology was evaluated using the segmentation performance techniques VOE and DSC. Umesha D.K., et.al [19], proposed a study of different color models to segment the arecanut bunches. In order to separate the arecanut bunch from an input image, the HSV, YCbCr, YUV, YCgCr and YPbPr color models are applied to arecanut bunch images with manual threshold. Color models HSV and YCgCr were efficient in segmenting arecanut bunches from other color models used in the study.

According to the literature, there have been few studies conducted on arecanut bunches before harvest. A maturation level-based segmentation of arecanut bunches is needed in order to categorize and segment arecanut bunches. To conclude, the proposed research involves classifying arecanut bunches prior to harvesting. Arecanut bunches must be segmented from the input image in order to retrieve features. A well-segmented dataset yields a well-classified output. So, under the same conditions, a comparison of saturation(S), Cr, and Pr channels of HSV, YCbCr, and YPbPr color models (one of the most often used). This purpose aims to identify the best color model for segmenting arecanut bunches in a real-world image of an arecanut bunch. To identify the arecanut bunches according to maturity stage, the most efficient channel of the color model was used.

B. Color Models

The color model explains how small sets of primary colors can be combined to form an entire range of colors. In RGB space, the chrominance and brightness components are intermingled, making it impossible to distinguish a single color So there is a

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

need for transformation of image in RGB color space into other color space. The HSV, YCbCr and YPbPr color models has a good capability of representing properties of human vision. The problem of brightness is solved because the brightness is independent from color component. To segment arecanut bunch from the input image, the S, Cr, and Pr channels of the HSV, YCbCr, and YPbPr color models were utilized. The brief discussion of HSV, YCbCr and YPbPr color model is given [20].

HSV color model

Three elements make up this color model: hue (depth of color), saturation (purity of color), and value of intensity (color-brightness). Hue uses the terminology green, red, or magenta to denote pure color. To convert an image to HSV, the Red, Green, and Blue channels must be taken. In proposed work saturation channel were used to segment arecanut bunch from an input image. Saturation is the colorfulness of a stimulus relative to its own brightness [21].

YCbCr color model

The YCbCr model utilizes the characteristics of the human eye to mimic human vision [22]. A YCbCr model is another name for a YPbPr color model. The color model is the same as YCbCr. The value of Y ranges from 0 to 255. The value of Cb and Cr ranges from 16 to 240.

YPbPr color model

In addition to luminance (Y), YPbPr also includes chrominance (Pb/Pr). Blue in the RGB model is subtracted from the luminance channel to form Pb, which is represented by (Blue -- Y). In a RGB color model, (Red --Y) represents the Red color subtracted from luminance. YPbPr is also called as the analog version of YCbCr color model. The value of Y ranges from 0 to 1. The value of Pb and Pr ranges from -0.5 to 0.5.

C. Segmentation Performance Metrics

Segmentation results are normally verified with ground truth images. Measurement of segmentation accuracy is achieved using different metrics in this method. It is common to use Dice Similarity Coefficients (DSCs) and Volumetric Overlap Errors (VOEs) for statistical analysis. Using the provided method, S1 may be partitioned into regions based on G1, which can be used as the ground truth image. Evaluation of DSC and VOE are done with respect their definitions [19]. VOE is computed by applying given Equation.

$$VOE = ((|S1 \cap G1|/|S1 \cup G1|) - 1). \quad (1)$$

DSC [23] is obtains by using Equation.

$$DSC = 2 * |S1 \cap G1|/(|S1| + |G1|) \quad (2)$$

D. Feature Extraction for classification

A wide variety of classification processes employ color and texture feature analyses. The maturity levels of arecanut bunch were classified using color and texture features. To classify the matured and un-matured arecanut bunches the three color features (i.e., the Average intensity on the R, G, and B bands: RAvg, GAvg, and BAvg) were calculated by applying the Equations.

RAvg = sum of (red pixels) / the entire amount of pixels occupied by Arecanut Bunch Image (3)

GAvg = sum of (green pixels) / the entire amount of pixels occupied by Arecanut Bunch Image (4)

BAvg = sum of (blue pixels) / the entire amount of pixels occupied by Arecanut Bunch Image (5)

Texture Features : To classify the arecanut Bunch along with color features entropy and correlation features of texture were used.

E. Classifier

Random Forest Classifier: There are a variety of classification problems that can be solved using the Random Forest classification method. There are different kinds of Random Forests, but most are based on tree topologies. The forest trees each vote for the class label that is most likely based on their inputs. An ensemble of nonlinear patterns can be detected using this quick and noise-resistant approach. Both numerical and category data may be handled with ease. The advantage of Random Forest does not cause overfitting even as the forest gets bigger is one of its greatest benefits [24]. Weka 3.8.4 was used to build the Random Forest Classifier in this study. Weka provides a unified interface to a variety of learning algorithms, as well as pre- and post-processing methods and ways for analyzing the results of learning schemes on any given dataset.

F. Classification Performance metrics

The confusion matrix is used to evaluate classification performance based on elements from the matrix. To compare class labels, the terms TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) are employed. The precision and accuracy of the classifier were estimated based on the values acquired for the confusion matrix in the proposed study to evaluate its performance [25].

- True Positive: These are cases in which predicted as matured (Yes), and they were actual Matured.

  TP Rate = TP / Actual Yes (6)

- True Negative Rate: These are cases in which predicted as un-matured (No), and they were actually un-matured.

  TN Rate = TN / Actual No (7)

- False positive Rate: Predicted matured, but they actually un-matured. (Also known as a "Type I error").

  FP Rate = FP / Actual No (8)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

- False positive Rate: Predicted un-matured, but they actually matured. (Also known as a "Type II error.").

  FN Rate = FN / Actual Yes (9)

- Precision Rate: When it predicts matured, how often is it correct?

  Precision Rate = TP/ Predicted Yes (10)

- Accuracy: The most typical indicator to assess how well the categorization process is working. The ratio of properly identified samples to the total number of samples is used to determine accuracy.

Accuracy = (TP + TN) / (TP + TN + FP + FN) (11)

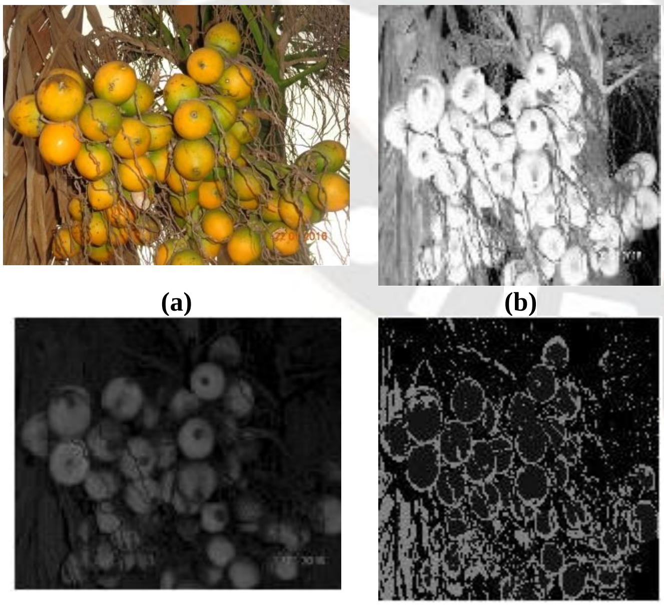
II. METHODOLOGY

This section describes the Database used for experimentation. The segmentation techniques to segment arecanut bunches using different color models were discussed. And also the tool and the classifier used to classify the arecanut bunch based on maturity level were mentioned.

A. Database

The Database contains 1017 colored images which are captured with help of a smart phone fitted on selfie stick. These images were taken at different time of the day. This database consists of 388 immature and 629 matured arecanut bunch and their ground truth images [26].

B. Segmentation and feature extraction of Arecanut bunch



(a) (b) (c) (d)

Figure 4**:** (a) (RGB) input image, (b) S-channel image of HSV , (c) Cr-channel image of YCbCr, (d) Pr-channel image of YPbPr color model.

Features were collected from the region of interest to classify arecanut bunches depending on maturity level. As a result, segmentation is critical for locating the desired region. Matlab 14.b is utilized in the proposed study to conduct segmentation and feature extraction experiments. In order to process an image further, segmentation is necessary to extract useful features.

An arecanut bunch's RGB image is read to segment it, and it is scaled to 255 x 255. The scaled image is transform to HSV, YCbCr and YPbPr image. The saturation, Cr, and Pr channel images of the HSV, YCbCr, and YPbPr color models are shown in Figure 4. The S, Cr, and Pr channels were exploited for segmentation. Because brightness is independent of color component, the brightness problem is overcome.

The image is then transformed to a binary image using otsu's approach of automated thresholding [27]. The resultant binary picture is jam-packed with noise and extraneous information. To remove noise the median filter is used, while morphological procedures are used to remove superfluous information and fill gaps in an image. A series of mathematical procedures known as morphological operations is used to transform an input image into an identically sized output image. While doing so, the structural element is taken into consideration. Erosion and closing are the two most important morphological processes. Pixels from object edges are removed using the erosion process. Discs are used as structural components in erosion. As soon as erosion has been applied, a closure operation is carried out. To fill the gaps in an image Closing method are applied with structural element as a disc. The closure method is divided into two stages: dilution and erosion. After that, the picture was transformed back to its RGB format. RGB components were recovered independently from the RGB segmented image to compute the average of each component. To extract correlation and entropy texture features, an RGB segmented image is transformed to a gray scale image. The computed feature values were saved in a.csv file.

C. Cclassification

The .csv file is used as an input by the Weka 3.8.4 utility. The arecanut bunches are classified depending on their maturation stage using a random forest classifier. The values of the confusion matrix were used to calculate precision and accuracy metrics to evaluate the performance of classification.

III. EXPERIMENTATION

In order to categorize the maturity level of 1017 images of immature and mature arecanut bunches, experiments were conducted. To classify them, certain features were required. To extract key features, arecanut bunches are segmented from the input image. All 1017 images are converted to HSV, YCbCr, and YPbPr color models to choose the S-channel of HSV, Cr-channel of YCbCr, and Pr-channel of the YPbPr color model. The optimal color model channel for segmenting arecanut bunches is identified by comparison of the segmented and ground truth images.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Figure 5 (a) and (b), shows input image and appropriate ground truth images of a immature arecanut bunch image. Figure 5 (c) (d),(e) (f) and (g) (h) shows segment mask and overlay of result images obtained using S-channel, Cr-channel and Pr-channel of HSV, YCbCr and YPbPr color model respectively. Figure 5, demonstrates that when the S-channel of HSV was used instead of the Cb and Pr channels of the YCbCr and YPbPr color models, the resultant image of an immature arecanut bunch was the closest to the ground truth (Figure5(b)) image.

From Figure 7, it is observed that resultant image was close to ground truth(Figure 7(b)) image of mature arecanut bunch using S, and Cr channel of HSV and YCbCr color model. Figure 7, also shows using Pr channel of YPbPr color model resultant image not accurately segmented compare to ground truth image.

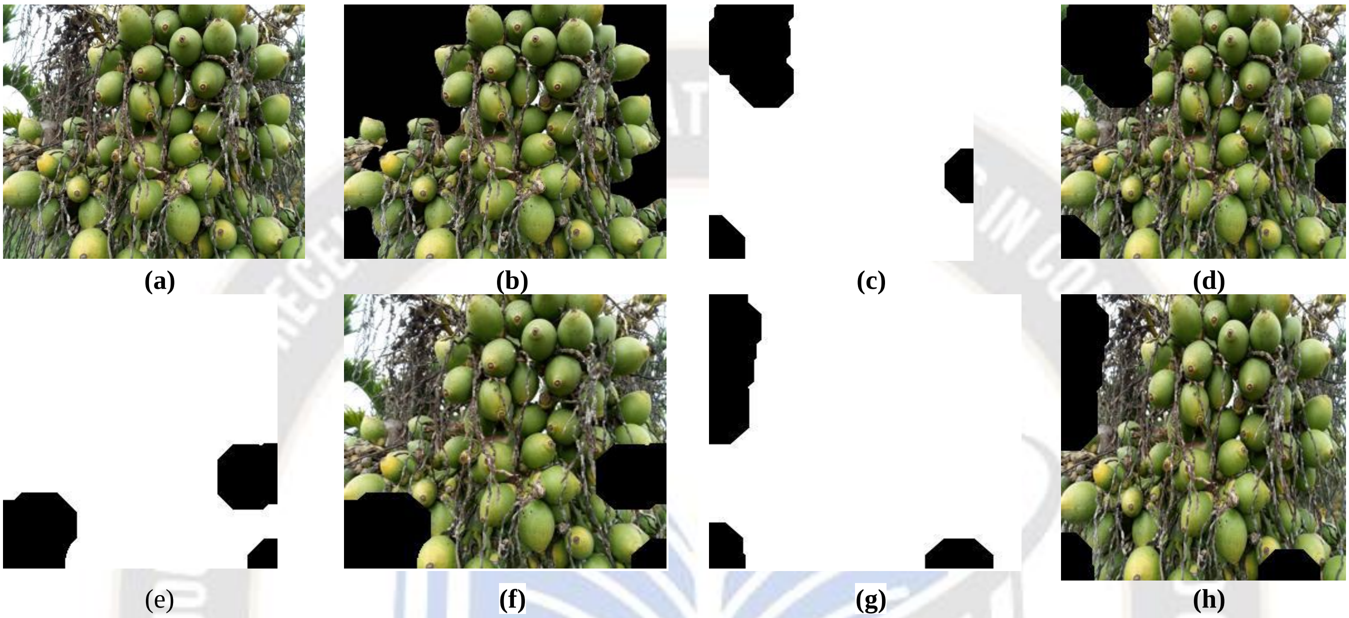


Figure 5**:** Results of different color model used to segment immature arecanut bunch image (a) input image, (b) ground truth image, (c) & (d) represents segmented mask & overlay of result using HSV S-channel , (e) & (f) using YCbCr Cr-channel, (g) & (h) using YPbPr Pr-channel respectively.

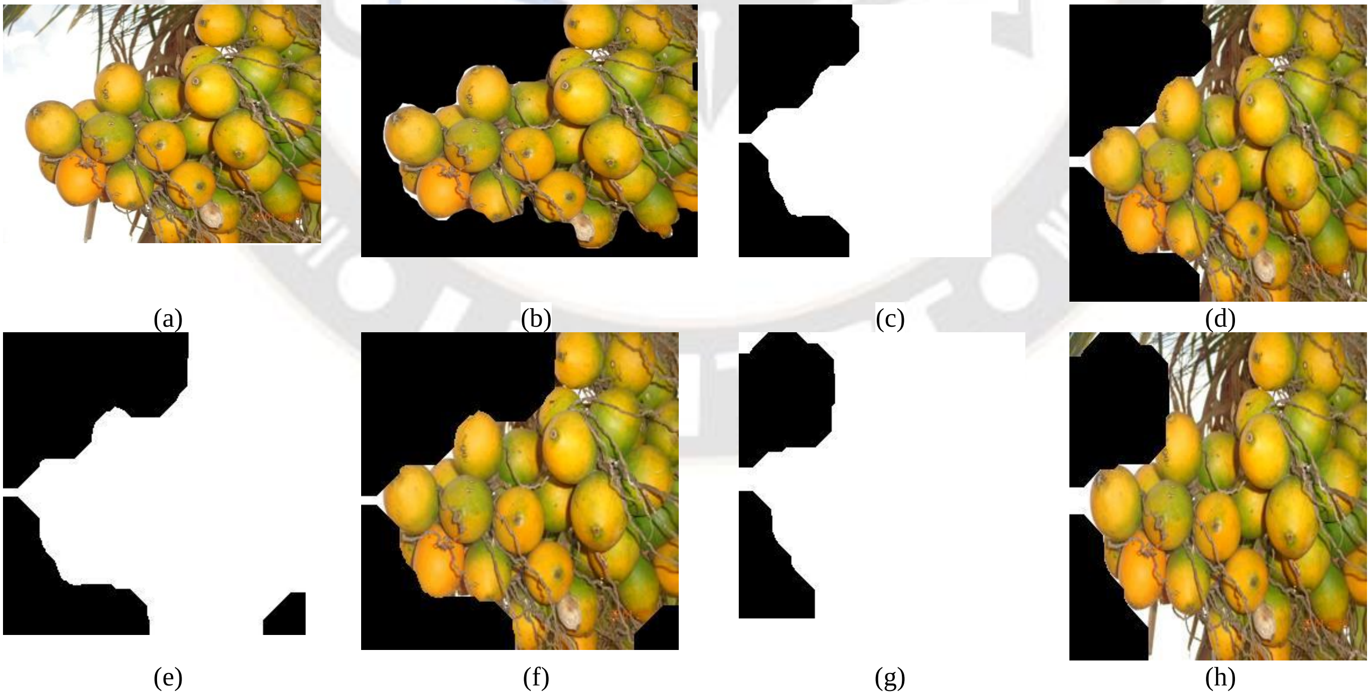
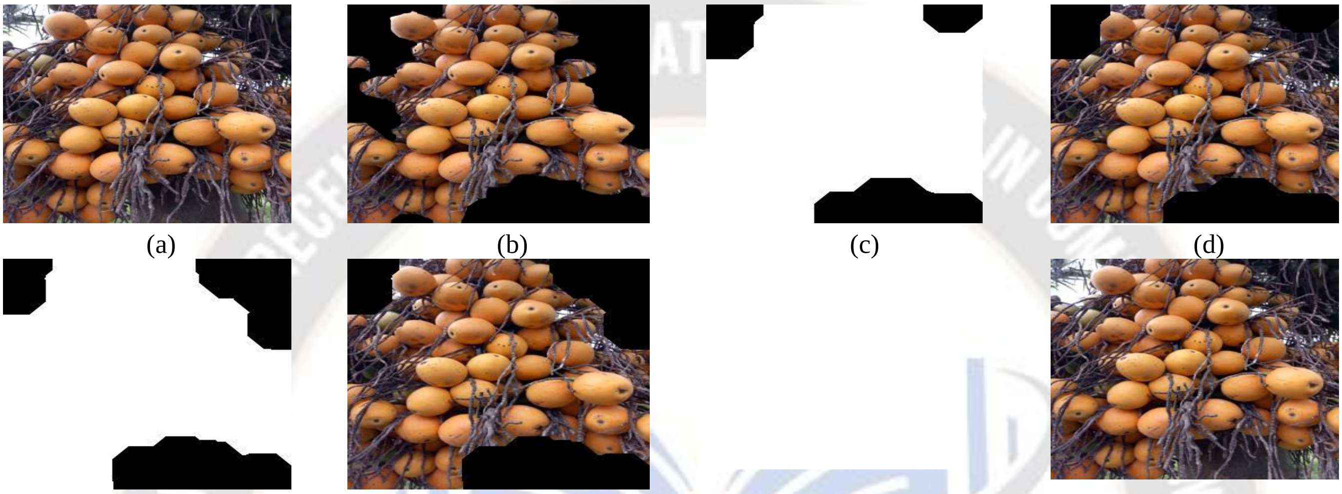


Figure 6: Results of different color model used to segment matured arecanut bunch image (a) input image, (b) ground truth image, (c) & (d) represents segmented mask & overlay of result using HSV S-channel , (e) & (f) using YCbCr Cr-channel , (g) & (h) using YPbPr Pr-channel respectively.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

By observing Figure 6(d) S channel of HSV and Figure 6(f) Cr channel of YCbCr color model segmented well when it is compared to ground truth image.(Figure 6(b)). Pr-channel of YPbPr color model is comparatively far segmented by noticing ground truth image. When compared to the ground truth image in Figure 6(d) S channel of HSV and Figure 6(f) Cr channel of YCbCr color models segmented effectively. By observing the ground truth image, the Pr-channel of the YPbPr color model is comparably widely segmented. Figure 5, Figure 6 and Figure 7, illustrates that both immature and mature arecanut bunch images were correctly segmented using the S-channel of the HSV color model. In contrast, the Cr-channel of the YCbCr color model is better at segmenting matured arecanut bunches than immature arecanut bunches, since the Cr-channel of the YCbCr color model is strong in the regions where reddish color appears [22]. YPbPr is the analog version of YCbCr color model, is also gives good results when reddish content is in input image [28].

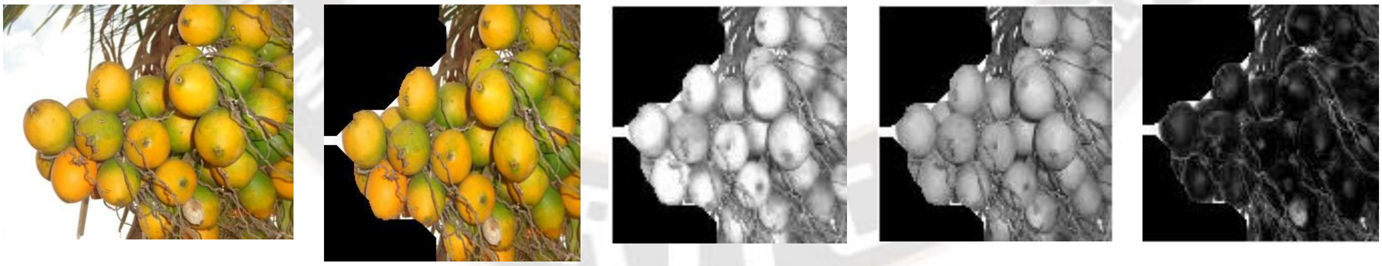


(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 7: Results of different color model used to segment over matured arecanut bunch image (a) input image, (b) ground truth image, (c) & (d) represents segmented mask & overlay of result using HSV S-channel , (e) & (f) using YCbCr Cr-channel , (g) & (h) using YPbPr Pr-channel respectively.

By distinguishing Figure 5, Figure 6 and Figure 7, the Pr-channel of YPbPr color model is relatively good for segmenting matured compare to immature arecanut bunch images. As a result, the S-channel of the HSV color model was more effective in segmenting arecanut bunches from the input image than the Cr and Pr channels of the YCbCr and YPbPr color models, respectively.



(a) (b) (c) (d) (e)

Figure 8**:** Arecanut Bunch. (a) Input image (RGB) (b) Segmented image (c)Red Channel (d) Green Channel (e) Blue Channel

Color and texture features were taken from segmented images to execute the classification procedure. To extract the RGB components of an image, a segmented image is transformed to an RGB image. Figure 8 (a), (b), (c), (d), and (e) shows the input image, segmented image, the red component image excluding green and blue colors, the green component image excluding red and blue colors and the blue component image excluding green and red colors respectively. Texture features like correlation and entropy were extracted using GLCM method of MATLAB 14b.

The retrieved feature values were entered into the Weka 3.8.4 tool (.csv file). The experiment was carried out using a random forest classifier. The parameters set in random forest classifier are as follows:

- **Cross-Validation Fold:** is a method of the Evaluation class that is employed to carry out cross-validation using a single dataset and an untrained

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

classifier. Cross-Validation Fold is given a value of 10 (default value).

- **BagSizePercent:** Size of each bag, as a percentage of the training set size. Value assigned is 100.
- **BatchSize:** The preferred number of instances to process if batch prediction is being performed. Value assigned is 100.
- **NumIterations:** The number of trees in the random forest. Value assigned is 100.

Other parameters were set to default value assigned by Weka 3.8.4. The findings were discussed in the section Results and Discussion.

IV. RESULTS AND DISCUSSIONS.

Arecanut bunches are first segmented to extract required features before being classified according to maturity level. With dataset images provided with the proposed method, the segmentation results are compared to manually graded ground truth (GT). The efficiency of segmentation was measured using DSC and VOE segmentation measures in this proposed work. table 1 & table 2 summarizes the results of the study .A successful segmentation produces values for DSC that are closest to 1 and VOE that are close to 0. A detailed examination of the values generated for each color model was examined in this section for each segmentation performance metrics.

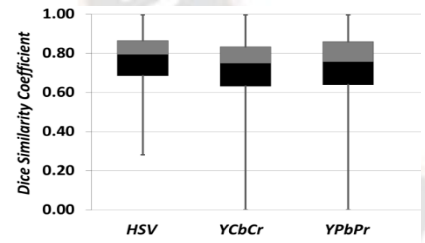
A. DSC Evaluation

Figure 9(a), shows box plot in which line inside the box indicates the median value and as a vital unwritten-rule. Any segmentation method, according to DSC, works well if the matching median in the box plot is high. The box's range represents the dispersion of DSC-values estimated from distinct arecanut bunch images in the dataset. Table 1, shows that the HSV color model's S-channel has the greatest median values when compared to the other color models utilized in this study.

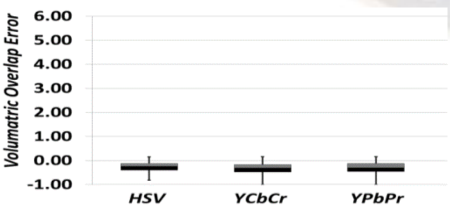
The S,Cr, and Pr channels of the HSV, YCbCr, and YPbPr color models have median values of 0.79,0.75, and 0.076, respectively. According to DSC, the S-channel of the HSV color model is effective for separating an arecanut bunch using an input image.

Table 1: Segmentation assessment using DSC metric

	Minimum	Q1	Median	Q3	Maximum	Hidden	Lower	Upper	Whiskers Top	Whiskers Bottom
HSV	0.28	0.69	0.79	0.87	1.0	0.69	0.11	0.07	0.13	0.40
YCbCr	0.0	0.63	0.75	0.83	1.0	0.63	0.12	0.09	0.16	0.63
YPbPr	0.0	0.64	0.76	0.86	1.0	0.64	0.12	0.10	0.14	0.64



(a)



(b)

Fig 9: Whisker Box plot for the DSC and VOE against color models w.r.t Table 1 and Table 2 respectively.

B. VOE Evaluation

The box plot used to analyze VOE is shown in Figure 9 (b). VOE produces good segmentation if the median value inside the box plot is closest to zero, as indicated in the Section 1. C. The S-channel of HSV encountered median values were closest to zero when compared to the Cr and Pr channels of the YCbCr and YPbPr color models median values, as seen in table 2. Examining the median values of the Cr channel of YCbCr (-0.40) and the Pr channel of YPbPr (-0.39) for VOE. The S-channel of the HSV color model has an acceptable VOE value for segmenting Arecanut bunch.

The S and Cr channels of the HSV and YCbCr color models performed well, as shown in table 1 & table 2 and Figure 9 (a) & Figure 9(b) for 1017 images. By looking at the output images of the S and Cr channels of the HSV and YCbCr color models for 1017 images. When comparing images of matured and immature arecanut bunches, it was discovered that the Cr-channel of the YCbCr color model output images was well segmented for matured arecanut bunches. For both immature and matured arecanut bunch images, the S-channel of the HSV color model produces good results.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

It is clear from the segmentation findings that arecanut bunch images were segmented well using the S channel of the HSV color model. The value of the retrieved features is saved in .csv file with six columns labelled RAvg, GAvg, BAvg, Correlation, Entropy, and Class. There are two values in the Class column. Yes or No. Yes denotes matured arecanut bunches, whereas No denotes immature arecanut bunches.

The performance of classification is assessed using precision and accuracy metrics based on confusion matrix values, as shown in table 3. Table 3, shows that 893 samples out of 1017 dataset samples were correctly classified, resulting in a **87.81%** of accuracy. In the dataset, there were 388 matured arecanut bunches, of which 311 (or approximately **80.15 %**) were accurately classified as matured arecanut bunches. The dataset had a total of 629 immature arecanut bunches, of which 582, or **92.52%**, were properly classified as immature arecanut bunches.

Table 2: Segmentation assessment using VOE metric

	Minimum	Q1	Median	Q3	Maximum	Hidden	Lower	Upper	Whiskers Top	Whiskers Bottom
HSV	-0.84	-0.48	-034	-0.24	-0.01	-0.48	0.14	0.10	0.23	0.36
YCbCr	-1.0	-0.54	-0.40	-0.29	-0.01	-0.54	0.14	0.12	0.28	0.46
YPbPr	-1.0	-0.53	-0.39	-0.25	-0.01	-0.53	0.14	0.15	0.24	0.47

Table 3: Performance of classifier evaluation using Confusion Matrix

N=1017	**Predicted NO**	**Predicted YES**	**Correctly Classified**	**Precision**	**Accuracy(%)**
Actual YES(388)	FN=77	TP=311	311	0.868	**87.8%**
Actual No(629)	TN=582	FP=47	582		
Total(1017)	659	358	893		

Farmer failed to identify the matured arecanut bunches for harvesting than there is small amount of loss in profit, because arecanut bunches moves to next level of maturity as discussed in Section 1. Farmers should avoid harvesting immature arecanut bunches since it is in the form of liquid content. If immature arecanut bunches are picked, there will be a significant loss of earnings.

This work shows **92.52%,** towards identifying immature arecanut bunches, which may aid farmers in harvesting only matured arecanut bunches and increasing profitability. When images are acquired in poor light and contain a significant number of waste inflorescence's (male flowers), the proposed method may fail to properly differentiate arecanut bunches using the color models used in this investigation. It is difficult for the classifier to determine the maturity level of arecanut bunches since the images were not segmented appropriately.

V. Conclusion

Automatized system for segmenting and classifying the arecanut bunch as matured and immature was developed in this paper. Segmentation is more dominant but arduous process to collect vast amount of information. The S, Cr and Pr channel of HSV, YCbCr and YPbPr color models were used to identify the suitable color model image which is superior for segmenting an arecanut bunch from an input image. The simplest and fastest Otsu's thresholding method is used to get segment mask. To compute performance of segmentation DSC and VOE segmentation metrics were applied. The experimental results reveal that the S-channel of the HSV color model is superior for segmenting arecanut bunches from the input image. Three color features (i.e., RAvg, GAvg and BAvg) and two texture features (Correlation and Entropy) were chosen for classification. Random forest classifier is used to classify arecanut bunches based on maturity level. Results shows the classifier classifies arecanut bunches with an accuracy of 87.81%.

References

1. India Department of Agriculture. ANNUAL REPORT 2017-18, 2019.
2. Shivamogga UAHS. Areca referral laboratory at UAHS, Shivamogga, 2019.
3. K.S. Sheshagiri, H.Narayanaswamy et.al.. Methods of arecanut cultivation.
4. D. Surya Prabha and J. Satheesh Kumar. Performance evaluation of image segmentation using objective methods. Indian Journal of Science and Technology, 9(8), 2016.
5. MAMCOS. Arecanut Market Analysis. {https://mamcos.info/en/mamcos-areca-rates/} Date of Access:27/12/2020.

6. Van Huy Pham and Byung Ryong Lee. An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm. Computer Science, 2:25–33, 2015.
7. A.A. Tamouridou X.E. Pantazi, D. Moshou. Automated leaf disease detection in different crop speciesthrough image features analysis and one class classifiers. Elsevier Journal of Computers and Electronicsin Agriculture, 156:96–104, 2019.
8. Yanan Li, Zhiguo Cao et.al. In field cotton detection via region-based semantic image segmentation. Elsevier Journal of Computers and Electronics in Agriculture, 127:475–486, 2016.
9. NST Saia, Ravindra Patil et.al. Truncated dct and decomposed dwt svd features for image retrieval. In Proc. Elsevier Conf. 7th International Conference on Communication, Computing and Virtualization, volume 79, page 579 – 588, 2016.
10. Yaqoob Majeed, Jing Zhang et.al. Apple tree trunk and branch segmentation for automatic trellis training using convolutional neural network based semantic segmentation. In Proc. Elsevier Conf. IFAC(International Federation of Automatic Control) PapersOnLine, pages 51–17, 2018.
11. Daneshwari Ashok Noola et.al., “Corn leaf image classification based on machine learning techniques for accurate leaf disease detection”, Indonesian Journal of Electrical Engineering and Computer Science, Vol. 12, No. 3, June 2022, pp. 2509-2516, ISSN: 2088-8708, DOI: http://doi.org/10.11591/ijece.v12i3.
12. Uoc Quang Ngo et.al., “Digital image processing methods for estimating leaf area of cucumber plants”, Indonesian Journal of Electrical Engineering and Computer Science, Vol 25, No. 1, January 2022, pp. 317-328, ISSN: 2502-4752, DOI: http://doi.org/10.11591/ijeecs.v25.i1.
13. Lakshmana Megha.P.Arakeria. Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry. In Proc. Elsevier Conf. 7th International Conference on Communication, Computing and Virtualization, volume 79, page 426 – 433, 2016.
14. Yang Yu and et.al Sergio A. Velastin. Automatic grading of apples based on multi-features and weighted k-means clustering algorithm. Information Processing in Agriculture, 2019.
15. Santi Kumari Behera et.al. Maturity status classification of papaya fruits based on machine learning and transfer learning approach. Information Processing in Agriculture, 2020.
16. S. K Narasimhamurthy Suresha M, Ajit Danti. Classification of diseased arecanut based on texture features. Computer Applications (0975 – 8887) Recent Advances in Information Technology, 2014.
17. Suresha M Ajit Danti. Segmentation and classification of raw arecanuts based on threesigma control limits. Elsevier Journal of C3IT-2012, 4:215 – 219, 2012.
18. Shrinivasa Naika C. L. Dhanesha R. A novel approach to segment an arecanut bunches using active contour method. In Proc. Springer,2nd International conference on Integrated Intelligent Computing, Communication And Security, pages 677–682, 2019.
19. R. Dhanesha, D. K. Umesha, C. L. Shrinivasa Naika and G. N. Girish, "Segmentation of Arecanut Bunches: a Comparative Study of Different Color Models," 2021 IEEE Mysore Sub Section International Conference (MysuruCon), 2021, pp. 752-758, doi: 10.1109/MysuruCon52639.2021.9641680.
20. Jose M. Chaves-Gonz alez, et.al. Detecting skin in face recognition systems: A color spaces study. Elsevier Journal of Digital Signal Processing, 20:806–823, 2010.
21. M.D. Fairchild. Color appearance models. Reading, 1998.
22. P. Ganesan et.al . Satellite image segmentation based on ycbcr color space. Ecology, Vol 8(1)(3):35–41, 2015.
23. Lee R Dice. Measures of the amount of ecologic association between species. Ecology, 26(3):297–302, 1945.
24. Breiman L. Random forests. Mach Learning, pages 5–32, 2001.
25. Abinash Tripathy et.al.. Classification of sentiment reviews using n-gram machine learning approach. Expert Systems With Applications, pages 117–126, 2016.
26. Shrinivasa Naika C. L. Dhanesha R. Segmentation of arecanut bunches using ycgcr color model. In 1st IEEE International conference on Advances in Infromation Technology, pages 50–53, 2019.
27. N. Otsu. A threshold selection method from gray-level histograms. IEEE Transactions on Systems Man Cybernet, SMC-9(1):62–66, 1979.
28. Garrett M. et.al. Derivation of a color space for image color difference measurement. COLOR research and application, 35(6):387–400, 2010.