

Identification of Mustard Greens Freshness Level Based on RGB Leaf Color and Stem Shape Features using Image Thinning Morphology

Rosa Andrie Asmara*, Budi Harijanto, Mustika Mentari, Ekojono
and Afwika Chori Q

Information Technology Department
State Polytechnic of Malang, East Java, Indonesia
*rosa.andrie@polinema.ac.id

Abstract

*Fruit and vegetables are the commodity with high commercial value and have many benefits. One of the popular and favorable vegetables in Indonesia is Mustard greens (*Brassica juncea*) since it has good taste and easily gets at an affordable price. Indonesian and Chinese food often used this vegetable as soup or stir fry. Like other leaf vegetables, Mustard greens can easily damage. It can easily wither since it contains much water. Wide and large leaf surface make withering and evaporation process faster. Keep the freshness level in vegetable storage is one of the important criteria in the market to maintain high-quality products. To sort out a high volume of products, stores need an automatic, fast, and efficient Mustard Greens system for sorting out their freshness level. The improvement of digital image processing research and technology makes it possible to identify agricultural and plantation products. Combining this technology with some mechatronics automation, the sort out system could work automatically. The purpose of this research is to identify Mustard Greens Freshness by visual analysis of their images acquired using Smartphone camera. The feature extracted from the image is the shape of the stem and their leaf color. The feature extraction method for stem shape is done using Binary Morphology. First Order statistical method from the image histogram used for leaf color extraction is Mode. Single Layer Perceptron used as a Neural Network Classifier. The classification will determine best weight value in the network for 3 categories of freshness: fresh, slightly withered and very withered. From the experiments result, Classification accuracy for identifying the freshness of the Mustard greens using Single Layer Perceptron is 90%. Conducted classification accuracy results indicate that the methods good enough for classify freshness level of Mustard Greens using color and stem shape feature.*

Keywords: Image processing, Identification, Feature Extraction, Normalization, Neural Network

1. Introduction

The harvest and post-harvest distribution on Mustard greens starting from cut the base of the stem, leaching, disposal, sorting, preparation and distribution of Mustard greens, farmer should give a splash of water frequently to make withering process slower. Unlike other leafy vegetables, Mustard greens can easily damage and withered since it contains much water. Wide and large leaf surface make withering and evaporation process is done faster. Freshness in the vegetable is one of the criteria in selecting high-quality products. The age of the freshness of the mustard ranged between 10 – 14 days stored in cold rooms

Received (February 18, 2018), Review Result (May 20, 2018), Accepted (May 28, 2018)

* Corresponding Author

with temperature 0° – 5° degree Celsius and humidity of 70% to 90% (cold storage), while in the outdoors (room temperature) ranged of freshness is between 3 – 5 days [1].

The identification of freshness on Mustard greens can be done using visual and physical observations such as checking the color of fresh leaf, no yellowing or Browning, leaf veins can be seen, there are no patches of intact leaf, no hollow and the stem leaf can easily broke (not resilient). Physical observation, however, will be destructive and cause mustard greens damage. The manual freshness identification process is subjective so that under certain conditions it is not specific to the classification process or depends on the human perception of the color composition of the object (green mustard). But the sense of a human eye has weakness such as a level of fatigue may be occurred during the identification process when the process takes too long.

A previous study of Mustard greens image segmentation has been performed as preprocessing step in extracting color features. The features are extracted using a development method of the excess green color index to obtain the region from the background using a variety of color spaces. Excess Green value then normalized and convert to Hue color space. Hue color space values are extracted due to its low sensitivity against background noise and lighting conditions. Otsu threshold, morphology Opening, Labeling, RLE, bounding box, transformation geometry and produce 99.16% classification accuracy [2]. Research on green mustard objects has been done for identification of nitrogen and potassium in mustard leaf using color moment for color feature extraction, gray level co-occurrence matrix (GLCM) for textural feature extraction and artificial neural network back propagation for the classification, the result of accuracy reaches 97,82% [3].

Therefore, this research proposes easier application development, at a low-cost computation, and achieved high classification accuracy.

The rest of this paper is organized as follow. Section 2 presents theoretical background. Section 3 presents implementation and obtained results. Finally, Section 4 concludes this work.

2. Rudimentary

In this research, digital image processing technology used to preprocess the image dataset and classifying the freshness of the Mustard Greens using RGB color feature extraction based on the image from mustard leaf with the First-Order Statistical Method used Histogram. Features extracted from the stem shape using the Thinning Morphological Binary Method. Classification of Mustard Greens freshness level is done using Artificial Neural Network method on Single Perceptron learning method.

2.1. The Freshness Parameters

Mustard greens (*Brassica juncea*) is leafy vegetable used as food which is easily broken (perishable). In fact, the vegetables just after their harvest time are living tissue contains a lot of water and begin the respiration and transpiration process. The presence of high respiration cause vegetables easily withers and microbes also easily contaminate it. The high room temperature will stimulate microbial breeding. A low room temperature will inhibit the growth of microbes. However, A high humidity will prevent withering and drought on the vegetables [1]. To easily sorting out mustard greens between fresh or withered, this research created an output parameter based on the freshness level in three category, namely: fresh, slightly withered and very withered. Figure 1 shows different category of freshness by visual observation. The difference can be clearly seen by its color and stem shapes.

 <p>(a) Fresh</p>	 <p>(b) Slightly Withered</p>	 <p>(c) Very Withered</p>
<ul style="list-style-type: none"> • Fresh colored leaf (Dark green color, and sharp) • No yellowing or Browning • Veins on the leaf can be seen clearly • Whole leaf no spotting, not perforated • Stem mustard no wrinkles, and easily broken 	<ul style="list-style-type: none"> • Dull color of the leaf • Light green with some minor yellow color • Yellowing on the stem • Dull leaf veins • Mustard stem not straight • Mustard Stem leaning 	<ul style="list-style-type: none"> • Pale color of the leaf • The dominant color is yellow and Brown style • Stained leaf • Leaf Veins Volume shrinking • Mustard stem not straight (ducking) • Mustard stem difficult to break • Leaf veins is not visible or is very faint

Figure 1. Sample of the Parameters Freshness for Mustard Greens can be Seen by its Color and Stem Shape from Human Visual Perception

2.2. Image Segmentation

Image Segmentation is the process of the digital image is partitioned into multiple segments (sets of pixels). Image segmentation in this research will separate image into two-segment, background and foreground image. This research using Global Threshold to segment image and separate into two-segment. All the pixels in the image are converted into black and white with one thresholding value [4].

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq T \\ 0 & \text{if } f(x, y) < T \end{cases} \quad (1)$$

$g(x, y)$ represent binary image from grayscale image $f(x, y)$ and T (threshold) declared a threshold value limit [5]. From the experiments, the best threshold value used in this research is 241. The next regional segmentation produces white (255-value) that is used as a container to map pixel coordinate of RGB image results from the image segmentation step.

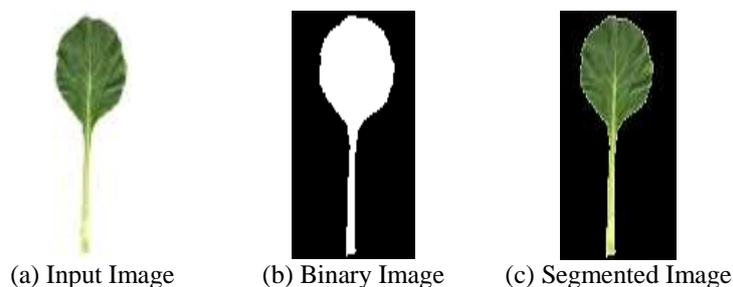


Figure 1. The Segmentation Process

Figure 2 show the image segmentation process, (a) is an Original image, (b) is a binary image after segmentation process, (c) is a foreground image successfully segmented from background image.

2.3. Color Features Extraction of Leaf

On the color feature extraction, this research proposed to extract the characteristics or information of objects in the image and use it as a feature to distinguish between class. The color information extracted are the value of RGB (red, green, blue) for each pixel in one image. The RGB pixel value then reduced the dimension using Mode First Order Statistical method. Statistical feature extracted from image dataset RGB histogram value. These values depends on the pixels individually and not a result of the interaction of pixel values that adjacent to others [6]. From the values of the histogram produced, Mode can be easily acquired. Mode represents the highest frequency data in one image.

2.4. The Shape features Extraction of Stem

Shape feature extraction used in retrieval characteristics of an object that can represents the characteristics of the object [7]. The shape of an object is the character of the surface configuration represented by lines and contours [7]. Shape Feature extraction in this research are used to measure stem curvature shape while smaller value represents fresher category. This extraction is done by calculating the curvature using Thinning Binary Morphology.

2.4.1. Binary Morphology

Binary morphology is a set of fundamental operations on binary images. To get a binary image, one simply converts grayscale to a binary image using threshold image method. It is a simple method of nonlinear convolution-like operation between two sets. Unlike linear convolution, morphology takes the minimum and maximum value of elements in the set. One set is the image pixel value the other is the kernel of the convolution, which is called a structural element and has a defined origin in the centre. Binary morphology is extremely important for fast and low-level image matching operations. Most of the Computer Vision system has binary morphology to pre-process image data because of its usefulness. Some of the basic binary morphological methods are dilation, erosion. Others binary morphological operations are the combination of those two methods such are opening, closing, thinning, thickening, skeleton, etc. this research proposed used Zhang-Suen Thinning morphological method to extract pixel wide from the shaped of Mustard Greens stem [8]. Thinning image is stalling or elongated shaped which can be expressed in a thinner version. The thinner form consisting of connected lines called the skeleton or core line. Ideally, the skeleton stretched along with the axis of the object. This algorithm is an iterative algorithm method for extracting the skeleton of a picture and removing all the contour points of the picture except those points that belong to the skeleton [8]. Thinning Zhang Suen has two sequence sub-iterations which are performed against the contour points or structural elements of the image. Contour point is any pixel with a value of 1 and has at least an 8-neighbor that has a value of 0. In order to preserve the connectivity of the skeleton, this method divides each iteration into two subiterations. In the first subiteration, the contour point P1 is deleted from the digital pattern if it satisfies the following conditions:

- a. $2 \leq N(p_1) \leq 6$;
- b. $S(p_1) = 1$;
- c. $p_2 \times p_4 \times p_6 = 0$;
- d. $p_4 \times p_6 \times p_8 = 0$;

Where $N(p_1)$ is the number of neighbors of the p_1 is not 0 (zero): that, $N(p_1) = p_2 + p_3 + \dots + p_8 + p_9$. $S(p_1)$ is the sum of the sequence of 0-1 transition at $p_2, p_3, \dots, p_8, p_9$, which is an 8-neighbor of P_1 as shown in Figure 3.

P_9	P_2	P_3
P_8	P_1	P_4
P_7	P_6	P_5

Figure 2. Points under Consideration and their Locations (8- neighbors)

And in the second step, the conditions (a) and (b) the same as the first step, while the conditions (c) and (d) is changed to:

$$(c') \quad p_2 \times p_4 \times p_8 = 0 ;$$

$$(d') \quad p_2 \times p_6 \times p_8 = 0 ;$$

The first step is performed on all pixels in the image border. If one of the four conditions above are not met or are violated then the corresponding pixel value is not modified. Conversely, if all the conditions are met then the pixels marked for deletion. Pixels that have been marked will not be removed before all border points processing are finished. This is useful to prevent changes to data structures. After step 1 completed for all border points then performed the deletion to the point marked (changed to 0). After that is done the second step in the data results from the same process of step one. Within a single iteration, the sequence of algorithms process are:

- a. Run the first step to mark the border points to be removed,
- b. Delete the marked points with replacing it with the number 0 (zero),
- c. Run the second step on the remaining border points on the first steps that have not been deleted then that comply with all the conditions that should be fulfilled in step 2 then marked for deletion,
- d. Delete the marked points and replace it with the number 0 (zero)

This procedure is done in an iterative process until there are no more points that can be removed, at the time of the algorithm is finished it will be generated from the initial image of the skeleton. Figure 4 below shows binary image before and after thinning process.

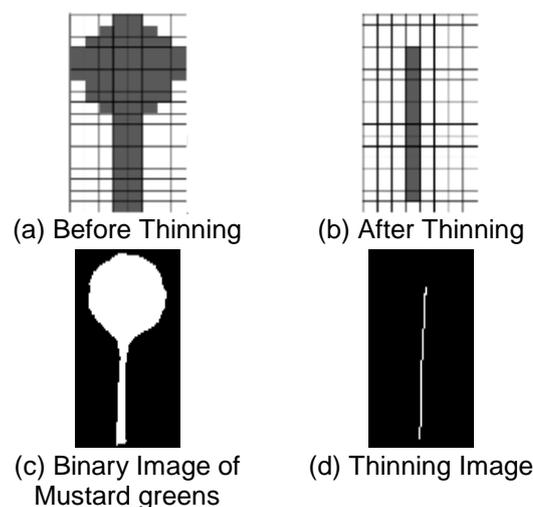


Figure 3. Thinning Process

2.5. Data Normalization

Normalization is the process of grouping data and determines the relationship between groups. Normalization is done to equalize the value of the results after RGB (Red, Green, and Blue) color extraction from the leaf and has a results value range 0 to 1. The normalization data in this research is using the value of the Max and Min of all data, as shown in the following formula (2):

$$Z = \frac{X - \min(X)}{\max X - \min X} \quad (2)$$

The Z for the variable data normalization result and X is the original data. The min (X) as the data for the minimum value and max (X) is the maximum value of the data, the normalization is done per column data.

2.6. Classification

The process of classification using NN (Neural Network) method with the Training and Testing approach on Single Layer Perceptron Algorithm. Figure 5 below shows Single Layer Perceptron Architecture. This architecture had 4 inputs from color (red, green, blue) feature, Thin, and Bias for additional input.

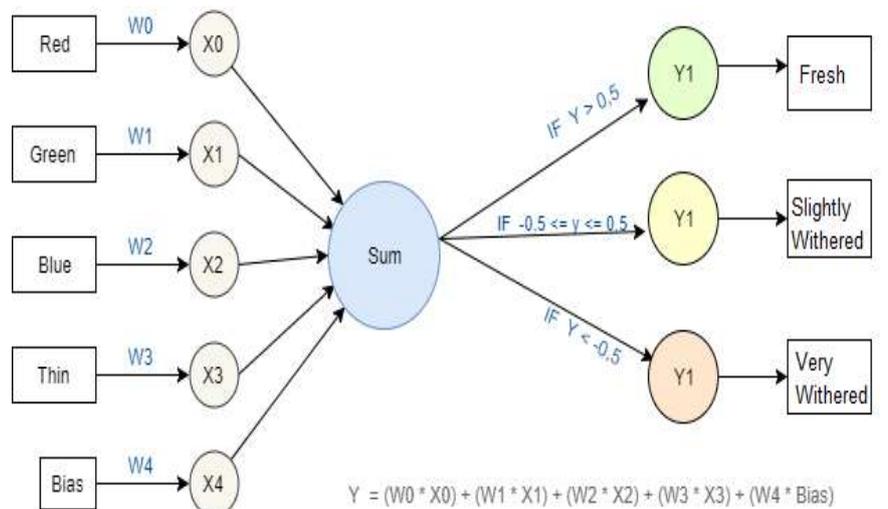


Figure 4. Single Layer Perceptron Architecture

2.6.1. Training

Learning or call training, is a process of giving pattern to the Neural Network (NN) which have known output label. The training will give initial value for each input weight, iterate to modify input weight and stop the iteration once the minimum error threshold achieved. The difference between the pattern of actual output with the desired pattern output is called error with the correct weights are used for NN, so NN can generate output as close as possible to the pattern of a known target [9]. The following steps are done to get the best weight:

- a. Initializing input the results extraction feature and rate of the initial weights

W_i = weight of input data (i)

X_i = input data (i)

Y_{net} = Result of activation function

Youtput = Output value

- b. Determine the rate of learning (Lr) or speed of learning with a value $0 < \alpha \leq 1$. This research gave initial value of $\alpha = 1$
- c. The activation Function = Saturating Linier (See Figure 6).

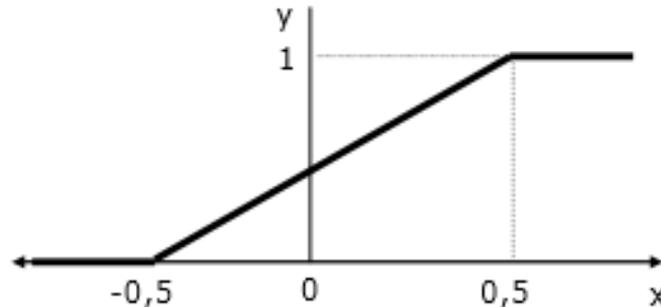


Figure 5. Saturating Linier Activation Function

$$Y_{net} = \begin{cases} 1, & Y_{net} \geq 0.5 \\ 0, & -0.5 \leq Y_{net} \leq 0.5 \\ -1, & Y_{net} \leq -0.5 \end{cases}$$

- d. Specify the target : (Fresh =1) , (Slightly Withered= 0), and (Very Withered = 0)
- e. Calculate the output of Activation Function:

$$f_{(net)} = b + X_i \times W_i$$

- b = Bias
 - X_i = Input
 - W_i = Weight of input
 - t = Target
- f. Modify the weighting pattern that contains mistake ($Y \neq t$), using following the equation:
 - $W_i(\text{new}) = W_i(\text{old}) + \Delta w$ ($i = 1 \dots n$),
 - With $\Delta w = \alpha * t * X_i$
 - $b(\text{new}) = b(\text{old}) + \Delta b$,
 - with $\Delta b = \alpha * t$
 - g. Modify W (new), b (new) and ($Y = t$) until the weights are stable or unchanged after a some iterations.

There are a few things to note in the single perceptron algorithm:

- a. Iteration performed continually until all pattern has output of the network with the same target ($Y = t$)
- b. Modify the weighting value is only done on the pattern that contains mistake ($Y \neq t$). These changes are the result of input unit with the target time and rate of comprehension ($\Delta w = \alpha * t * X_i$). Weight value modification will only occur if the input unit $\neq 0$.

- c. The speed of iterations is determined by the used pattern (with $0 \leq \alpha \leq 1$). The larger the value of α , the fewer iterations will be required. However, if the value of α is too large, it will break the correct pattern and the system could become slower.

2.6.2. Testing

Data testing is done using the best weight from the training results, which will be used as the multiplier on Green Mustards features.

- The training process produces the best Weights Input and Weight.
- Calculate the test object with the equality:
$$Y_{output} = (W0 * X0) + (W1 * X1) + (W2 * X2) + (W3 * X3) + (W4 * X4)$$
 - If value of Y_{output} more than 0.5 then the result is Fresh,
 - If result Y_{output} less than -0.5 then the result is Very Withered,
 - If Y_{output} in range of -0.5 to 0.5 then the result is slightly Withered

3. Implementation

This step explains how the method is used to solve problems systematically as well as analysis and experiments conducted to build application identification freshness of Green Mustards. In General, there are some steps done to identify the freshness level of Mustard Greens from a digital image is shown in Figure 7 below:

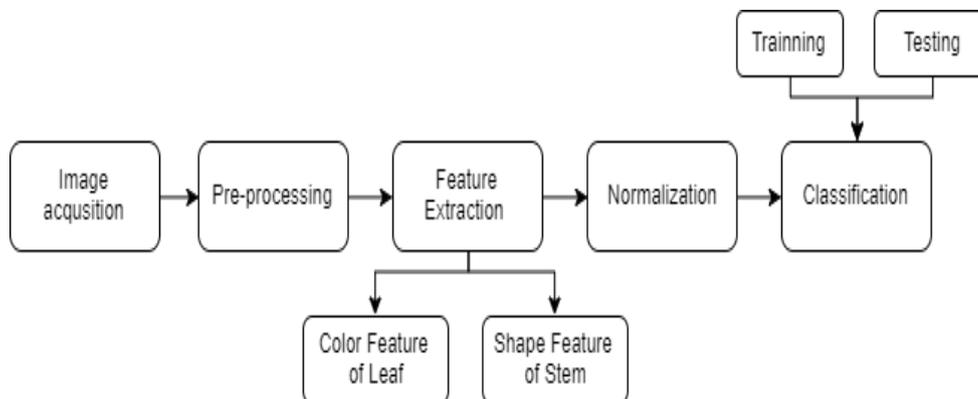


Figure 7. Process Flow Applications

From Figure 7, there are five main system block, image acquisition of Mustard Greens, image preprocessing step: segmentation, RGB normalization, grayscale and binary color conversion, morphological operation, and color and shape feature extraction, classification stage: training the dataset to get best weight value, and testing the best weight with data test.

Image Dataset has some considerations as follows:

- a. Types of Mustard Greens: local / Indonesian Mustard Greens, the image of Mustard Green are taken from the stem until its leaf, the stem length is around 20 - 35cm.
- b. Background image: White color object
- c. Camera resolution: 12Mp Smartphone Camera
- d. Distance object acquisition from the camera is around 20 – 40 cm
- e. Acquisition time at noon, indoor, with maximum light intensity
- f. Image Extensions: .jpg

Fresh Green Mustards samples obtained from a traditional market. The following data shows in Table 1 is the data amount used to classify the freshness levels:

Table 1. Data Sample for this Research

Data Sample of Mustard Greens Image			
Data Training			Data Testing
Fresh	Slightly Withered	Very Withered	Random Category
35 Image	35 Image	35 Image	22 Image

The detail flowchart of the system shows in Figure 8 below. The steps are as follows:

- a. Image acquisition is done and provide a dataset for training and testing the best weight.
- b. Image Pre-Processing steps:
 - **Resize**
Resizing the image to make smaller size image and ensure faster computation process
 - **Rotate**
Rotating the image to convinced the leaf is in the upper part of the image
 - **Crop**
Cutting off a portion of the image. This step is optional if camera capturing object further than required and make the image objects look smaller
 - **Image segmentation with binary tresholding**
Separating the object with the background image, in order to minimize the errors caused by background objects pixel value
- c. The extraction of features:
 - **Color features of Leaf (first order Statistics: mode) using image RGB pixel value**
 - **Stem Curvature Features (Binary Morphology Thinning) and thinning results will be calculated to determine the level of curvature**
- d. Classification:
 - **Training: search for the best weighting.**
 - **Testing: generate output results of identification using best weight**

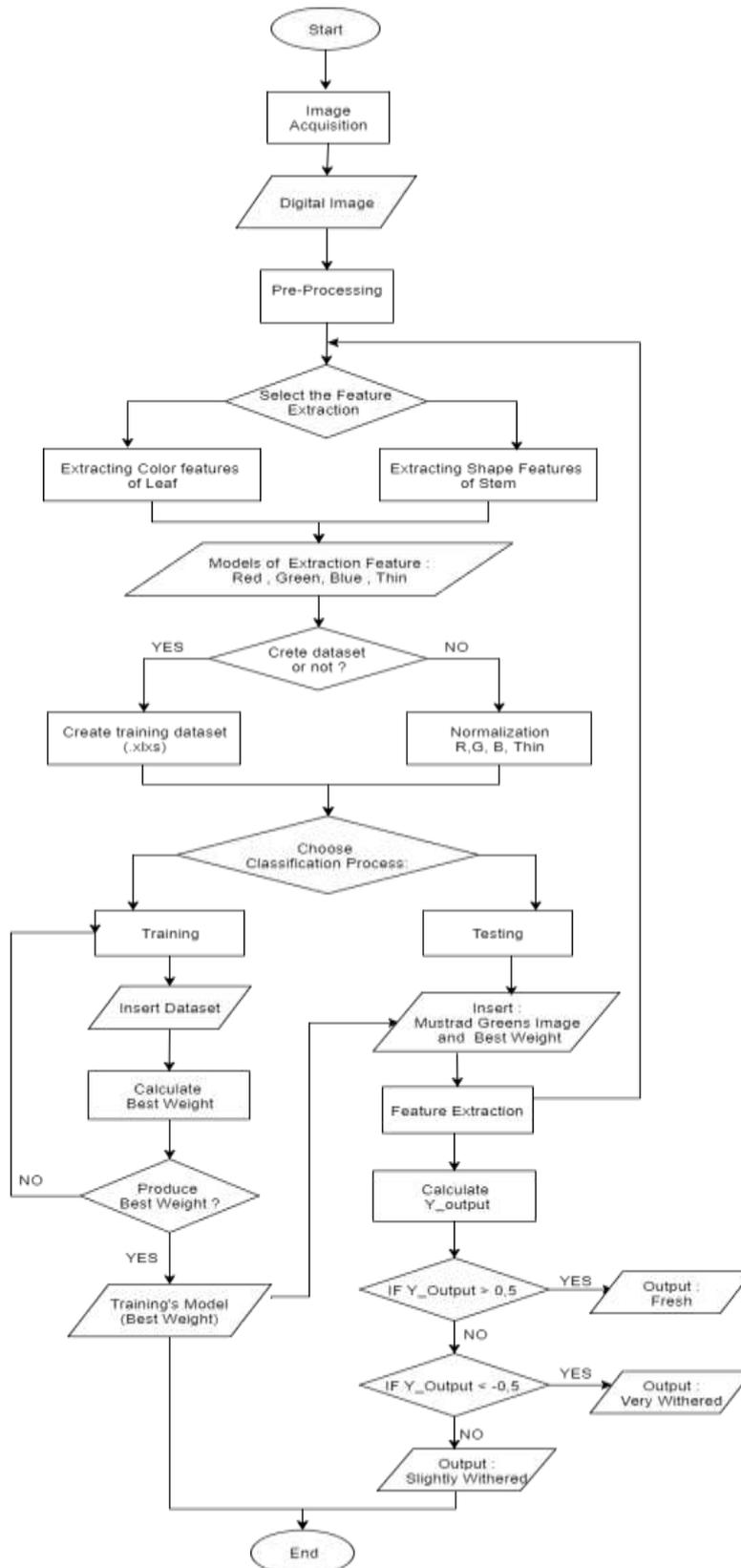
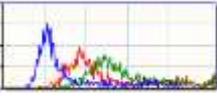
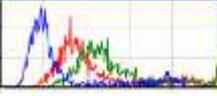
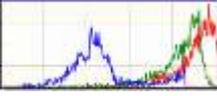


Figure 7. Flowchart of System Proposed

Table 2. Example Result of Extraction Feature

Input	Segmen	Thinning	Mode & Thin	Histogram Segmentation
			R: 91 G: 122 B: 53 T: 35	
			R: 81 G: 124 B: 46 T: 2	
			R: 242 G: 231 B: 105 T: 74	

The training dataset had a total of 105 data for each category (Fresh, Slightly Withered, and very Withered) taken from 35 image for each category, with the initial input:

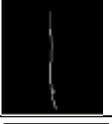
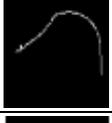
- The initial weight :
Red =W0 , Green = W1, Blue = W2,
Thin = W3, Bias = W4
All W0=W1=W2=W3=W4= 0 (zero)
- Bias = 1
- Learning Rate = 1
- Target: (Fresh = 1), (Slightly Withered = 0), (Very Withered = -1)

Table 3. Best Weight of Training Dataset from Experiments Result

Best Weight	
Epoch = 100000	
W0	-0.6045197740117
W1	-1.1125000000082
W2	4.83333333334696
W3	-7.169642857183
W4	1

From the training experiment data, best weight are achieved with the values shows in Table 3. Testing the data is done by comparing the freshness level from the visual perception and the output of the application. The following Table 4 is a result of the test conducted:

Table 4. Result Testing in Application

Testing Image	Thinning Result	Visual Perception	Feature Extraction	Output Application
		Slightly Withered	R : 80 G : 129 B : 65 T : 34	Slightly Withered √
		Slightly Withered	R : 133 G : 254 B : 90 T : 20	Slightly Withered √
		Slightly Withered	R : 84 G : 255 B : 56 T : 24	Very Withered x
		Slightly Withered	R : 111 G : 255 B : 69 T : 41	Very Withered x
		Very Withered	R : 69 G : 95 B : 49 T : 72	Very Withered √
		Very Withered	R : 81 G : 115 B : 73 T : 124	Very Withered √
		Very Withered	R : 105 G : 138 B : 71 T : 108	Very Withered √
		Fresh	R : 100 G : 135 B : 62 T : 5	Fresh √
		Fresh	R : 114 G : 145 B : 78 T : 5	Fresh √
		Fresh	R : 115 G : 145 B : 70 T : 9	Fresh √
		Very Withered	R : 77 G : 115 B : 57 T : 110	Very Withered √
		Very Withered	R : 92 G : 117 B : 60 T : 101	Very Withered √

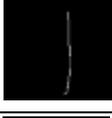
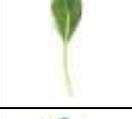
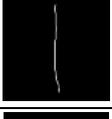
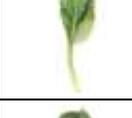
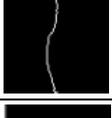
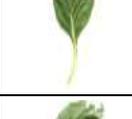
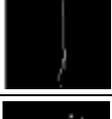
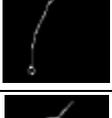
		Very Withered	R : 84 G : 122 B : 71 T : 101	Very Withered	√
		Very Withered	R : 99 G : 125 B : 65 T : 107	Very Withered	√
		Very Withered	R : 81 G : 109 B : 52 T : 110	Very Withered	√
		Fresh	R : 141 G : 255 B : 81 T : 7	Fresh	√
		Fresh	R : 103 G : 140 B : 58 T : 5	Fresh	√
		Fresh	R : 105 G : 223 B : 75 T : 15	Slightly Withered	x
		Fresh	R : 136 G : 196 B : 100 T : 8	Fresh	√
		Fresh	R : 95 G : 139 B : 62 T : 9	Fresh	√
		Slightly Withered	R : 143 G : 225 B : 116 T : 66	Very Withered	x
		Slightly Withered	R : 115 G : 145 B : 74 T : 27	Slightly Withered	√

Table 4 above shows 22 testing samples of the Mustard Greens image. 18 data can be identified the freshness level correctly and 4 data misclassified. The classification accuracy results can be calculated:

$$\text{Classification Accuracy} = \frac{\text{the number of images successfully identified}}{\text{the total number of testing image}} \times 100\% = \frac{18}{22} \times 100\% = 81.81\%$$

4. Conclusion

This research has proposed the identification of Mustard Greens freshness level using an RGB color feature and reduce the dimension using first order statistical Mode, achieved by calculating pixels from the histogram. The other feature is extracting the stem

curvature using thinning morphological method. Classification is done using Neural Network Single Layer Perceptron Classifier. From the experiment result of 22 image testing data of Mustard Greens with a random category has achieved 81.81 % classification accuracy. The use of the Thinning binary morphological method for extracting stem shape of Mustard Greens is very good to distinguish their shape characteristic.

References

- [1] I. Rukmana, "Bertanam Petsai dan Sawi", Yogyakarta: Kanisius, (2007).
- [2] I. P. Budisanjaya, "Perangkat Lunak Pengolahan Citra Untuk Segmentasi dan Cropping Daun Sawi Hijau", Prosiding Conference on Smart-Green Technology in Electrical and Information Systems, (2013).
- [3] I. P. G. Budisanjaya, "Identifikasi Nitrogen Dan Kalium Pada Daun Tanaman Sawi Hijau Menggunakan Matriks CoOccurence, Moments Dan Jaringan Saraf Tiruan", (2013).
- [4] D. Putra, "Pengolahan Citra Digital", Yogyakarta: Andi Offset, (2010).
- [5] R. A., "Pengolahan Citra", Telkom University, [Online]. Available: <http://achmadrizal.staff.telkomuniversity.ac.id>. [Accessed 19 06 2014].
- [6] D. Putra, "Feature Extraction (taken from Intelligent System)", Airlangga University, [Online]. Available: <http://alfian-p-p-fst10.web.unair.ac.id/>. [Accessed 31 03 2013].
- [7] R. Munir, "Pengolahan Citra dengan Pendekatan Algoritmik", Bandung: Informatika Bandung, (2004).
- [8] F. A. Fakhрина, "Thinning Zhang-Suen dan Stentiford Untuk Menentukan Ekstraksi Ciri (Minutiae) Sebagai Identifikasi Pola Sidik Jari Whorl dan Loop", Teknologi Elektro, (2016).
- [9] D. Puspitaningrum, "Pengantar Jaringan Syaraf Tiruan", Yogyakarta: Andi Offset, (2006).