Nutmeg grading system using computer vision techniques

I S Nasution*, K Gusriyan
Agricultural Engineering Department, Universitas Syiah Kuala, Jalan Tgk. Hasan Krueng Kalee 3, Darussalam-Banda Aceh 23111, Indonesia

*Email: i.nasution@unsyiah.ac.id

Abstract. Nutmeg is a popular spice and important commodity that has been recognized by world market. One of the problems during post-harvest handling of nutmeg is quality of the product, particularly: non-uniform quality, damage and contaminated by aflatoxin. This is caused by poor postharvest handling and the products were sorted traditionally. This study performed a computer vision system to sort or to classify quality of the nutmeg based on Indonesian national standard (SNI). In this paper, discriminant analysis and multilayer perceptron (MLP) are implemented to design an automatic classifier of the nutmeg quality. A number of 150 nutmegs are captured by using charge couple device camera. The models are then trained via the data of 124 nutmegas, and their accuracy is tested through the data of 26 nutmegas. Some visual features (texture, shape, and color) of each nutmeg are extracted by using image processing techniques. The results point out that, our proposed grading system could classify nutmeg based on the SNI accurately.

1. Introduction
Nutmeg fruit, a popular spice in food industry, consists of pericarp (fruit covering), seed shell, aril (reddish seed covering), and seed [1]. World nutmeg production was nominated by Indonesia, around 75% nutmeg production covered by this country [2]. Nutmeg have been used as a spice in the food industry for making soups, baked product, sauces, flavoring of meat products therapeutic, and aphrodisiac properties for a long time. In addition to its common use as a kitchen spice, in alternative medicine, nutmeg also utilized as a stimulant, antidiarrheal, carminative, stomachic, tonic, and aphrodisiac [3].

Due to these reasons, nutmeg has high economic value and is being a great potential market to support industrial in the field of spices, food, beverages, medicines, perfumes, and also cosmetics. However, lack of the quality control leads to decreasing price of the commodity, particularly the selection of nutmeg still done by hand. Grading of nutmeg is necessary and usually segments the nutmeg in different homogenous groups according to its specific characteristics like shape, weight, broken, and contaminated. According to Indonesian national standard SNI 01-0006-1993, nutmeg quality consists of three categories, includes: ABCD average, shrivel, and BWP (broken, wormy, punky) as shown in table 1.
Table 1. Nutmeg quality based on SNI 01-0006-1993

|                  | ABCD Average | Shrivels | BWP (broken, wormy, punky) |
|------------------|--------------|----------|-----------------------------|
| Weight (g)       | 5 - 8.33     | 4.11 - 4.99 | < 4.11                     |
| Shape            | all around and less wrinkles | not all around and wrinkles | wrinkles            |
| Broken           | no           | no       | yes                         |
| Contaminated (%) | < 10         | < 10     | not available                |

Some studies have developed various grading methods for nutmeg using image processing and computer vision. Paulus and Suryani [4] developed smart machine of nutmeg sorting by some criteria to categorize the quality of nutmeg such as: color, texture and shape features. They utilized two ultraviolet lights to help recognize the aflatoxin in the nutmeg. Computer vision has been used for classification of nutmeg quality. Artificial neural network analysis was employ based on quality attributes: color and shape [5], discriminant analysis have been used based on only texture features [6]. However, there is a little study about utilizing combination of features such as color, shape, and texture in order to classify nutmeg quality. The current study, aims to classify the nutmeg quality by employing discriminant analysis and neural network methods. Image analysis was used to extract the feature information related to the quality standard according to the Indonesian national standard (SNI 01-0006-1993).

2. Materials and methods

2.1. Nutmeg samples
A number of 150 nutmegs that have been dried and peeled from shell skin were used. The samples were categorized into three nutmeg qualities (ABCD average, shrivel and BWP) according to the Indonesian national standard aforementioned, as shown in figure 1.

![Nutmeg samples](image1.jpg)

Figure 1. Nutmeg samples according to Indonesian national standard. (a) ABCD average; (b) Shrivel; (c) BWP (broken, wormy, punky)

2.2. Image acquisition
The images were capture with the kinect v2 camera. The kinect v2 has a color image resolution of 1920 x 1080 pixels and provides field of view 84.1 x 53.8 degrees, resulting in an average of about 22 x 20 pixels per degree. This improves the color image detail with a factor of two in horizontal and vertical direction. This is an open improvement for scenario’s that use the color image for capturing videos or images, segmentation, object recognition and so on.
2.3. Image segmentation
Nutmegs and background objects were extracted according to L*a*b* color space and Otsu’s thresholding method [7]-[8]-[9]. Otsu’s method based on L* channel was used to distinguish between the objects and background. The Otsu thresholding is an optimal threshold for binarizing an image with a bimodal intensity histogram. Then, unwanted small objects (noise) were removed.

2.4. Features extraction (colour, shape and texture)
In order to determine the color of the nutmeg, the mean of the color array for red, green and blue channel was calculated. Some shape features such as area, perimeter, diameter, and circularity were utilized to describe a given object. Area was calculated based on the number of pixels in a shape. Perimeter is the number of pixels in the boundary of the object. If X₁, …, Xₙ is a boundary list, the perimeter is given by:

\[ Perimeter = \sum_{i=1}^{N-1} d_i = \sum_{i=1}^{N-1} |X_i - X_{i+1}| \]  

(1)

The distances \( d_i \) is equal to 1 for 4-connected boundaries and 1 or \( \sqrt{2} \) for 8-connected boundaries.

Circularity measures the ratio of the area of an object to the area of a circle with the same convex perimeter. The result is equal to 1 for a circular object and less than 1 for an object that departs from circularity. Moreover, diameter is defined as the maximum distance between two boundary points of nutmeg region.

\[ \text{Circularity} = \frac{4\pi \times \text{Area}}{\text{Convex perimeter}^2} \]  

(2)

The five texture features are extracted from each color channel using gray level co-occurrence matrix (GLCM) of the image. The features extracted from the GLCM are given below:

\[ \text{Contrast} = \sum_x \sum_y (x - y)^2 \cdot p(x, y) \]  

(3)

\[ \text{Correlation} = \frac{\sum_x \sum_y (x, y) \cdot p(x, y) - \mu_x \cdot \mu_y}{\sigma_x \sigma_y} \]  

(4)

\[ \text{Energy} = \sum xy p(x, y)^2 \]  

(5)

\[ \text{Homogeneity} = \sum_{x, y} \frac{p(x, y)}{1 + |x - y|} \]  

(6)

\[ \text{Entropy} = -\sum_x \sum_y p(x, y)^2 \log[p(x, y)] \]  

(7)

Where, \( p(x, y) \) = intensity of image, \( \mu \) is mean image intensity and \( \sigma \) is standard deviation.

2.5. Classification
2.5.1. Linear discriminant analysis (LDA). The LDA examines for a group of basis vectors, which has the smallest within class scatter and the largest between class scatter. Beside dimensionality reduction, the LDA also searches the directions for maximum discrimination of the classes. The basic steps of the LDA can be followed according to Soleimanipour et. al. [10].
2.5.2. Multilayer Perceptron (MLP). A classification with single-layer neural nets is not adequate in many classification applications. In order to get a classifier that can separate classes, more layers or hidden layers can be added to the net. The obtained multi-layer neural net (see figure 2) consists of an input layer, one or several hidden layers and output layer. MLP utilizes a supervised learning technique called backpropagation for training.

![Figure 2. Multi-layer neural network: input layer, hidden layer, and output layer (from left to right)](image)

In this study, hyperbolic tangent function was used as an activation function of the hidden layers, as follows:

\[ f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \]  

(8)

Whereas, softmax activation function is used for the output layer.

\[ f(x) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \]  

(9)

3. Results and discussion

3.1. Features extracted

The texture features which comprise the contrast, correlation, energy, homogeneity, and entropy of the nutmeg quality are shown in figure 3. The entropy has a significant influential in distinguishing the three nutmeg qualities, continued by homogeneity, correlation, and contrast. However, energy has insignificant values that difficult to distinguish among the three nutmeg qualities. Figure 4 depicts the variation of mean values of Red, Green, and Blue channel, respectively. As it can be seen in figure 4, the value of the pixels in the R channel is much higher than both Green and Blue channel. However, among each channel tend to have less significant. Different shape features are calculated particularly on area, diameter, perimeter and circularity of nutmegs. Shape features calculated for different nutmegs are shown in table 2.

According to Wilks’ Lambda in the discriminant analysis, six features (perimeter, entropy, homogeneity, area, circularity, and correlation) were chosen as the best predictors (data not shown). These predictors were used as input features for MLP classification. Some features used in this investigation are similar to that used by other studies. Another study defines four data predictors to classify nutmeg quality such as area, perimeter, compactness and roundness [5]. In term of texture features, entropy and correlation were preferred as data predictors to distinguish nutmeg quality [6]. This study produced results which corroborate the features of the previous work, in which it may explain relatively good classification of nutmeg qualities.
Figure 3. Nutmeg texture features (Error bars represent standard deviation of mean, n= 50)

Figure 4. Colour features of nutmeg according to the Indonesian national standard (Error bars represent standard deviation of mean, n= 50)

Table 2. Shape features of nutmeg (pixels value ± standard deviation)

| Category | Area        | Diameter   | Perimeter   | Circularity |
|----------|-------------|------------|-------------|-------------|
| ABCD     | 42868.28 ± 3370.299 | 232.657 ± 8.658 | 741.949 ± 29.165 | 0.8124 ± 0.036 |
| Shrivel  | 36772.44 ± 2455.857  | 217.173 ± 5.941    | 685.724 ± 19.955   | 0.7929 ± 0.047  |
| BWP      | 32555.72 ± 2831.301  | 202.877 ± 11.285   | 629.174 ± 30.056   | 0.7247 ± 0.072  |
3.2. Classification using discriminant analysis

All features were extracted and a data matrix was formed. Then this data matrix was processed by a linear discriminant analysis (LDA) for visualizing the clusters. It can be observed from the LDA plot that nutmeg classes are clustered in a well-defined manner (figure 5). It is interesting to see that ABCD average and BWP are distinguishable. Shrivel and BWP are close to each other although they are distinguishable. Overall nutmeg classes are entirely separable.

![Figure 5. Linear discriminant analysis (LDA) of nutmeg quality](image)

According to table 3, 94.7% of original grouped cases correctly classified, whereas 94% of cross-validated grouped cases correctly classified. Studies have compared this results and another investigation [6] and found that they are essentially identical.

| Category        | Predicted results (%) | Original | Cross-validated |
|-----------------|------------------------|----------|-----------------|
|                 |                        | ABCD average | Shrivel | BWP | Mean | ABCD average | Shrivel | BWP | Mean |
| Original        |                         | 94       | 6    | 0    | 94.7 | 94       | 4      | 2    | 94   |
| Shrivel         |                         | 6        | 94   | 0    |       | 6        | 94     | 0    | 94   |
| BWP             |                         | 0        | 4    | 96   |       | 0        | 6      | 94   |       |

N = 150 samples

3.3. Classification using MLP

The MLP method was utilized to predict quality of nutmeg. This method employed six inputs as above-mentioned, and chooses two hidden layers as the best layers. The first hidden layer has five nodes, and the second hidden layer has four nodes. Table 4 displays the percentage of classification of nutmeg quality, in average, which able to classify 99.2% of accuracy for data training and 100% for data testing, respectively. These results are similar to that found in Dinar et al. [5] who stated the accuracy of nutmeg classification was 100%. In contrast to earlier findings, however, greater input...
nodes and hidden layers were utilized in this study. Whereas, more input nodes and hidden layers tend to be more steady result.

According to table 4, cross entropy error of training and testing are 5.75 and 0.19, respectively. The cross entropy function considers a binary classification error function. It brings to faster training and improved generalization and has better convergence rates. Moreover, the cross entropy function leads to minimization of error with a small range of number of epochs.

### Table 4. Classification results of nutmeg quality using multilayer perceptron (MLP)

| Sample  | Number of observations | Predicted results (%) | Cross entropy error |
|---------|------------------------|-----------------------|--------------------|
|         |                        | ABCD average          | Shrivell BWP Mean  |
| Training| 124                    | 100                   | 97.8 100 99.2      | 5.75               |
| Testing | 26                     | 100                   | 100 100 100        | 0.19               |

### 4. Conclusions

The computer vision techniques were investigated in order to classify nutmeg based on the Indonesian national standard. The MLP method is gave higher results compared to the Least Discriminant Analysis (LDA) method. The adoption of the MLP method in grading of nutmeg will be of immense benefit for further study, such as implementation on a real time grading system. Therefore, research in this area has dealt with trials on a laboratory scales and hence it desires more intensive and comprehensive study.

### References

[1] Kiani S, van Ruth S M, van Raamsdonk L W D and Minaei S 2019 Hyperspectral imaging as a novel system for the authentication of spices: A nutmeg case study Lwt *104* 61–9

[2] Leela N K Nutmeg and mace. *Chemistry of spices* (Wallingford: CABI) pp 165–89

[3] Evans W C, Evans D and Trease G E 2009 *Trease and Evans pharmacognosy* (Saunders/Elsevier)

[4] Paulus E and Suryani M 2019 Image analysis for smart machine of nutmeg sorting *J. Phys. Conf. Ser.* *1196*

[5] Perhimpunan Teknik Pertanian Indonesia L, Suyantohadi A and Fallah M A F 2012 *Jurnal keteknikan pertanian*. vol 26

[6] Dinar L, Suyantohadi A, Affan M and Fallah F 2013 Quality Criteria for Determination of Seeds Nutmeg (Myristica fragrans Houtt) Based Texture Analysis Using Digital Ima *Agritech* *33* 81–9

[7] Otsu N 1979 A Threshold Selection Method from Gray-Level Histograms *IEEE Trans. Syst. Man. Cybern.* *9* 62–6

[8] Nasution I S 2017 Non-Destructive Measurement for Estimating Leaf Area of Bellis perennis *Rona Tek. Pertan.* *10* 58–64

[9] Nasution I S 2018 Technical optimization of a laser marking process of banana fruits and poinsettia bracts (Gottfried Wilhelm Leibniz Universitat Hannover)

[10] Soleimanipur A, Chegini G R and Massah J 2018 Classification of anthurium flowers using combination of PCA, LDA and support vector machine *Agric. Eng. Int. CIGR J.* *20* 219–28