Classification Quality of Tobacco Leaves as Cigarette Raw Material Based on Artificial Neural Networks

Yuslena Sari[*1], Ricardus Anggi Pramunendar[*2]

[1]*Departement of Technology Information of Lambung Mangkurat University, Banjarmasin, Indonesia*

[2]*Department of Computer Science, Postgraduate Program, Dian Nuswantoro University, Semarang, Indonesia*

Abstract - *Determination of the current tobacco grade classification performed by the tobacco commonly called grader with a variety of human frailties. Therefore it is necessary to develop classification automation tools. But earlier experiments need to be done first, in this case using Backpropagation Neural Network classification approach.From this research was obtained increased accuracy for the classification grade tobacco leaf with Backpropagation Neural Network method obtained an accuracy of 77.50%. This indicates that the feature extraction parameters such as shape, color, and texture applied to a Neural Network Backpropagation method can produce a level of accuracy that is quite accurate. Tests were also carried out to produce a level of precision and recall satisfactory as well. Based on the data testing eksperimet of 40 tested for classification grade tobacco leaf there are 8 different datasets that result accuracy between Backpropagation Neural Network with a grader.*

Keywords —*image processing, classification, tobacco, backpropagation neural network.*

I. INTRODUCTION

One of the most highly valued plantation commodities in many countries such as China, India, Brazil, America, Europe, Zimbabwe, Malawi, Russia and Indonesia is the tobacco commodity [1]. High economic value on tobacco is one of the causes of tobacco as the main ingredient for cigarette raw materials. The role of cigarette production in Indonesia in the national economy is in the receipt of excise duty [2]. Various types of tobacco [3] are cultivated in Indonesia, both in plantations and plantations managed more professionally by tobacco companies.

However, the quality of cigarettes in addition to being influenced by the production process can also be strongly influenced by the quality of tobacco leaves as the main raw material. Tobacco leaves have a very diverse type depending on the seeds, planting land, and maintenance process until post-harvest processing [4]. Consequently to obtain a mixture / herb favored, the necessary raw tobacco in various grade.

Determining the classification of class or grade tobacco is currently done by experts of tobacco or so-called grader. Grader has the expertise measure and analyze the quality of tobacco in order to obtain the grade tobacco in groups separated according to category. Graders must have a keen sight and smell in order to distinguish the good quality of tobacco leaves. Such capability must be maintained continuously in order to obtain a stable and consistent classification decision meeting the expected quality [5]. It is devastating for cigarette production, especially in terms of quality, so it requires the determination of classification class or grade of tobacco in a more precise, accurate, and fast.

II. REVIEW OF LITERATURE

Research quality non-local tobacco leaf is encountered Xinhong Zhang research on tobacco leaf quality by using the color, texture, shape through the Fuzzy Comprehensive Evaluation. By using this approach Xinhong Zhang managed to classify the tobacco leaves into three grades, but accuracy is achieved just 72% [5]. High accuracy in the classification of tobacco leaf quality is desirable to determine the quality of tobacco leaf can be trusted to be an alternative quality determination manually. Therefore, it is necessary to improve the accuracy of tobacco leaf quality classification more than previous research.

There are many methods related to the classification [6] [7] [8]. Some have been compared by Shahin in classifying apples. Shahin tried to compare three methods: neural networks, fuzzy logic, and Bayesian Neural Network with better results, ie 88% compared to the fuzzy logic that is 80%. Bayesian classifier has a slightly lower performance than the fuzzy logic that is 79% [11]. Based on the research Shahin, Neural Network has better accuracy in classification.

III. EXPERIMENTAL SET UP

A. Proposed Methodology

Experiments in reseach is a focus in step of recognition grade tobacco leaf. Prior to this recognition stage, did image acquititon (initial acquisition of data), pre-processing (pre-treatment data) and feature extraction. To test the data to be tested, first create training data. From the training data, the algorithm BPNN testing on test data by comparing the data training, this is called the process of pattern recognition of learning.

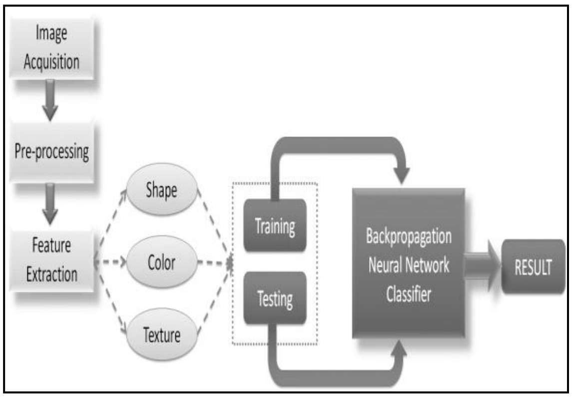


Figure 1. Proposed stages.

Performance measurement is done by calculating the average error that occurred through massive Mean Square Error (MSE). The smaller the MSE value says the closer the predicted value to the true value. Thus can be known level of accuracy of BPNN algorithm.

B. Data Acquisition

This research uses experimental research type with the first step is data collection. The data is taken directly on the tobacco leaf plantation in Temanggung. 200 obtained from various grades of tobacco leaf to be photographed as a dataset in this study. For equipment used is a red cloth fabric, tripod and Canon 350D DSLR camera with telephoto lens. The settings on the DSLR camera is auto focus and ISO 800. The distance between the camera and the tobacco leaf to be photographed is 30 inches. The 30-inch distance is the highest distance of the largest tobacco leaves. So that the biggest tobacco leaves will still look big and the small ones will still look small due to the same capture scale. The tobacco leaves that have been taken next will be inserted into the computer manually.

C. Preprocessing Images

Initial data processing is required to prepare completely valid data before it is processed further. Performed several stages before the leaf image can be extracted its features. This step applies to data training and data testing. At the start of image processing will use MATLAB tools in order to get the object leaves without being accompanied by a background that is obtained when shooting. Formula to remove the background color is by applying the Euclidean distance search method.

2a

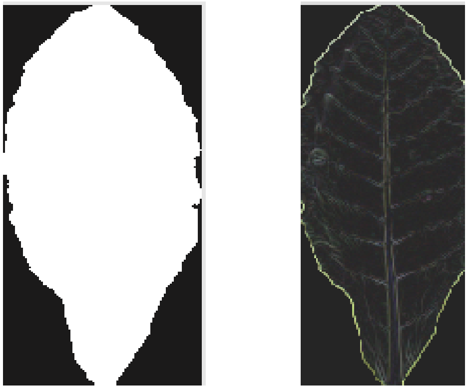
2b

Figure 2. Sort tobacco leaves (2a) and a photo shoot leaf (2b)

D. Feature Extraction

Three (3) feature extractions were performed in this study based on shape, color, and texture. The first process for color extraction is done by obtaining distance color. Furthermore in obtaining the form feature is done by utilizing morphology erodi and delasi. The morphology to remove any noise. And last in getting the texture features used by search entropy value of a previous image by using edge detection Robert methods.

(a) (b)



(c) (d)

Figure 3. (a) The image of tobacco leaves before extracted (b) extracted color (c) extracted form (d) the extracted texture

IV. IMPLEMENTATION AND RESULT ANALISYS

Experiments using Backpropagation Neural Network produced an accuracy of 77.50% with 77.78% recall and precision value 77.42%. The following table generated accuracy:

Table 1 Confusion Matrix To Backpropagation Neural Network

	True A	**True B**	**True C**	**Class precision**
pred A	17	4	1	77.27%
pred B	2	8	0	80.00%
pred C	2	0	6	75.00%
Class recall	80.95%	66.67%	85.71%	

From the table above obtained recall value to True A, B, C and precision for Pred values A, B, C was calculated as follows:
True Recall A = 17 (17 + 2 + 2) = 0.8095 = 80.95 %
Recall True B = 8 (4 + 8 + 0) = 0.6667 = 66.67 %
Recall True C = 6 (1 + 0 + 6) = 0.8571 = 85.71 %
Precision Pred A = 17 (17 + 4 + 1) = 0.7727 = 77.27 %
Precision Pred B = 8 (8 + 2 + 0) = 0.8 = 80.00%
Precision Pred C = 6 (2 + 0 + 6) = 0.75 = 75.00%
From Table 1 above can be calculated the value of its accuracy, as follows:
Accuracy = (17 + 8 + 6) 40 = 0775 = 77.50% accuracy of 77.50% obtained from the correct amount of tobacco leaf is classified in a particular grade in comparison with the total dataset of the tobacco leaf is classified. The correct amount of tobacco leaf is classified in grade A (true positive A) is 17 datasets, grade B (true positive B) is 8 datasets and grade C (true positives C) as 6 dataset. The number of the three is compared with the total number of dacquets of tobacco leaves of 40 dataset so as to produce accuracy of 77.50%.

V. CONCLUTION

The results of research experiments on improving the accuracy of grade classification of tobacco leaf is able to improve the accuracy of previous studies with different methods. Prior to the classification of tobacco leaf grade, this research performs image acquisition stage, image pre-processing and feature extraction to obtain dataset in the form of numeric value which then done classification step. From the correct dataset classified into a certain grade compared to all tested datasets obtained the accuracy of the grade classification of the tobacco leaf. Increased accuracy for the classification of tobacco leaf grade with Backpropagation Neural Network method obtained accuracy of 77.50%. Besides the value of accuracy, also obtained the value of precision and recall value from the results of this experiment. A comparison of 40 data testing tested for the classification of tobacco leaf grade with Backpropagation Neural Network contained 8 datasets with different accuracy results from graders. Before classifying the tobacco leaf grade, a grader must be in a good or ideal state to classify so as to minimize misclassification.

In addition, for future research there are several suggestions, namely:

1. Reproduce the dataset of tobacco leaf image to get more accurate results.
2. Improve the process of retrieval and preliminary data processing to obtain a good tobacco leaf image to be classified.

REFERENCES

[1] D. Guru and P. B. Mallikarjuna, Spots and Color Based Ripeness Evaluation of Tobacco Leaves for Automatic Harvesting. 2010.

[2] S. Barber, A. Ahsan, S. M. Adioetomo, D. Setyonaluri, U. Indonesia., and Lembaga Demografi, Ekonomi tembakau di Indonesia. 2008.

[3] T. Wibowo, "Potret industri rokok di indonesia," Kaji. Ekon. Dan Keuang., vol. 7, no. 2, pp. 83–107, 2003.

[4] F. Zhang and X. Zhang, "Classification and quality evaluation of tobacco leaves based on image processing and fuzzy comprehensive evaluation.," Sensors (Basel)., vol. 11, no. 3, pp. 2369–84, Jan. 2011.

[5] X. Zhang and F. Zhang, "Images Features Extraction of Tobacco Leaves," 2008 Congr. Image Signal Process., no. X, pp. 773–776, 2008.

[6] G. S. K and S. N. Deepa, "Analysis of computing algorithm using momentum in neural networks," Learning, vol. 3, no. 6, pp. 163–166, 2011.

[7] K. Husnul and S. Yuslena, "Prediksi kualitas hasil hutan lahan basah menggunakan Backpropagation," JTIULM, vol. 1, pp. 1–9, 2016.

[8] R. A. Pramunendar and A. Syukur, "AN IMPROVED TECHNIQUE OF COLOR HISTOGRAM IN IMAGE CLUSTERING USING IMAGE MATTING," J. Theor. Appl. Inf. Technol., vol. 51, no. 2, pp. 196–200, 2013.

[9] S. Kai, "A Research of maize disease image recognition of Corn Based on BP Networks L : / o . O," no. 2009921090, pp. 246–249, 2011.

[10] R. A. Pramunendar, C. Supriyanto, D. H. Novianto, I. N. Yuwono, G. F. Shidik, and P. N. Andono, "A classification method of coconut wood quality based on Gray Level Co-occurrence matrices," 2013 International Conference on Robotics, Biomimetics, Intelligent Computational Systems. pp. 254–257, 2013.

[11] M. a. Shahin, E. W. Tollner, and R. W. McClendon, "Artificial Intelligence Classifiers for sorting Apples based on Watercore," J. Agric. Eng. Res., vol. 79, no. 3, pp. 265–274, Jul. 2001.

[12] (Size 8) S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.