

Artificial Neural Network for Classification of Immature and Mature Coffee Beans Using RGB Values

William R. Eustaquio¹, Jesusimo L. Dioses Jr.²

¹College of Arts and Sciences, Isabela State University, Echague, Isabela, Philippines,
william.r.eustaquio@isu.edu.ph

²College of Computing Science, Information and Communication Technology, Isabela State University,
Echague, Isabela, Philippines, jesusimo.l.dioses@gmail.com

ABSTRACT

This paper is about differentiating immature coffee beans from mature coffee beans using their red, green and blue color values, and using these features as variables for classification using the feed forward back propagation artificial neural network. This paper is an extension of the previous article wherein the classifier used are the 23 machine learning algorithms of MATLAB's Classification Learner App. The same dataset was used from the previous study but the classifier was replaced by the feed forward back propagation artificial neural network. In the previous study the highest classification accuracy was achieved by the Quadratic Support Vector Machine with 94 % accuracy and 0.62 seconds training time. In this paper the feed forward backpropagation artificial neural network achieved 95 % classification accuracy and training time of 5 minutes. Comparing the two classifier in differentiating immature coffee beans from mature coffee beans, higher accuracy was achieved by feed forward backpropagation artificial neural network but it also required longer training time.

Key words : Immature coffee beans; Mature coffee beans; ; RGB values; Feed Forward Back Propagation Artificial Neural Network.

1. INTRODUCTION

Immature coffee beans can be differentiated from mature coffee beans in terms of their red, green and blue color values extracted thru image processing as established by the previous article. In that paper, the best classifier is the Quadratic Support Vector Machine with 94 % accuracy and 0.62 seconds training time. The authors were curious if these results can be bested by artificial neural network (ANN) or that the data mining algorithms are still better classifiers. The research question being answered in this paper is "Which one will yield best result, the previous

study which integrated image processing with data mining algorithms or the present study which combined image processing with artificial neural network.

Image processing is the method of using different manipulation techniques and algorithms so that a desired features can be extracted such as morphology, color and texture [1] from an image. A number of studies have extracted the features such as morphology [2-13], color [14-18] and texture [19] and used the values for classification. Using image processing, the values extracted from the image's features signifies characteristics existent only on a particular class and different from other classes.

Feature values between classes are mostly overlapping and to distinguish and establish the differences between classes a classifier is needed. There are a number of classifiers that are already formulated and their application and performances depends on a particular application. It cannot be generalized that if this classifier is best in classifying coffee bean species, it will also be the best in classifying different classes of other agricultural crops. There are two main classifiers in which their performances are most of the times being compared. The two major groups of classifiers are the data mining algorithms and the artificial neural network. The main difference between the two groups is that training and testing are combine in data mining algorithms using a single continuous dataset, while in artificial neural network, from a single continuous data it is further divided into two, one is the training data and the other is the test data. Data mining algorithms are mathematical in approach while artificial neural network mimics how a human brain works in differentiating classes, it must be trained first, meaning the neurons must identify first that this value is from this class and that values belongs to the other class. After the training process, the model was created, and this model is tested whether it can achieved high accuracy in classification. If the model has low accuracy it can be trained again until a high accuracy is achieve. The rule of

thumb in ANN is the longer the training time , the higher the classification accuracy in testing.

In the following papers [20] and [21] both approaches were compared in classification of coffee bean species using the same dataset. In [20] the ANN was used as classifier while in [21] the data mining algorithms were used. It was found out that ANN has an accuracy of 96.66 % while 94.1 % was achieved by Coarse Tree Algorithm , a data mining algorithm. Another study [22] was done to fit ANN vs data mining algorithm in classification of civet coffee from non-civet coffee wherein features are not extracted by image processing but by near infrared spectroscopy. In [22], ANN achieved 94 to 100 % accuracy while Fine Gaussian Support Vector Machine, another data mining algorithm achieved only 86.5 % accuracy.

2. METHODOLOGY

This paper is the continuation of the previous paper of the authors wherein immature coffee beans was differentiated from mature coffee beans using image processing for feature extraction and data mining algorithms for classification. The image processing part was not repeated in this paper instead the classifier was replaced with artificial neural network. The data set from the previous work of the authors were utilized. There are 100 mature coffee beans and 100 immature coffee beans used in the original dataset. In this paper , the dataset was divided into two groups. For mature coffee beans , 80 samples were allocated as training data and the remaining 20 samples were set aside as test data. The test data were randomly selected from the original total of 100 samples, thus making 80 training samples and 20 test samples. The same was done for immature coffee beans samples. Using MS Excel the training samples were composed of 80 mature coffee beans and 80 immature coffee beans. The first column is the red values, the second column is for green and the third column is for blue. The fourth column is the output or classes indicator. The value “1” signifies the mature coffee and “2” for immature coffee beans. This means that for the first 80 rows the fourth column has “1” values and for rows number 81 to 160 the fourth column has a values of “2” .

The dataset was inputted to the MATLAB workspace. The three variables used are input, output for training data and test for test data The input is the RGB values or the three columns of training data. The input is composed of 80 mature coffee beans and 80 immature coffee beans. The output is the fourth column, composed of 80 rows of “1” and 80 rows of “2”. The test data is composed of 20 rows of RGB values of mature coffee beans and 20 rows of RGB values of immature coffee beans.

After the variables were inputted to the MATLAB’s workspace the “nntool” was called in the workspace , so that operations can transfer to the neural network toolbox of MATLAB.

3. RESULTS AND DISCUSSION

The original data set has been used in this study and values for RGB is shown in Table 1. There is overlapping in the RGB values of mature and immature coffee beans. Mature coffee bean’s RGB values occupied the lower range.

Table 1: RGB Range of Values for Mature and Immature Coffee Beans

Coffee Bean Type	Red	Green	Blue
Mature Coffee Beans	74 to 143	89 to 163	89 to 159
Immature Coffee Beans	90 to 150	100 to 168	104 to 164

Figure 1 show the neural network tool used in this study

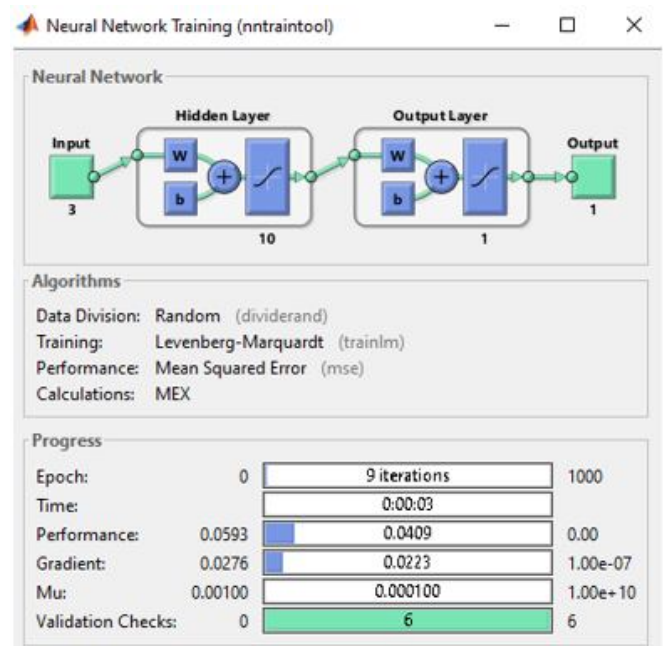


Figure 1: Feed Forward Back Propagation Artificial Neural Network. Used.

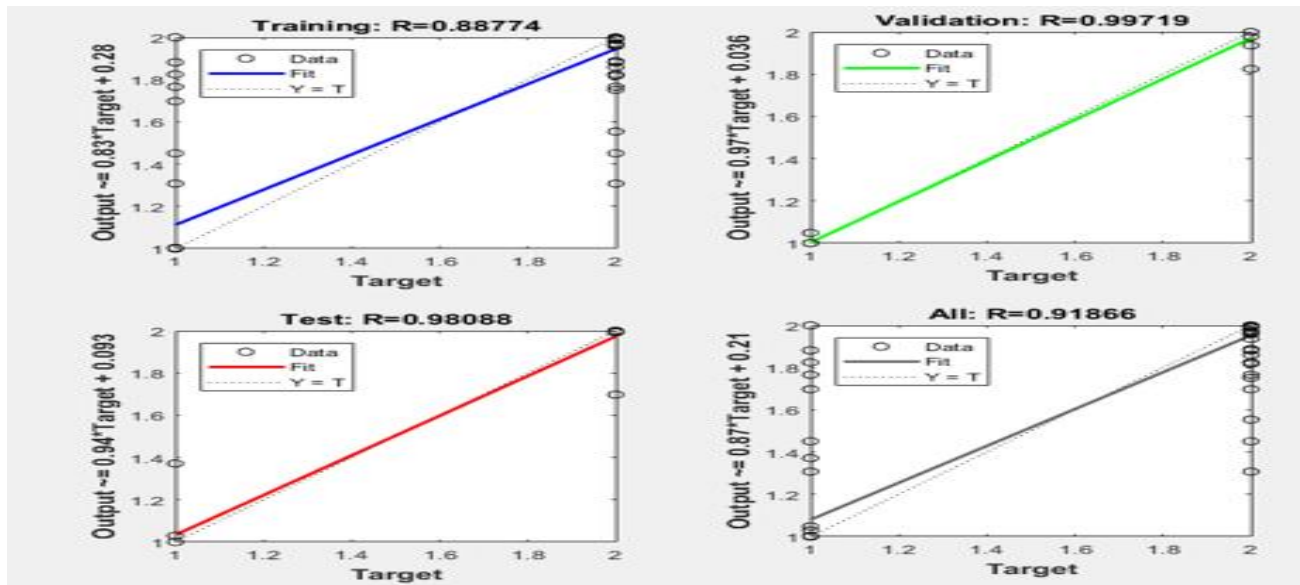


Figure 2: Result of the Training Process

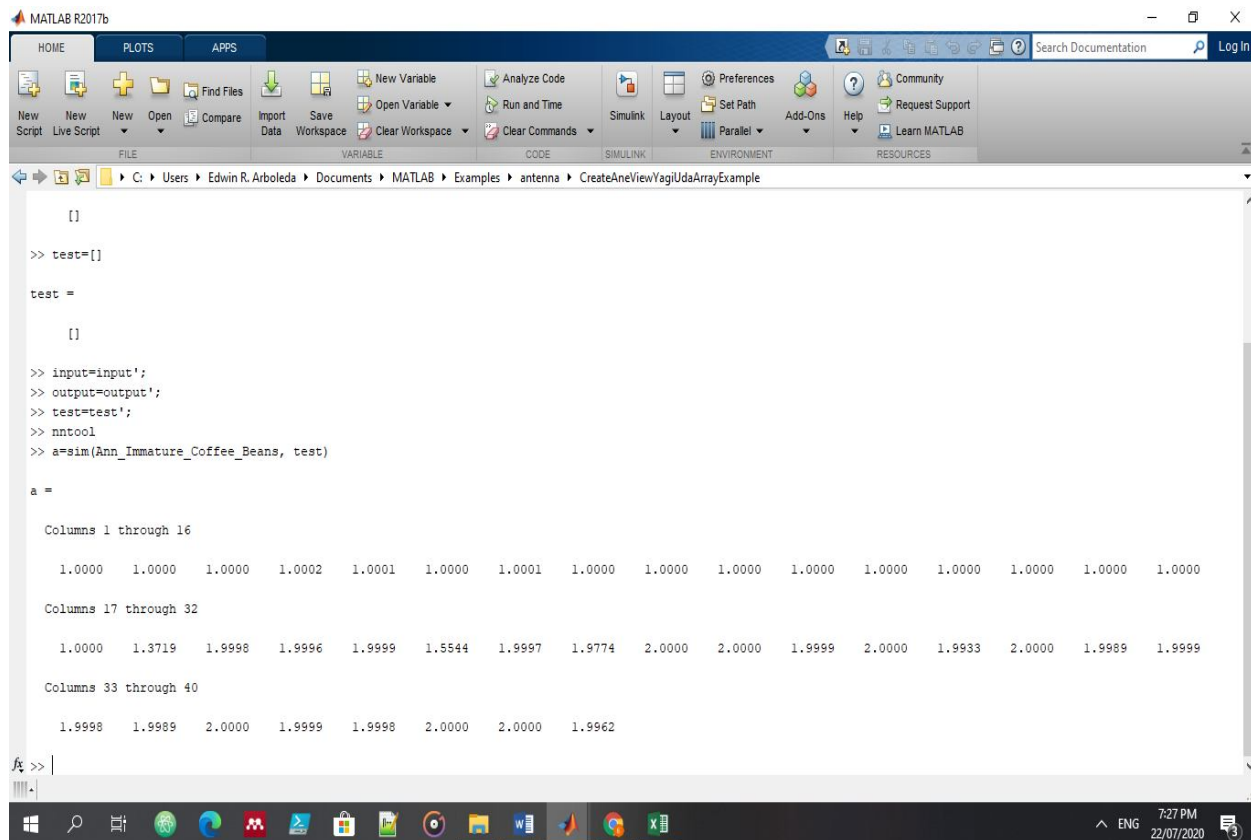


Figure 3: Result of the Testing Process

In Figure 1 there are three input variables. These are the red, green and blue values of mature and immature coffee beans. There are 10 hidden layer and 1 output layer. The output is either 1 or 2, representing mature and immature coffee beans. The training samples were subjected to the developed neural network tool. The training process is to instruct the neural network, which features belong to a particular class. In data mining algorithms the training and testing processes are combined. In ANN the user needs to train the network first, then comes testing. In this study, the training process was done until a desired validation percentage was achieved as shown in Figure 2.

The ANN network can be trained according to the desired level of accuracy that the user desires. The training time can be as little as a few minutes or as lengthy in terms of hours, days, weeks or even months. The longer the training time yields better classification accuracy. For this study the training of 10 times for the ANN network takes about five minutes, but it can be extended. The authors are satisfied with overall R= 0.91866 which is the average classification accuracy for the network. After achieving the desired overall R, the network was imported back to the workspace, to evaluate if it can differentiate mature from immature coffee beans. Test data is composed of 20 mature coffee bean samples and 20 immature coffee bean samples. The developed ANN model ideally should classify twenty 1's and twenty 2's. The result of classification is shown in Figure 3.

It can be seen in Figure 3 that columns 1 to 18 have been classified as 1 or mature coffee beans, which are correct classification while columns 19 and 20 have values of 1.9998 and 1.9996, which means it is '2' or the two beans have been classified as immature beans. There are 18 out of 20 correct classification for mature coffee beans. For columns 21 to 40, all were classified as 2 or immature coffee beans. There are 20 out of 20 correct classification or 100% classification accuracy for immature coffee beans. Overall the classification accuracy is 38 out of 40 or 95% classification accuracy.

The comparison of results of this study and the previous study is shown in Table 2.

Table 2: This Study Vs. Previous Study

Study	Classifier	Accuracy	Training Time
Previous Study	Quadratic Support Vector Machine	94 %	0.62 seconds
This Study	Feed Forward Backpropagation Artificial Neural Network	95 %	300 seconds or 5 minutes

Table 2 shows that there is an increase in accuracy in this study which has a value of 95% accuracy as compared to the previous study with 94% accuracy. However the training time of Quadratic Support Vector Machine is faster which is 0.62 seconds as compared to the training time of the feed forward back propagation artificial neural network which is 300 seconds or 5 minutes. The accuracy of this study can further be increase if the training time is longer.

4. CONCLUSION

The answer to the research question which classifier is the better classifier paper in discriminating immature coffee beans from mature coffee beans has been answered in this study. It can be concluded that artificial neural network is better in classifying immature coffee beans and mature coffee beans. The feed forward back propagation ANN has higher accuracy but longer training time than Quadratic Support Vector Machine.

REFERENCES

[1] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "Green coffee beans feature extractor using image processing," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 18, no. 4, p. 2027, 2020.

[2] E. J. Olaes, E. R. Arboleda, J. L. D. Jr, and R. M. Delloso, "Bell Pepper And Chili Pepper Classification : An Application Of Image Processing And Fuzzy Logic," *Int. J. Sci. Technol. Res.*, vol. 9, no. 02, pp. 4833–4839, 2020.

[3] D. L. Bersabal, E. R. Arboleda, and E. M. Galas, "Coffee Bean Recognition Using Shape Features Using Decision Trees And Ensemble Classifiers," *Int. J. Sci. Technol. Res.*, vol. 9, no. 02, 2020.

[4] A. E. Andaya, E. R. Arboleda, A. A. Andilab, and R. M. Delloso, "Meat Marbling Scoring Using Image Processing with Fuzzy Logic Based Classifier," *Int. J. Sci. Technol. Res.*, vol. 8, no. 08, pp. 1442–1445, 2019.

[5] K. E. N. T. Rabe, E. R. Arboleda, and R. M. Delloso, "Fuzzy Logic Based Vehicular Congestion Estimation Monitoring System Using Image Processing And KNN Classifier," *Int. J. Sci. Tehnol. Res.*, vol. 8, no. 08, pp. 1377–1380, 2019.

[6] C. L. Macalalad, E. R. Arboleda, A. A. Andilab, and R. M. Delloso, "Morphological Based Grain Comparison of Three Rice Grain Variety," *Int. Journal Sci. Technol. Res.*, vol. 8, no. 08, pp. 1446–1450, 2019.

[7] R. E. T. Bae, E. R. Arboleda, A. Andilab, and R. M. Delloso, "Implementation Of Template Matching, Fuzzy Logic And K Nearest Neighbor Classifier On Philippine Banknote Recognition System," *Int. J. Sci. Technol. Res.*, vol. 8, no. 08, pp. 1451–1453, 2019.

- [8] J. G. C. Rancapan, E. R. Arboleda, J. L. Dioses Jr, and R. M. Dellosa, "Egg Fertility Detection Using Image Processing And Fuzzy Logic," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 3228–3230, 2019.
- [9] N. O. Delgado, E. R. Arboleda, J. L. Dioses Jr, and R. M. Dellosa, "Identification Of Mango Leaves Using Artificial Intelligence," *Int. J. Sci. Technol. Res.*, vol. 8, no. 12, pp. 2864–2868, 2019.
- [10] C. I. A. Baluran, E. R. Arboleda, M. G. Dizon, and R. M. Dellosa, "Crab Gender Classification Using Image Processing , Fuzzy Logic And K Nearest Neighbor (KNN) Classifier," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 1458–1462, 2019.
- [11] V. M. D. Manalo, E. R. Arboleda, J. L. D. Jr, and R. M. Dellosa, "Differentiation Among Lettuce (L . Sativa) Seed Varieties Grown In Gourmet Farms , Silang Cavite , Philippines Using Image Processing With Fuzzy Logic And Knn As Classifiers," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 8–11, 2019.
- [12] F. B. Vidad III, E. R. Arboleda, J. L. Dioses Jr, and R. M. Dellosa, "Morphological Based Seedling Comparison Of Natural And Hybrid Lemon," *Int. J. Sci. Technol. Res.*, vol. 9, no. 02, 2020.
- [13] J. A. A. Garcia, E. R. Arboleda, and E. M. Galas, "Identification Of Visually Similar Vegetable Seeds Using Image Processing And Fuzzy Logic," *Int. J. Sci. Technol. Res.*, vol. 9, no. 02, pp. 4925–4928, 2020.
- [14] R. A. Magbayao, E. R. Arboleda, and E. M. Galas, "Identification Of Asian Green Mussel Perna Viridis ' Sex Using Image Processing , Fuzzy Logic And K – Nearest Neighbor," *Int. J. Sci. Technol. Res.*, vol. 9, no. 01, 2020.
- [15] M. J. L. Jacoba and E. R. Arboleda, "Identification of rain cloud amount, rain intensity and possible chance of flooding using image processing and fuzzy logic," *ARPJ J. Eng. Appl. Sci.*, vol. 14, no. 14, pp. 2563–2568, 2019.
- [16] A. J. G. Gonzales, J. A. D. Sosa, E. R. Arboleda, and E. M. Galas, "Filtering Of Faded Coffee Beans Using Image Processing," *Int. J. Sci. Technol. Res.*, vol. 8, no. 12, pp. 3634–3637, 2019.
- [17] V. A. C. Austria, E. R. Arboleda, and E. M. Galas, "Image Processing Of Clean And Dirty Dishes To Design And Construct A Fuzzy Logic Dishwasher," *Int. J. Sci. Technol. Res.*, vol. 8, no. 12, pp. 8–11, 2019.
- [18] J. N. C. Sarino, M. M. Bayas, E. R. Arboleda, E. C. Guevarra, and R. M. Dellosa, "Classification Of Coffee Bean Degree Of Roast Using Image Processing And Neural Network," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 3231–3233, 2019.
- [19] A. M. Castillo, R. D. Aradanas, E. R. Arboleda, A. A. Dizon, and R. M. Dellosa, "Coffee Type Classification Using Gray Level Co- Occurrence Matrix Feature Extraction And The Artificial Neural Network Classifier," *Int. J. Sci. Technol. Res.*, vol. 8, no. 10, pp. 1463–1465, 2019.
- [20] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "Classification of Coffee Bean Species Using Image Processing , Artificial Neural Network and K Nearest Neighbors," in *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, 2018, no. May, pp. 1–5.
- [21] E. R. Arboleda, "Comparing Performances of Data Mining Algorithms for Classification of Green Coffee Beans," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 5, pp. 1563–1567, 2019.
- [22] E. R. Arboleda, "Discrimination of civet coffee using near infrared spectroscopy and artificial neural network," *Int. J. Adv. Comput. Res.*, vol. 8, no. 39, pp. 324–334, 2018.
<https://doi.org/10.19101/IJACR.2018.839007>