Cacao Bean Quality Assessment Procedure: A Method for Classification Process

Dindo C. Obediencia, Deborah Go Brosas, and Ritchell Solitario Villafuerte

Abstract

The study focus on Cacao Bean Quality Assessment. The method starts on image processing of sample cacao bean to detect the defected beans on the sample submitted for quality assessment. After the Image Processing, this paper proposes to utilize the Adaptive Neural-Fuzzy Inference System (ANFIS) technique for classification of the cacao bean quality level. The ANFIS technique serves as the medium for identifying the quality level of the cacao bean sample submitted for evaluation and categorization. The input variables are Bean Count, Moldy, Slaty, and Defected Beans. The Output Grade level of Cacao. The Two hundred (200) data sets used were taken from the image processing output and normalized to be trained in ANFIS using the tool in Matlab. Eighty (80) percent for training and Twenty (20) percent for checking. During the training with the use of Matlab software, it resulted in an accuracy rate value of 99.715% with an error rate of 0.285%. The study shows that ANFIS is a technique that can be used efficiently to classify the cacao bean quality based on their category level.

Keywords: Image Processing, Adaptive Neural-Fuzzy Inference System, ANFIS, Matlab, Cacao, Philippines

1.0 Introduction

Cocoa is the main ingredient in the manufacture of chocolate, making it one of the world’s most highly traded tropical crop. In 2015, this industry was at approximately US$9 billion. The cocoa-chocolate Global Value Chain, for which cocoa beans is an essential product, is also growing steadfast. Total trade in the cocoa-chocolate Global Value Chain has doubled over the past ten years, reaching nearly US$44 billion in 2015. Chocolate exports got 56% in the industry, followed by cocoa beans 20%, cocoa butter 12%, cocoa liquor/ paste 7% and cocoa powder 5% (de Vries, 2000).

The cocoa-chocolate global value chain attaches cocoa bean producers with manufacturers and consumers of chocolate in a multimillion-dollar industry. The manufacture mostly takes place in agricultural countries, while the manufacturing and distribution of end products take place in developed
countries. In the past ten years, the demand for chocolate has grown steadfast (DTI Policy-Brief, 2017), motivating several countries already in the industry to improve their production further, and for other countries to take part in the value chain. It is in the attentiveness of the Philippines with its capacity to produce cocoa beans, to take advantage of the increased global demand and the need for additional players in the value chain (de Vries, 2000).

![Image](Asia-Pacific cocoa production: situation in 2014/2015)

Figure 1. Asia Pacific cocoa production (Ramos, 2016)

The increasing global demand represents an opportunity for the Philippines to grow its economy. However, the Philippines’ participation in the cocoa-chocolate (DTI Policy-Brief, 2017) Despite many competitive advantages, the nation has manufactured cocoa for centuries and has excellent climatic and geographical environments for its manufacture. However, exports remain low. Globally, the country ranks 72nd in exports with a global market share of less than 0.01%. Despite a history of cocoa bean production dating back to the 17th century, the Philippines’ current output is limited compared with other global and even regional players, generating a total of just over US$24 million in exports in 2015 (DTI Policy Briefs, 2017). While exports grew steadily over the past years, increasing by 288% from 2005 to 2015, the nation’s participation in the global value chain remains lesser compared to other countries in the area. Manufacture of cocoa beans in the Philippines peaked in the 90s at 35 thousand tons and had since deteriorated due to aging trees, depressed producer prices, and climatic conditions. As a result, by 2015, the Philippines produced roughly 6 thousand tons of cocoa beans, though Industry players evaluated it from 10 thousand to 12 thousand tons. The Philippine Statistics Authority (PSA) places the figure lower, reporting roughly 6 thousand tons in 2015 (de Vries, 2000).

The deteriorating resource has also reduced the number of large-scale processors. Most businesses are either small-scale chocolate companies or cocoa tablet “table” makers, with few firms having significant export operations. Cocoa-chocolate global value chain exports amounted to US$24,000,000 in 2015. The Philippines in its place was a net importer with imports exceeding US$132,000,000 for all products in the cocoa-chocolate global value chain. Overall, only eight Filipino firms exceeded that threshold in 2014. Producing cocoa beans is a major
requirement for participation in the cocoa chocolate global value chain in the Philippines (G. E. de Vries, 2000).

Quality Control

Quality control and food safety are severe trepidations of food manufacturers, governments and consumers that have especially gained attention with some foodborne illness outbreaks in recent years. Because of the processing in cacao bean manufacturing, including heat treatment and the removal of excess moisture, such illness from chocolate products is comparatively less likely, but the cocoa, candy, and chocolate industry still face the same challenge of ensuring that raw materials, including cacao beans, are of high quality and safe. Cacao beans are spontaneously fermented seeds, and as such are subject to a high level of variability depending on growing conditions, genetics, postharvest fermentation and drying of the cocoa beans before shipment or handling. This variability can have a significant impact on the finished chocolate product, so it is imperative to be able to distinguish if beans are correctly fermented, of high quality, and lacking defects. In addition to trying to distinguish correctly fermented beans and maintain a high-quality raw material, there are also the food safety concerns. According to R. Dand (2010), there is the potential for a variety of moisture damage related chemical processes occurring due to high humidity, that can result in mold colonization and/or microbial growth during storage and transport, or the unintentional inclusion of other chemical alterations (or adulterants) to the raw material itself. While the presence of chemical degradation products due to moisture damage does not always indicate an unsafe product, it can often lead to unpleasant off-flavors in the finished chocolate product. Thus, moisture damage of cacao beans is both food safety and quality control concern for this industry (Humston, Knowles, McShea, & Synovec, 2010).

The present method in monitoring the quality control of cacao is somewhat subjective and limited in its ability to quantify bean quality. Typically, 50 to 100 or more cacao beans are cut into half (cut test) to show the core for investigation so that the color, a clear indicator of the state of fermentation can be observed (Humston et al., 2010).

For the determination, 300 beans are opened or cut lengthwise through the middle, visually checking both halves of each bean in full daylight or equivalent artificial light to expose the maximum cut surface of cotyledons tallying each defected bean type separately, and the result for each kind of defect expressed as a percentage of the 300 beans examined. According to M.S. Fowler (2009), the test identifies the beans that are visibly moldy, slaty (i.e., unfermented), infested,
germinated, flat (i.e., containing no nib or cotyledon) and purple or brown. It is, however, very subjective, and Lopez and Dimick noted that it is at best limited to the measurement of bean defects and color (Kongor, Takrama, Budu, Mensah-brown, & Afoakwa, 2013).

The problem with the current method is that it is subjective and laborious. The assessment and segregation process can be inconsistent depending on who is tasked to do the evaluation. This sorting is low speed and less accurate; this is related to human tiredness and discrimination. Thus, the researchers propose an automated identification and classification process of the defected cacao bean. This submitted research is an approach which aims to make use the image processing technique for defected bean identification. To categorize the grade level value of the sample beans will be classified using Adaptive-Neuro Fuzzy Inference System is a kind of neural network based on the Takagi–Sugeno fuzzy inference system. It integrates the features of both neural networks and fuzzy logic principles. This study aims to develop an ANFIS model for cacao bean categorization based on their grade level after the fermentation process with the purpose of having a quick and specific analysis. Classification process will be grade value input based and the value of defective beans found on the sample submitted for categorization.

There are six quality classifications for cacao beans set by the Philippine National Standards Table 1 shows these classes and their respective specifications (PNS/BAFPS 58:2008).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Bean count (per 100 g)</th>
<th>Moldy</th>
<th>Slaty</th>
<th>Defects such as insect damaged, infected beans and germinated beans</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>≤ 100</td>
<td>3</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>1B</td>
<td>101 – 120</td>
<td>3</td>
<td>3</td>
<td>2.0</td>
</tr>
<tr>
<td>2A</td>
<td>≤ 100</td>
<td>4</td>
<td>8</td>
<td>8.0</td>
</tr>
<tr>
<td>2B</td>
<td>101 – 120</td>
<td>4</td>
<td>8</td>
<td>8.0</td>
</tr>
<tr>
<td>Sub-standard &gt; 120</td>
<td>&gt; 4</td>
<td>&gt; 8</td>
<td>&gt; 8</td>
<td>8.0</td>
</tr>
</tbody>
</table>

The given table shows that Substandard Cacao beans which exceed one of the limits accepted for Grade 2 shall be viewed as sub-standard and marked ‘SS.’ Marketed sub-standard cacao shall only be under exclusive contract (Bureau of Product Standards - Philippine National Standard, 2008).

This paper proposed the digital image processing technique based on Convolutional neural network (CNN) approach for detection and classification. The result of the processed image is the categorization of the cacao bean if it is moldy, slaty and defective.

Image processing refers to the processing of the digital image, i.e., removing the noise and any irregularities present in an image using the digital computer. The noise or irregularity may
creep into the image either during its formation or transformation. For mathematical analysis, an image is a two-dimensional function $f(a,b)$ where $a$ and $b$ are spatial (plane) coordinates, and the amplitude of at any pair of coordinates $(a, b)$ is the strength or gray level of the image at that point. When $(a, b)$, and the intensity values of $f$ are all finite, discrete quantities, we call the image a digital image. It is essential that a digital image be a finite number of elements, each of which has a particular location and value. (Chitradevi & Srimanthi, 2014).

The main advantage of Digital Image Processing method is its versatility, repeatability and the preservation of original data precision. The various Image Processing techniques are (Chitradevi & Srimanthi, 2014):

- Image preprocessing
- Image enhancement
- Image segmentation
- Feature extraction
- Image classification

After the Image Processing, this paper proposes to utilize the Adaptive Neural-Fuzzy Inference System (ANFIS) technique for classification of the cacao bean quality level. The ANFIS technique serves as the medium for identifying the quality level of the cacao bean sample submitted for evaluation and categorization.

Adaptative Neuro-Fuzzy Inference System (ANFIS) is a hybrid neuro-fuzzy inference expert system, and it works in Takagi-Sugeno-type fuzzy inference system, which was established by Jyh-Shing Roger Jang (1993). The technique provides a scheme for the fuzzy modeling procedure to learn information about a data set, in order to calculate the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data. This learning method works like that of neural networks (Shastry, Sanjay, & Hegde, 2015).

2.0 Methodology

The quality management in cocoa should be done in a more efficient and precise manner than the previously exercised manual ‘cut-test’ defected bean identification to increase the production of quality cocoa. The proposed application helps to perform the following operations.

An approach for detection and classification of the different cacao bean sample characteristics consists of phases, namely, image acquisition, feature extraction, classification and categorizing input samples according to their graded level value as shown in Figure 2:
This proposed approach has two processes. One is to detect the defected cacao bean after the fermentation process with the use of image processing to classify the bean input sample if it is moldy, slaty and defected (insect damaged, infested and germinated). The second approach is to classify the bean quality level value using the Adaptive Neuro-Fuzzy Inference System.

**Image Processing**

a) Defective Cacao Bean Detection Process

(1) Cacao Bean Image Acquisition
   To decide the collected sample images of bean, if it is defective or not.

(2) Image Pre-processing
   The aim of pre-processing is an improvement of the image data that suppresses unwanted distortions or enhances some image features vital for further processing.

**Image Enhancement and Noise Removal**

Image enhancement changes the visual impact that the image has on the interpreter in a way that improves the information content. Image enhancement techniques are applied where subjected excellence of images is essential for individual interpretation (Kuruvilla, Sukumaran, Sankar, & Joy, 2016).

**Figure 3. Sample Images of Defected Cacao Beans**

(3) Detection and Classification
   The author opted to use the Convolutional Neural Network (CNN). It is a type of feedforward artificial neural network. CNN uses images directly as input. (CNN was proposed by (LeCun, Bottou, Bengio, Haffner, 1998)). CNN consists mainly of three kinds of layers as shown in Figure 3.
During this process, we utilize the preprocessed cacao bean images as input for the detection and classification process. Classified Cacao Bean based on their defect category is the projected output of this phase.

This paper chose to use the CNN technique because according to (Pinto, Furukawa, Fukai, & Tamura, 2017) they employed deep convolutional neural networks, the state-of-the-art machine learning technique, for the image processing. As the results, they succeeded to sort defect beans from 72.4% to 98.7% of accuracies based on the types of defects.

(4) Bean Sample Grade Level Classification
After the image processing procedure, defected cacao beans are classified base on their defect characteristics. The output of this process will serve as the input on the next process, which is the categorization of the sample bean evaluated according to their quality grade level value. This paper opted to use the Adaptive Neuro-Fuzzy Inference technique to predict the grade level value of the assessed sample cacao bean.

**Neuro-Fuzzy Model Development (ANFIS)**

We recommended the Sugeno fuzzy model for a systematic approach to generating fuzzy rules from a given input-output data set.

Stating a standard Sugeno fuzzy rule in the following form (J. Kaplan, 2016):

\[
\text{IF } x_1 \text{ is } A_1 \\
\text{AND } x_2 \text{ is } A_2 \\
\text{AND } \ldots \\
\text{AND } x_m \text{ is } A_m \\
\text{THEN } y = f(x_1, x_2, \ldots, x_m)
\]

where \( x_1, x_2, \ldots, x_m \) are input variables; \( A_1, A_2, \ldots, A_m \) are fuzzy sets; and \( y \) is either a constant or a linear function of the input variables. When \( y \) is a constant, we obtain a zero-order Sugeno fuzzy model in which the consequent of a rule is specified by a singleton. When \( y \) is a first-order polynomial, i.e.

\[
y = k_0 + k_1 x_1 + k_2 x_2 + \ldots + k_m x_m
\]

We obtain a first-order Sugeno fuzzy model.

A six-layer feedforward neural network usually represents Jang's ANFIS.

Figure 5 shows the ANFIS architecture that corresponds to the first order Sugeno fuzzy model.
Layer 1 is the input layer. Neurons in this layer pass crisp external signals to Layer 2. That is,

\[ y_{i}^{(1)} = x_{i}^{(1)}, \]

where \( x_{i}^{(1)} \) is the input and \( y_{i}^{(1)} \) is the output of input neuron \( i \) in Layer 1.

(2)

**Table 1. Propose Input Parameter**

<table>
<thead>
<tr>
<th>Input Variable</th>
<th>Linguistic Value</th>
<th>Numeric Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BeanCount (W)</td>
<td>A</td>
<td>&lt;=100</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>101-120</td>
</tr>
<tr>
<td>Moldy (X)</td>
<td>One</td>
<td>&lt;=3</td>
</tr>
<tr>
<td></td>
<td>Two</td>
<td>4</td>
</tr>
<tr>
<td>Slaty (Y)</td>
<td>One</td>
<td>&lt;=3</td>
</tr>
<tr>
<td></td>
<td>Two</td>
<td>8</td>
</tr>
<tr>
<td>DefectedBeans (Z)</td>
<td>One</td>
<td>&lt;=2.5</td>
</tr>
<tr>
<td></td>
<td>Two</td>
<td>5.0</td>
</tr>
</tbody>
</table>

The created crisp input are the linguistic variable to be passed to Layer 2.

It is the fuzzification level. Neurons in this level execute fuzzification. In Jang, (1993) model, fuzzification neurons have a bell triggering the task. A bell activation task, which has a regular bell shape, is specified as:

\[ y_{i}^{(2)} = \frac{1}{1 + \left( \frac{x_{i}^{(2)} - \alpha_{i}}{\epsilon_{i}} \right)^{2}}, \]

where \( x_{i}^{(1)} \) is the input and \( y_{i}^{(2)} \) is the output of neuron \( i \) in Layer 2; and \( \alpha_{i} \) and \( \epsilon_{i} \) are parameters that control, respectively, the centre, width and slope of the bell activation function of neuron \( i \).

(3)

The output generated below is the membership value of BeanCount using Matlab.

```
Figure 5. Adaptive Neuro-Fuzzy Inference System Architecture (J. Kaplan, 2016)
```

```
Figure 6. Membership Function of BeanCount
```

Layer 3 is the Sugeno-type fuzzy rule layer. A rule neuron receives inputs from the particular fuzzification neurons and calculates the firing strength it denotes. In an ANFIS, the operator product is used to measure the conjunction of the rule antecedents. Thus, as the obtained, the output of the neuron \( i \) in Layer 3 is,
The fuzzy rules were generated using the set of input and output; the antecedent part has eight sections. The antecedent parts of the rules are combined using the maximum and product operator.

Below are the 81 rules generated using the Matlab.

Figures 7b. 81 Generated Rules for Cacao Bean Quality Classification Generated
Layer 4 is the normalization layer. The output of neuron \( I \) as presented in Layer 4,

\[
y_{i}^{(4)} = x_{i}^{(4)}(k_{0i} + k_{1i}x_{1} + k_{2i}x_{2}) = \mu_{i}(k_{0i} + k_{1i}x_{1} + k_{2i}x_{2}),
\]

\[
y_{i}^{(4)} = \frac{\mu_{i}}{\sum_{j=1}^{n} x_{j}^{(4)} + \sum_{j=1}^{n} \mu_{j}} = \mu_{i},
\]

where \( x_{i}^{(4)} \) is the input from neuron \( J \) located in Layer 3 to neuron \( I \) in Layer 4, and \( n \) is the total number of rule neurons. For example,

\[
y_{i}^{(4)} = \frac{\mu_{i}}{\mu_{1} + \mu_{2} + \mu_{3} + \mu_{4}} = \mu_{i}
\]

(5)

Layer 5 is the defuzzification layer. Each neuron in this layer links to the respective normalization neuron, and also receives initial inputs, \( x_{1} \) and \( x_{2} \). A defuzzification neuron computes the weighted consequent value of a given rule as,

\[
y_{i}^{(5)} = \frac{\mu_{i}}{\mu_{1} + \mu_{2} + \mu_{3} + \mu_{4}} = \mu_{i}
\]

(6)

A summation neuron denotes layer 6. Thus computes the sum of yields of all defuzzification neurons and resulted in the overall ANFIS output, \( y \).

**Data Description**

**ANFIS Training**

Matlab software was used to develop a neuro-fuzzy model by presenting a training dataset that contains the desired input/output pairs of the system to ANFIS toolbox for training.

The stages involve are listed below:
- Loading the training and testing data.
- Generate or load an initial FIS model.
- View the FIS model structure once an initial FIS has been generated or loaded.
- Choose the FIS model parameter optimization method: a mixture of back propagation and least squares (hybrid method).
- Choose the number of training epochs and the training error tolerance for training the system.
- Train the FIS model by clicking the Train Now button. This training adjusts the membership function parameters and plots the training error plot(s) in the plot region.
- Validating the trained FIS.

**Figure 8. Proposed ANFIS Structure for Cacao Bean Quality Classification**
3.0 Results And Discussion

Fuzzy Inference System for Cacao Bean Quality Level Classification

The figures show the diagram of the Fuzzy Inference System for Cacao Bean Quality Level classification. There are four input variables which as shown in Table 1 and one output variable which show the bean grade quality levels. Input one (1) was identified as the bean count variable, input two (2) was used to identify the moldy variable, input three (3) for slaty and for the input four (4) is for identifying the defects. Figures 9a to 9d shows the different membership function for the input in classifying the grade quality of the cacao bean.

Membership Function

**Figure 9a. Bean count**

**Figure 9b. Moldy**

**Figure 9c. Slaty**

**Figure 9d. Defects**

**Figures 9a to 9d. FIS Structure of the System**

Experiments were conducted to determine the testing datasets by taking into consideration the accuracies of the system. For the experiment, two hundred (200) data set was loaded to the FIS system. From the data set, eighty percent (80%) was used for training while 20% was used for testing. A training data was used to test the proposed approach using
the Matlab, and a result shows that ANFIS generated a total of 193 nodes, the results of linear parameters were 81, and the non-linear parameters are 24.

During the ANFIS training, it results in an average error of 0.285 and the ANFIS training completed at epoch 2. Test data was used to train the prediction quality of the network during the training. Figure 10a and 10b show the Training and Checking Data.

![Figure 10a. Training Data](image)

![Figure 10b. Checking Data](image)

**Validation**

Validation is needed to test the trained FIS model against the checking data. Figure 11 shows the correctly identified data against the training set.

![Figure 11. Validation Check](image)

**4.0 Conclusion and Future Works**

During the training, the accuracy rate value resulted in 99.715% with an error rate of 0.285%. The current study shows that ANFIS is a technique that can be used efficiently to classify the cacao bean quality based on their category level.

The result generated was only to test the efficiency of the adaptive neural-fuzzy inference system to determine the quality level of the cacao bean sample. For future
works, the researcher will have to explore different image processing technique that will enhance the digitized image of the defected cacao bean sample to come up with a better result.

References


