Copra Meat Classification using Convolutional Neural Network

Rufo I. Marasigan Jr[1], Alvin Sarraga Alon[2], Mon Arjay F. Malbog[3], Joshua S. Gulmatico[4]
[1]Technological Institute of the Philippines, Philippines, rufo.marasigan.jr@tip.edu.ph
[2]Technological Institute of the Philippines, Philippines, aalon.cpe@tip.edu.ph
[3]Technological Institute of the Philippines, Philippines, malbog.monarjay@gmail.com
[4]Technological Institute of the Philippines, Philippines, engr.gulmatico@gmail.com

ABSTRACT

Abstract – Copra in the Philippines is one of the by-products from coconut which contributes as one of the major sources of income of Filipino farmers. During the process of selling the Copra in the market, farmers usually lose in the price competition from the buyers due to the unidentified quality of their Copra. Copra which commonly either overcooked or undercooked are paid as half of the price of the perfectly cooked. This happened due to the lack of information of the farmers in assessing the quality of the processed copra meat. In this study, a Convolutional Neural Network had been evaluated in terms of its *accuracy* by varying the numbers of convolutional layer filters, the size of filters, and its activation function. The identified best parameters were used to develop a CNN algorithm that classifies the quality of Copra. The algorithm was implemented using Tensorflow in a python environment. A series of tests were applied to the final CNN model. Random images of Copra with identified quality were used as testing data. Out of 120 sample images, the final CNN model performs an overall 86% accuracy. The model was also implemented into a simple android application for validation. The confusion matrix and f-score were used to evaluate the performance of the system.

Key words: Android-based Copra Classifier, Convolutional Neural Network, Copra Meat Classifier, Copra Quality, Image Processing,

1. INTRODUCTION

Copra, is dried coconut meat and biodiesel raw materials [1]. In the process, the firm white coconut meat of a matured coconut is either being dried in sunlight or being smoked to remove the moisture content. The results are dry and lightly browned coconut meat called Copra. Then it is deshelled from the outer shell and is further processed for various uses. In the Philippines, as of 2015, there have been 3.517 million hectares of coconut production, 26% of the total agricultural land is the main contribution of coconut. 68 out of 81 provinces are coconut areas [2]. In 1996 it was estimated that there are about 11 million hectares, and around 93 percent coconut area is found in the Asian and Pacific region.

Indonesia and the Philippines having the two largest areas has 3.7 million hectares and 3.1 million hectares respectively. While India is still the is the third-largest producer. On the other hand, Papua New Guinea is the leading roduce r in countries of the South Pacific. The largest source of Copra is from the Philippines, where the value of annual production exceeds US$80 million [3].

Copra production is undergone to a certain process and once finished it will be sorted out to different classes based on its appearance. First is getting the right coconut from the tree. It is better if the selected coconuts are dried or in a tinge of a brown or coffee brown. Next would be the removal of the outer shell (coconut kernel) and the inner shell (coconut meat) that copra workers acquire will be going to the drying process. There are two kinds of drying (also called cooking): sun-drying wherein it will be staying under the sun for hours, and when it has the brown color, it achieves the sun drying. Another way is the smoke drying where the Copra will be placed on the grill, and once it is dry enough or the Copra will be cracked itself, it is now considered dried.

Copra industry in the Philippines proliferates, and people are facing emerging problems. Copra is processed, dried over the course of a few days by coconut farmers. It is difficult, labor-intensive, and dirty work that usually place in remote locations and mostly limited to men. Consequently, the identification of the quality of Copra as of now is checked manually by the farmers resulting in possible confusion about the quality and price of the Copra when selling to the market.

The study aims to (1) Evaluate the performance of the Convolutional Neural Network (CNN) in terms of accuracy by varying the different parameters like number of convolutional layer filters, filter size and activation function for each layer in classifying the smoked Copra based on their appearance and their cooking conditions (undercooked, perfectly cooked or overcooked); and (2) Evaluate the accuracy of the final CNN model derived from the evaluation of the first objective.

2. REVIEW OF RELATED LITERATURE

2.1 Copra Classification Standards

In the Philippines, there are four known classes of Copra based on the method of drying. They are designated as class A, B, C and D with seven grades, from 1 to7, based on moisture content. generally, pricing and trading of Copra are essentially based on moisture content. In the Philippines, where Copra is the main coconut product , with approximately 90 to 95% of total production being sold to the village trader. Copra with low moisture content (about 20-25 percent moisture content) is bought at a lower price. This concept is referred to as the "pasa system" of copra buying where a lower price of the Copra is based on moisture. Therefore, Copra is commonly classified based on its moisture content even at the first point of sale. Although there exist moisture meters, most of the time it is expensive and not readily available in the villages. Thus, moisture content determination is done manually using visual inspection or by splitting the Copra by hand and feeling. This practice requires an experienced and highly skilled person to do this [4].

2.2 Quality of Copra

Copra is one of the major products of the coconut. Usually, the drying of Copra is either kiln or direct sun drying. It is necessary to reduce the moisture content of the meat by about 50% (w.b.) to 6% (w.b.) to reduce the weight, prevent microbiological deterioration and concentrate the oil. On average, five nuts are required to produce 1 kg of Copra. However, this conversion rate varies (plus or minus 40%) from country to country [5].

2.3 Grading of Copra

Table 1 below is the qualifications for the determination of the quality of Copra - According to the Philippine National Standard [6].

Table 1: Grading of Copra

Parameters	Grade 1	Grade 2	Grade 3
Moisture Content (%)	67.9	810.9	1113.9
Oil (%, min)	62	60	58
Free Fatty Acid (oleic, % max)	0.5	4.0	5.0
Color of meat	Clean, white to pale yellow	Brown to dark brown	Brown to dark brown
Extraneous matter (% max)	0.25	0.75	1.0

2.4 Convolutional Neural Network

Way back the 1990s, fruit quality was always checked manually. Inspection is sought to be labor-intensive, slow, and can be inconsistent due to fatigue and due to the relatively large staff turnover caused by boredom. In modern fruit processing, several algorithms were employed to improve the checking process. Some methodology uses the fruit's color ratio, which was calculated with HSV color space to run the classification feature. The classic Bayes classifier, whose parameters were obtained by a study module [7][8] is usually used in this type of process.

In Image Recognition and Detection, There are various methods to make the system perform a certain task, like Classifying Images. One of the methods is the Convolutional Neural Network. CNN is an architecture of Artificial Neural Network which mimics the human brain [9]. (According to Ujjwalkarn, "Convolutional Neural Networks (CNNs) are a category of Neural Networks that have proven very effective in areas such as image recognition and classification. ConvNets have been successful in identifying faces, objects, and traffic signs apart from powering vision in robots and self-driving cars" [10].

Convolutional Networks works by moving small filters across the input image called patch, kernel or window. This means the filters are re-used throughout the entire input image for pattern detection.This makes Convolutional Networks far stronger than Fully-Connected Networks with the same number of variables. This, in turn, makes the Convolutional Networks faster to train [11].

2.5 Convolutional Neural Network Activation Functions.

Activation Functions are functions connected at the end of each convolutional layers. Their primary purpose is to convert an input signal of a node in the layer to produced modified meaningful output. It is also known as Transfer Function. It can also be attached in between two Neural Networks. On CNN, there are several activation functions that are implemented. Generally, they are categorized as linear and non-linear. Between these two, non-linear activations functions are mostly used, which is why these activations functions are also used in this study [12]. Under non-linear activation, functions are Logistic (Sigmoid), Hyperbolic Tangent (Tanh), and Rectified Linear Unit (Relu).
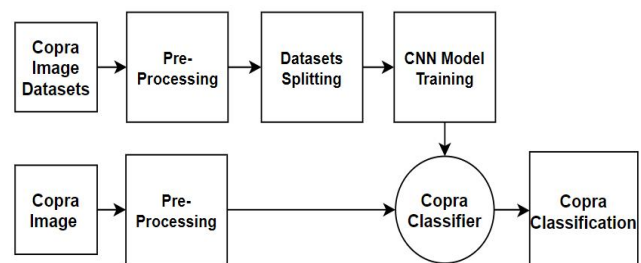

Figure 1: Copra Quality Classifier System Architecture

3. METHODOLOGY

The system mainly consists of two general sets of processes. The first set is for training the model, which consists of accepting the input images, pre-processing, data splitting, and training the CNN model. The other set is for classifying an input individual copra image. Figure 1 shows these different processes on each set.
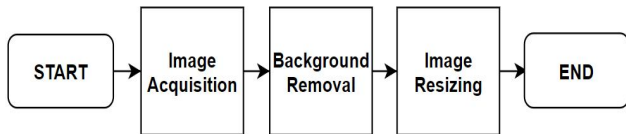
3.1 Pre-Processing

Figure 2: Pre-Processing

The pre-processing, as shown in figure 2, starts from accepting the input image and resizing it. During the same process, the properties of the image are also acquired and save for future use in the process. The background in the image is also removed. The acquired image is then resized to the desired size. In this study, a 50x50px sized was used as a parameter for resizing the image

3.2 Data Splitting

To develop the system, it is necessary to train the model and test the model. Datasets in creating models are split into a 75-25% partitions. Seventy-five prevent of the datasets will be used for training the data, and twenty-five percent will be used for validation of the model. This is necessary for the process to determine its performance.

3.3 Copra Data Sets

A total of 150 copra samples coming from the province of Quezon City, Philippines, was gathered in the development of the model. One hundred twenty of these samples were used for creating the model (training and testing) and the remaining 30 samples were used to validate the system.

To create the training datasets, Copra image is manually obtained by taking a picture of Copra with different quality using the camera of Vivo Y85. The Vivo Y85 is placed into a handle 1ft from the copra holder. White cloth is placed under each Copra to have a white background on the image (see Figure 3). The image was captured and saved on the computer. Below are sample images of the Copra.
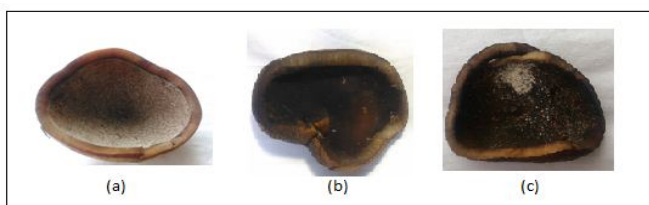

Figure 3: Image of copra datasets: (a) Undercooked Copra

3.4 Convolutional Neural Network Model (CNN)

The algorithm will follow a simple CNN model (see figure 4) consisting of 4 convolutional layers, one fully connected layer, and one output layer. Max pool filters are implemented every after the Convolutional layer; this helps the model to extract more specific features in the image. As mentioned during the design, an attempt to vary the size of the filter for each layer will be applied, which means the value of the n x m filter will be evaluated, and the best value will be used in the final design.

The output of each layer will be fed to the next layer until the classification processed is done. During the processed, imaged fed in the Convolutional layer 1 and 2 will be downsampled to a lower pixel value. The output of the Convolutional layer two will be flattened so that it will be compatible as input for the fully connected layer and an output layer. The output layer consists of a 3x1 column vector, where each cell corresponds to its classification. Classification is undercooked (UC), perfectly cooked (PF), and overcooked (OC).
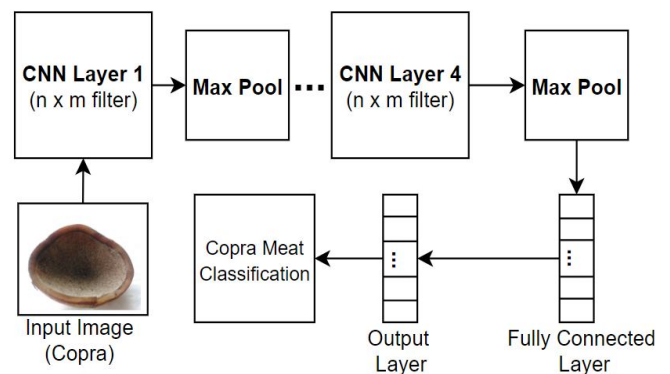

Figure 4: Copra Quality Classifier CNN Model

4. HYPERPARAMETER TUNING

4.1 Number of Filters

The tuning of the filters was made to evaluate the best parameter in developing the final model. On the first test, all Convolutional layers will be tested with a filter value of 32, 64, 128, 256, and 512 with a filter size of 5 and a max pool filter size of 5. Table 2 shows the output of varying the number of filters between Convolutional layers. The system is run on ten epochs with a snapshot step of 500.

Table 2: Result of convolutional layers parameter evaluation

Test No.	No of Filters	Loss Initial Value	Loss Final Value	Accuracy
1	32	0.65362	0.41192	0.6829
2	64	0.75397	0.51157	0.6138
3	128	3.82212	2.15215	0.2887
4	256	9.33121	6.65742	0.4844
5	512	6.06364	9.60820	0.6238

4.2 Activation Function Tuning

Based on the result of Table 2, the filter with 32 layers for each convolutional layer will result in a decreasing loss with good accuracy. Therefore, this filter size was used in the model.

To tune the activation function, the selected filter will be used to evaluate the performance of the model. Accuracy of the tuning will be tabulated. Table 3 below shows the result of evaluating different activation function for each convolutional layer.

Table 3: Activation function evaluation result

Test No.	Activation Function	Accuracy
1	Sigmoid	0.2937
2	Tanh	0.4205
3	ReLu	0.6829

4.3 Filter Size Tuning

After testing the best activation function, which is the ReLu, the size of the filter will be tuned. The system will run on a series of iteration test. Each test will have a value of n x n filter size. The result is tabulated in Table 4.

Table 4: Filter Seize evaluation result

Test No.	Filter Size (n x n)	Accuracy
1	5 x 5	0.7829
2	10 x 10	0.5926
3	20 x 20	0.5287
4	30 x 30	0.6370
5	40 x 40	0.3791

4.4 Final Model Accuracy Testing

Combining the best parameters on the result of the evaluation of activation functions, size of filters, and numbers of filters, the final model was created and tested. The 25% testing datasets of random copra images are and feed to the model. Images of the Copra classified were displayed on the system. Each Copra was labeled based on their correct or incorrect classification. Labeled with green color means correct classification and red means incorrect classification. Figure 5 shows sample images taken from the testing.
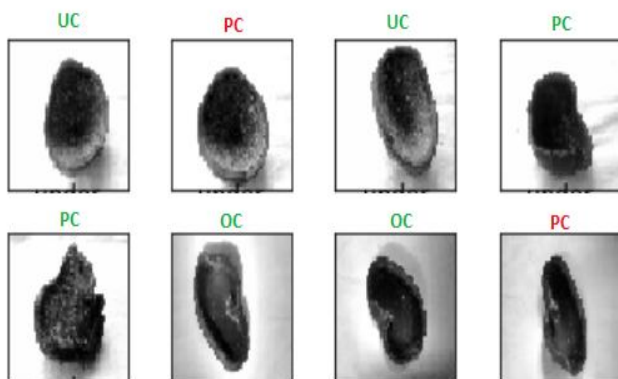

Figure 5: Sample testing output of the final model

The model correctly identifies 26 copras out of 30 samples giving an accuracy rating of 86.67%. The confusion matrix of the result, as shown in figure 6.
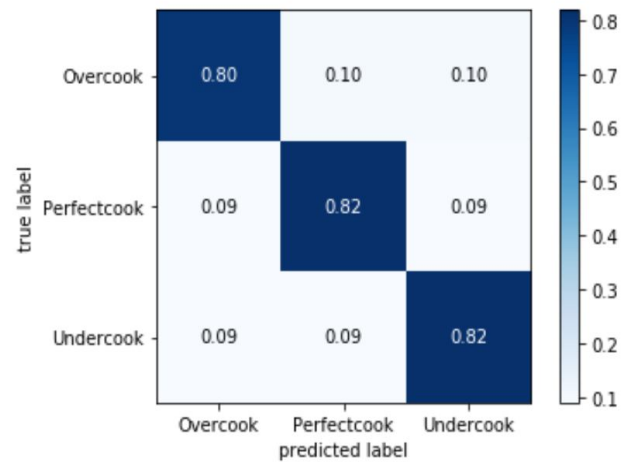
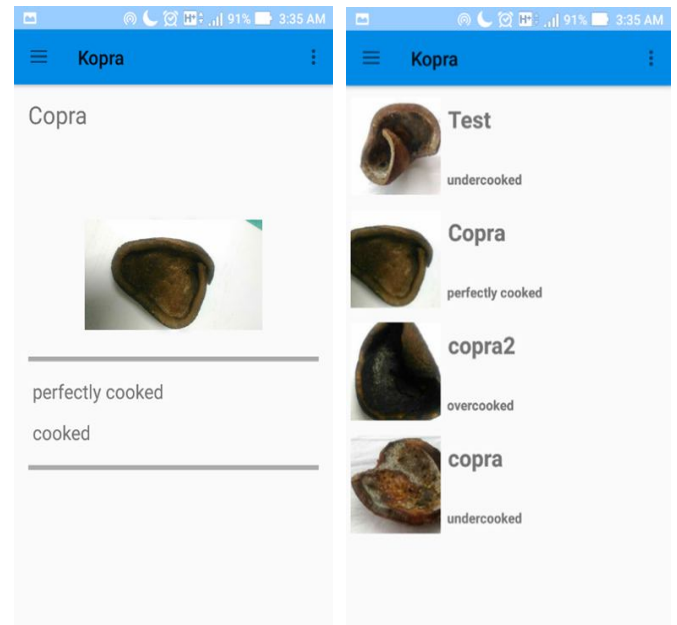

Figure 6: Final Model Confusion Matrix

Figure 7: Kopra android application running on Vivo Y85

4.5 Validation

To validate the model, a simple android application called "Kopra" was developed. The application is capable of taking the pictures of the Copra and classifying the captured copra image. The application was installed on Vivo Y85 Smart Phone. The Sample output is as shown in figure 7.

The 30 copra-validation datasets were used to validate the accuracy of the system. The application was used to capture the image of the Copra and recorded the classification result.

The confusion matrix of the classification was created to evaluate the performance of the application (see figure 8). Using the Vivo Y85, the application accurately classifies the 21 copras accurately with a f-score of 0.81212.
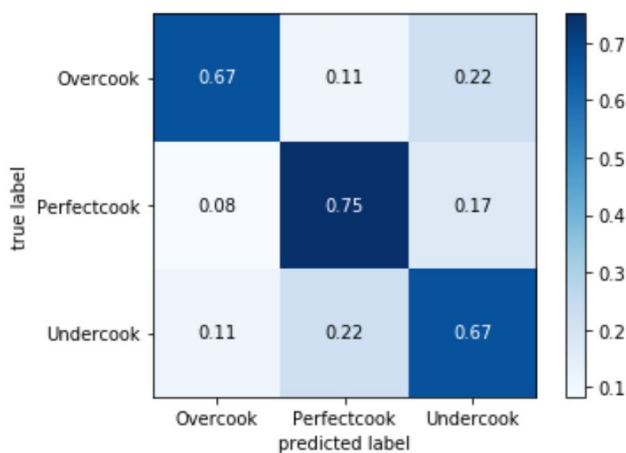

Figure 8: *Kopra* Confusion Matrix

5. SUMMARY AND CONCLUSION

With the limited number of datasets in the development of the model, a simple CNN model is evaluated with different parameters and was tested with the final model. During the testing, it is found out that it is best to use a 5x5 filter size on 32 numbers of layers of filters with an activation function of the Rectified Linear Unit (ReLu). While not officially written as a result of the testing, it was observed that increasing the filter size on filter layers results in longer training the model. Combining all the best parameters, the final CNN model for Copra Classification obtained an 86.67% accuracy, which is good since it does not result in overfitting the training on a small number of datasets. The model was also implemented using the android application and validated. The application was called *kopra* and has a validating f-score of 0.81212.

With the Experimental process on the hyperparameters of the CNN, a model was created that can classify the grade of the Copra. The model can also be used in an android application to be used for classification purposes.

The model implemented in the android application gives more opportunities for the development of other application that supports copra classification that would aid copra producers.

6. RECOMMENDATIONS

CNN model works best on a large number of sample images. In this study, 150 in total images seems small, and it can be further improved by adding more copra images. Possible testing on having a different number of filter layers would be a good improvement in the study.

For the android application, it could be improved by adding pricing determination for each grade of the Copra. With this, it will lessen the confusion Copra pricing by the seller, and of course, the buyer.

REFERENCES

1. Subramaniyam, S. **Performance Analysis on Diesel Engine using Neem and Soya Bean Oil,** International Journal of Emerging Technologies in Engineering Research (IJETER), Vol. 5, Issue 8, August 2017.
2. J deepa, p rajkumar & t arumuganathan. **Quality analysis of Copra dried at different drying air temperatures**, August 2015, 1-88.
3. P. Punchihewa and R. Arancon. **COCONUT: Post-harvest Operations**, 1999.
4. Philippine Coconut Authority. October 2015. [Online]. Available: **http://pca.da.gov.ph/index.php/2015-10-26-03-15-57/ 2015-10-26-03-22-41#area**.
5. G. Patterson and P. Perez. **Solar Drying in the Tropics**, p. 13, 1981.
6. Philippine Coconut Authority (PCA). **Revised Rules and Regulations to Implement New Copra Classification Standards,** 1991.
7. G. Feng and C. Qixin. **Study on Color Image Processing Based Intelligent Fruit Sorting System**, 2004.
8. Balika J. Chelliah, et al. **Classification of Mushrooms using Supervised Learning Models**. *International Journal of Emerging Technologies in Engineering Research (IJETER),* Vol. 6, Issue 4, April 2018.
9. Aaron Don M. Africa. **ANN Distance Protection for Transmission Lines,** *International Journal of Emerging Trends in Engineering Research (IJETER),* Vol. 7, No. 8 August 2019. https://doi.org/10.30534/ijeter/2019/07782019
10. Ujjwalkarn. **An Intuitive Explanation of Convolutional Neural Networks**, 2016.
11. Pedersen m.e.h. C**onvolutional neural network, (2017), github repository, https://github.com/hvass-labs/tensorflow-tutorials/blo b/master/02_convolutional_neural_network.ipynb**
12. Sharma s. A**ctivation functions: neural networks, [retrieved from https://towardsdatascience.com**, June 2019].