# Discrimination of civet coffee using near infrared spectroscopy and artificial neural network

### Edwin R. Arboleda\*

Assistant Professor, Department of Computer and Electronics Engineering, College of Engineering and Information Technology, Cavite State University, Indang, Cavite, Philippines

Received: 06-September-2018; Revised: 03-November-2018; Accepted: 15-November-2018 ©2018 ACCENTS

### Abstract

Civet coffee is regarded as the most expensive coffee in the world. Owing to its high price the conventional method of harvesting civet coffee in the wild has been replaced by a farming method, wherein the wild civet cats are being captured, caged, and then force-fed with ripe, hand-picked, coffee cherries. This is contrary to the production of wild civet coffee from the ripest and sweetest coffee cherries, carefully picked by the civet cat. As the wild civet coffee is quite hard to obtain, most of the civet coffee commercially sold as authentic civet coffee is actually from a caged civet cat. Traders and consumers have no way of differentiating the civet coffee from other types of coffee. This study aimed to differentiate caged civet coffee beans from ordinary green coffee beans. The technique used was the near infrared spectroscopy (NIRS) as it is non- destructive and can generate quick results. A total of 218 samples were scanned, which generated absorbance in the entire 780 wavelength capacity of the Cavite State University Indium, Gallium and Arsenic (CvSU InGaAs)-based NIR instrument, ranging from 904 to 1684 nanometres (nm). A total of 170,040 spectra were generated. The average spectral absorbance's having major differences between the two groups, were chosen, which are 907nm, 1088 nm, 1540 nm and 1650 nm, respectively. Out of 218 samples, 130 samples were used as training data, 40 samples as testing data, and the remaining 48 samples for validation purposes. The training data were subjected to the 4 layers, 15 neurons feed forward back propagation artificial neural network (FFBPANN) for training. Classification scores of 95% to 100% were achieved. Using the combined NIRS and FFBPANN, the civet coffee can be successfully discriminated from coffee beans not eaten by a civet.

# **Keywords**

Civet coffee, Near infrared spectroscopy, Artificial neural network, Absorbance, Classification learner app.

# **1.Introduction**

Civet coffee refers to the coffee that includes partlydigested coffee cherries eaten and defecated by the palm civet cat (Paradoxurus hermaphrodites). Civet cats seek out and eat only the ripest, reddish coffee cherries. The coffee cherry fruit is sweet and is completely digested, but the beans are excreted in their feces. It is one of the rarest and priciest coffees in the world [1]. At present, there is still no accepted standard in determining the authenticity of the civet coffee [2]. Due to its high price, the traditional method of collecting feces from the wild civets has given way to a farming method where civets are kept in cages and force-fed with the cherries. Most of the previous studies done in discriminating civet coffee from ordinary coffee beans used a destructive method of sampling, preparation wherein coffee beans were grounded and results can take a very long time. The forte of the NIRS technique lies in its speed [3].

NIRS is an analytical technique that uses no chemicals, gives accurate and precise results in minutes or even continuously, and is simple to install, and is safe to use. It is a rapid method for qualitative and quantitative analysis of a very wide range of materials, powders, slurries, and even solid materials and gases. A "rapid" method is defined as a method that will provide accurate and precise results in two minutes or less, including the time for sample preparation, and accessing the sample too, and removing it from the instrument. The NIRS spectrum analysis does not require samples to be grounded because when the light passes through it, light can reveal a certain spectra of the characteristics or features that are unique to the class of that sample.

Artificial neural network (ANN) is a computational system that employs parallel processing based on connections and distributional methodology. ANN is similar in operations to a biological neural network [4,5]. It can be trained to learn and once trained it can be used as a very good classifier. Just like the

<sup>\*</sup>Author for correspondence

human brain, after its exposure to different classes of things or objects and studying its properties eventually it will learn to associate certain properties to a particular class and can successfully discriminate that particular class from other classes.

The civet coffee has been subjected to numerous discrimination and characterization studies which all aimed to describe its unique properties. It was discovered in [6] that the sensory and microbial characteristics of both wild and caged civet coffee are similar in profile. Results showed that the cupping score of wild civet coffee is slightly higher than the caged civet coffee. Wild civet coffee got an excellent specialty classification and the caged civet got a very good specialty classification on the scoring key that they used. Overall they concluded using ttest analysis, that the cupping scores of wild and caged civet coffee are similar. In terms of lactic acid bacteria count of civet feces, the caged civet has higher colony count that the wild civet. The lactic acid count could be affected by life condition and the diet taken by the civet. Based on this study, although the lactic acid colony differs in wild civet and caged civet coffee the taste of the two coffee groups are similar, this is according to the expert that have done the sensory evaluation of the coffee taste. In [7] civet coffee was discriminated from two other Indonesia's specialty coffee the peaberry and the pagar alam using the fluorescence spectroscopy, principal component analysis (PCA) and the soft independent modelling of class analogies(SIMCA) method. According to this study, the civet coffee has its unique characteristics distinguishable from other specialty coffee of Indonesia.

Although the fluorescence spectroscopy successfully discriminated the three specialty coffee groups, the major drawback of this method is that the coffee beans are needed to be ground and liquid samples are to be extracted and pipette into a cuvette for fluorescence spectral acquisition. This is clearly different from the NIRS method in which the whole coffee beans are used and spectra data can be acquired without destroying the beans by grinding. In [8] their research discriminated the civet coffee from non-civet coffee using UV-visible spectroscopy and the SIMCA and partial least squarediscrimination analysis (PLS-DA). They successfully discriminated with 100% accuracy the civet coffee from non-civet coffee by using ground coffee beans as samples. In [9] the civet ground coffee was discriminated from non-civet coffee using the electronic tongue, the multi-dimensional scaling (MDS), and the PCA. In [10] gas chromatography/ mass spectroscopy was applied in combination of

multivariate analysis to discriminate ground civet coffee from non-civet coffee. According to [2] the gas chromatography-mass spectroscopy (GCMS) chromatographic profile of civet coffee is almost similar to non-civet coffee but of different concentrations. In [1, 2] using the e-nose, aroma of civet coffee is different from the non-civet coffee. The major drawback of mentioned researches [1,2,6-10] is that they require complicated and time consuming sample preparation such as grinding the samples. Some require the use of chemicals. The coffee beans are needed to be ground before a feature can be extracted. In this study the samples were not destroyed by grinding, no chemicals were used; features were extracted in several seconds and subjected to a trained artificial neural network where results were achieved in seconds also.

Different studies proved that civet and non-civet coffee differs in aroma [1, 2], taste [9] and other distinguishable properties [7, 8, 10]. According to [6] wild civet coffee and caged civet coffee has almost the same taste. NIRS has been the accepted technology for rapid, accurate, non-destructive analysis of food quality and properties [11]. This has been used as a tool for extracting spectra that reveal attributes and properties for classification of vegetables [12,13], nuts [14,15], fruits [16–19], meat [20–22], grains [23–25] and other agricultural products. Many studies have been published about the utilization of NIRS in the classification of coffee properties. It has also been used to determine the compositions of coffee such as caffeine [23, 24, 26], the obromine and theophylline [27], water content [28] and moisture content [29]. According to these studies, coffee properties and features produce specific spectra that can be used to differentiate one group from another.

Feed forward back propagation artificial neural network (FFBPANN) is a powerful classifier and has been used in recent studies with high percentage of classification accuracy. In agriculture, FFBPANN has been used in the classification of coffee species into excelsa, robusta and liberica with a 96.67% accuracy [30]. It classified mature from non-mature apple fruit with 98.1% accuracy [31]. FFBPANN is also used in music, wherein it classified melakartna ragas with 96.4% accuracy [32]. In the field of informatics FFBPANN classified email spam with 91% accuracy [33]. In medicine, it has a 87.5 to 97.5% accuracy into classifying future osteoporotic fracture risks for low mass group, post-menopausal and old aged women [34]. It was also used in classifying damaged and undamaged tablets with 94.4% accuracy [35]. It successfully classified lesion

in mammogram with 96% accuracy [36]; renal calculus disease with 96.88% accuracy [37]; lung columnar cells, 92.9% [38]; major lung diseases, 92% [39] and brain magnetic resonance imaging with accuracy of 96.67% [40].

This study was conducted to determine if passage of coffee beans in the digestive tract of civet cat can alter its properties, and whether this alterations can be detected by NIRS and be discriminated by ANN. It is the also the aim of the study to detect the spectra equivalent to the unique properties of the caged civet coffee that makes it different from the non-civet coffee.

# 2.Materials and methods

### **2.1Conceptual framework**

The conceptual framework of the study is shown in Figure 1. SA and SB are the sample names. The first part is the use of NIRS to acquire absorbance values in a chosen wavelength of 85 single-bean samples civet and 85 single-bean samples of non-civet coffee. The second part is the classification of absorbance values by FFBPANN. This is done by training the network to classify that a particular absorbance value belongs to either civet or non-civet coffee. A total of 65 samples of civet coffee, samples SA1.1 to SA1.65 and another 65 samples of non-civet coffee, samples SB1.66 to SB1.131 comprises the training samples of the FFBPNN. After training the network multiple times an acceptable accuracy was achieved. Test samples are inputted into the network to determine if it can precisely classify the civet coffee and the non-civet coffee. Test samples, SA2.1 to SA2.20 for civet coffee and SB2.1 to SB2.20 for non-civet coffee is input into the network to determine if it can precisely classify the civet coffee and the non-civet coffee.

### **2.2Civet coffee samples**

Based on [6], civet coffee and non-civet coffee cannot be differentiated in terms of taste and due to the scarcity of civet coffee beans in the wild; this study used the caged civet coffee. The samples shown in *Figure 2* were from the coffee season of January to March 2018. Handpicked red Robusta

coffee beans were divided into two different groups. The first group consisted of Robusta beans offered to a caged civet cat and the second group were the control group or the non-civet coffee. Civet coffee beans were harvested in the morning and were kept in a plastic bag with zip lock and were placed in a freezer. For the non-civet coffee beans, the outer skin layers (epicarp) were stripped but the pulp layers underneath it (mesocarp) were not removed. This is because when the civet cat eats the coffee beans, only the epicarp is digested, but the mesocarp remains as shown in Figure 2. After removing the outer skin from the seed, the beans were also kept in a plastic bag with zip lock before putting in the freezer. After the collection period, the two groups of samples were taken out of the freezer and were washed in running water, then air dried and suns dried for a period of 8 days and were kept in dried plastic containers.

The conditions between the two groups were made the same in terms of storage time in the freezer, wash time, air drying time and the sun drying time, so that the moisture content between the two groups would remain the same. The recommended sun drying time for washed coffee is six to seven days to achieve a moisture of 15% [41]. In this study, the sun drying time was made eight days for the two groups of sample to attain an estimated moisture value of 9 to 12% or an average of 10.8% which is the natural moisture range for green coffee beans [42].

The conditions between the two groups were made the same in terms of storage time in the freezer, wash time, air drying time and the sun drying time so that moisture content between the two groups would remain the same. The recommended sun drying time for washed coffee is six to seven days to achieve a moisture of 15% [41]. In this study, the sun drying time was made eight days for the two groups of sample to attain an estimated moisture value of 9 to 12 % or an average of 10.8% which is the natural moisture range for green coffee beans [42].



Figure 1 Conceptual framework of the study



Figure 2 Civet coffee vs. non-civet coffee

### 2.3Collection of NIR single-bean spectra

The NIRS instrument used in this study is the CvSU InGaAs-based NIR instrument. This instrument is a patent pending, NIRS instrument developed by engineers, scientists and professors from Cavite State University that is based on Indium Gallium and Arsenic (InGaAs) NIRS sensor. It can obtain the absorbance data in the range of 950 to 1684 nm. The technique used in this study is the single seed NIR spectra acquisition as described in [43], where a single bean is placed in the scanner of CvSU In GaAs-based NIR instrument. The instrument was allowed to warm-up for a period of at least 30 minutes so that the light output and electronics related to spectrometer would stabilize prior to spectra collection. After warming up the instrument, the spectral precision for the two groups of sample was determined. This was done by scanning ten different samples (re-loaded group) of civet coffee and the tenth sample scanned for another ten times (non-reloaded group).

Then spectral data at two wavelength points, (e.g. 1210 and 1630 nm) were recorded for all the 20 scans and the standard deviation (SD) and coefficient of variation(CV) for both reloaded and non-reloaded sets of ten scans for the civet coffee and the non-civet coffee were calculated. The standard deviation of the reloaded set is the spectral precision that will be achieved in the day to day work. The CV should be no higher than 4%. The SD and CV were also calculated for the reloaded set. This is the error of the instrument itself, because there should be no operator error. The SD and CV of the non-reloaded set of 10 scans should be much lower.

After calibrating the NIR instrument, the total samples composed of the civet coffee and the noncivet coffee was subjected to scanning to the CvSU InGaAs-based NIR instrument in its entire

wavelength capacity of 904 to 1684 nm. From these scanned values, the average spectra of the civet coffee and non-civet coffee were placed in a graph. Also, the spectral average of the two groups was tabulated in Microsoft Excel to determine their differences. Based on the visual difference in the graph and the computed differences in MS Excel, four spectra values were chosen. The differences in values of spectra between the two groups were also inspected and evaluated to find out if their differences are not due to their water content. There are spectral values in coffee that are indications of water differences. The biggest differences that occur in the areas of 960, 1155, 1420 and 1930 nm indicate differences in water [44].



Figure 3 Structure of the feed forward back propagation artificial neural network

# 2.4The feed forward back propagation artificial neural network

After the appropriate spectral values were chosen, were subjected through analysis thev for classification using feed forward back propagation artificial neural network. A total of 65 samples per group were subjected to multiple training until an acceptable overall accuracy of above 99% was met. Testing samples were used to evaluate whether the network can classify which samples belong to their respective groups. The output of the ANN is the number of correct classifications. The artificial neural network is inspired on how the neurons work in a biological brain. Its structure is shown in Figure 3. It consists of a large network of interconnected neurons, which have the ability to gain knowledge through training. It can effectively use the knowledge acquired through training into associating to solution patterns in which it was trained for. An example on how the principle in biological brain works and is applied in ANN is the ability of the K9 bomb sniffing dogs. As a result of the numerous times of training the dog's brain neurons can associate certain scents of compounds used as bomb components. The pattern of solving, whether there is a bomb in a container is already created and stored in the dog's brain.

A FFBPANN is the first and simplest artificial neural network devised. The network is termed as feed forward because the information moves in a forward direction and there are no loops from the input layer to the hidden layer and then to the output layer. In back propagation, the output values are compared to the correct answer to compute the value of some

predefined error function. The error is fed back through the network using various techniques. As error is fed back through the network, the weights are given to each connection; the algorithm adjusts the weights of each connection with the objective of reducing the error's value in small amount in every training cycle. Nonlinear optimization, called gradient descent, is applied to adjust the error properly in such a way that he error decreases. The network calculates the derivative of the error function with respect to network weights such that the error decreases. Repetition of reducing the error in a large number of training cycles would result to a state of convergence where the error is very small. Upon reaching this state, the network is called a learned network [45].

# **3.Results**

Implementation of the proposed methodology was carried out using Intel core i7 with 4 GB RAM and 3.5GHz processor. The FFBPANN algorithm was implemented in MATLAB R2017b. A total of 218 samples of single bean civet and non-civet coffee were scanned in the CvSU IGaAs based NIR equipment. A total of 170,040 absorbance values were produced from a wavelength of 904 to1604 nm. Range of absorbance values of the chosen spectra is shown in *Table 1*. It can be seen in *Table 1* that the values of absorbance for 907, 1088, 1540 and 1658 nm are overlapping. The average absorbance for the two groups is displayed in *Figure 4*. It can be seen in Figure 4 that civet coffee beans have higher values of average absorbance compared to non-civet coffee for

the entire wavelength of 904 to 1604 nm. *Figure 5* shows the difference value in the wavelengths 960, 1155 and 1420nm which were evaluated to make sure that the difference between the two groups is not because of water contents. Further, the difference in the spectra between two groups does not peak at 960, 1155 and 1420 nm which means that the difference in absorbance between two groups is not due to water. From *Figure 5*, the four wavelengths where there is a large difference in absorbance were chosen. These are the 907, 1088, 1540 and 1658nm.

International Journal of Advanced Computer Research, Vol 8(39)

A total of 130 training samples, 40 testing samples and 48 validation samples were used in this study and were subjected to FFBBANN. They were trained in the network shown in *Figure 6*, which has a 4 input, 2 hidden layer and 10utput layer until an overall accuracy of 99.98% were reached as shown in *Figure 7*.

The trained ANN network was validated ten times using different set combinations of 20 samples per validation from 48 validation samples and an accuracy ranging from 95 to 100 % was achieved.

 Table 1 Range of values of absorbance for the chosen wavelength

Coffee type	907nm	1088nm	1540nm	1658nm	
Civet Coffee	0.7501 to 1.2103	0.80201 to 1.17116	0.934344 to 1.254289	0.91242 to 1.26597	
Non Civet Coffee	0.7130 to 1.0349	0.77854687 to 1.04549289	0.937738 to 1.146626	0.904525 to 1.14468226	



Wavelength

Figure 4 Average spectra of civet coffee vs. non-civet coffee



Wavelength in nm

Figure 5 Difference in absorbance between civet coffees vs. non-civet coffee



Figure 6 FFBBANN used in this study



Figure 7 Overall accuracy of 99.98 % was reached during training

### **4.Discussion**

# 4.1Comparison of the proposed method with the previous studies

Table 2 is a summary of past studies, discriminating civet coffee from non-civet coffee using different technologies and analysis techniques. Free range

civet coffee and caged civet coffee have similar aroma and flavoured characteristics [6]. However, other techniques stated that civet and non-civet coffee are different from each other in terms of properties and features. Using an electronic nose and GCMS, it was found that civet and non-civet coffee differ in aroma and volatiles [2]; using electronic tongue, they differ in sensory attributes [9], and using Gas chromatography with flame ionization detector (GC/FID), they differ in metabolites [10]. Using other spectroscopy techniques, the ultraviolet-visible spectroscopy and fluorescence spectroscopy, both yielded the same results that civet and non-civet coffee can be differentiated by statistically analyzing their spectra [7, 8]. Although these studies differ in terms of instruments and the algorithm used, their results corroborate the result of this study, that civet and non-civet coffee are different. The only difference is that the samples used in this study were not destroyed by grinding and all physical attributes before and after the data were gathered from the samples were retained. In the context of commercial application, samples not destroyed when data are extracted for testing are very advantageous because samples are not wasted and can still be sold.

Table 2 Comparison of past studies differentiating civet coffee from non-civet coffee

<b>Tuble 2</b> comparison of past studies unterentiating erver contee from non erver contee							
Author (s)	Techno	logy	Algorithm		Sample preparation : Destructive or non- destructive	Findings	
Ongo et al. [2]	Electror	nic nose (E-Nose)	Principal analysis (P	component CA)	Destructive	Civet and n differ in	on-civet coffee aroma and
	Gas mass	chromatography spectroscopy	Cluster anal	ysis		volatiles.	

International Journal of Advanced Computer Research, Vol 8(39)

Author (s)	Technology	Algorithm	Sample preparation : Destructive or non- destructive	Findings
	(GMCS)			
Muzaifa et al. [6]	Cupping quality was performed by a group of three certified judges	Paired T-test	Destructive	Free range civet coffee and caged civet coffee have similar aroma and flavour characteristics.
Suhandy and Yulia [7]	Fluorescence spectroscopy	PCA	Destructive	Civet and non-civet coffee differ in sensitivity and specificity values of fluorescence spectra.
Yulia and Suhandy [8]	Ultraviolet-visible spectroscopy	PCA and partial least square (PLS)	Destructive	Civet and non-civet coffee differ in sensitivity and specificity values of chosen ultraviolet-visible spectra.
Lopetcharat et al. [9]	Electronic tongue	PCA	Destructive	Civet and non-civet coffee differ in sensory attributes.
Jumhawan et al. [10]	GC/FID	Multivariate Analysis	Destructive	Civet and non-civet coffee differ in metabolites.
This study	NIRS	FFBPANN	Non-Destructive	Civet and non-civet coffee differ in chosen near infrared spectra.

# 4.2Comparison of the FFBPANN classifier with other machine learning algorithms

The classification performance of the FFBPANN was compared to other machine learning algorithms to determine if FFBNN is the best method. Classification Learner App in MATLAB 2017b was used to train the NIRS absorbance's of civet coffee and non-civet coffee using the same parameter used in training the ANN. In ANN, 65 samples each were used for training and 20 samples each were used for testing and were validated ten times.

To do the same training and testing in Classification Learner App, 85 samples were used for each group. The four columns containing the absorbances of each bean group in the 907 nm, 1088 nm, 1540 nm and 1650 nm were the input. The fifth column contains the response, whether it is 1 for civet coffee and 2 for non-civet coffee. In the response column there are eighty-five 1's and eighty five 2's. All the 22 machine learning algorithms in the Classification Learner App were selected and the 24-fold cross validation was used to set the training and testing data for the model. The 24-fold cross validation method was used to set 24% data for testing and 76% data for training. That is, 24% of 85 is 20.4 or 20 samples for testing and 76% of 85 is 64.6 or 65 samples of the training. The 24-fold cross validation also means that the training and testing process are repeated for 24 times. The results of the classification using the 22 machine learning algorithms are shown in *Table 3*.

It can be seen from the data in *Table3* that the highest classification accuracy is 86.5%, which was produced by quadratic support vector machine (SVM) and fine Gaussian SVM, followed by cubic SVM with 85.9% and the third in rank is the fine k-nearest neighbor (KNN) with 85.3%.

Classifier	Classifier type	Accuracy (%)	Prediction	Training
			speed (Objects/Seconds)	time
				(Seconds)
	Fine Tree	81.2	760	72.27
Decision Trees	Medium Tree	81.2	1000	65.93
	Coarse Tree	80	1300	83.26
	Linear Discriminant	65.3	1300	83.26
Discriminant	Quadratic Discriminant	64.1	110	44.89
Analysis	(Using Diagonal Covariance)			
SVM	Linear SVM	62.9	730	81.255
	Quadratic SVM	86.5	840	79.33
	Cubic SVM	85.9	920	101.7

Table 3 Classification results for each classifier using the classification learner app without using the PCA

Classifier         Classifier type         Accuracy (%)         Prediction speed (Objects/Seconds)         Training time (Seconds)           Fine Gaussian SVM         86.5         680         79.12           Medium Gaussian SVM         77.6         1100         78.85           Coarse Gaussian SVM         68.2         1200         78.547           KNN         Fine KNN         85.3         1500         78.04           Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91	Arboleda				
KNN         Fine Gaussian SVM         86.5         680         79.12           Medium Gaussian SVM         77.6         1100         78.85           Coarse Gaussian SVM         68.2         1200         78.547           KNN         Fine KNN         85.3         1500         78.04           Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91	Classifier	Classifier type	Accuracy (%)	Prediction speed (Objects/Seconds)	Training time
Fine Gaussian SVM         86.5         680         79.12           Medium Gaussian SVM         77.6         1100         78.85           Coarse Gaussian SVM         68.2         1200         78.547           KNN         Fine KNN         85.3         1500         78.04           Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91					(Seconds)
Medium Gaussian SVM         77.6         1100         78.85           Coarse Gaussian SVM         68.2         1200         78.547           KNN         Fine KNN         85.3         1500         78.04           Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Fine Gaussian SVM	86.5	680	79.12
Coarse Gaussian SVM         68.2         1200         78.547           KNN         Fine KNN         85.3         1500         78.04           Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Medium Gaussian SVM	77.6	1100	78.85
KNN         Fine KNN         85.3         1500         78.04           Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Coarse Gaussian SVM	68.2	1200	78.547
Medium KNN         80.6         1400         77.0           Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91	KNN	Fine KNN	85.3	1500	78.04
Coarse KNN         50         2100         76.89           Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Medium KNN	80.6	1400	77.0
Cosine KNN         70.6         1500         76.79           Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Coarse KNN	50	2100	76.89
Cubic KNN         82.4         1800         76.37           Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Cosine KNN	70.6	1500	76.79
Weighted KNN         84.1         2300         76.24           Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Cubic KNN	82.4	1800	76.37
Ensemble classifiers         Boosted Trees         82.4         350         75.99           Bagged Trees         84.1         310         73.9           Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Weighted KNN	84.1	2300	76.24
Bagged Trees84.131073.9Subspace Discriminant55.315073.5Subspace KNN82.915072.2RUS Boosted Trees64.129070.91	Ensemble classifiers	Boosted Trees	82.4	350	75.99
Subspace Discriminant         55.3         150         73.5           Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Bagged Trees	84.1	310	73.9
Subspace KNN         82.9         150         72.2           RUS Boosted Trees         64.1         290         70.91		Subspace Discriminant	55.3	150	73.5
RUS Boosted Trees         64.1         290         70.91		Subspace KNN	82.9	150	72.2
		RUS Boosted Trees	64.1	290	70.91

The training time for FFBPANN is dependent on the human trainer. The human trainer can extend the training time if higher accuracy is desired. The data in this study were trained for 300 seconds; it was stopped when the overall accuracy of 99.984% was reached. Validating the trained FFBPANN yields classification accuracy in the range of 95 to 100%. This is above any classification accuracy produced by 22 machine learning algorithms. The prediction speed of the trained FFBPANN network is 60 seconds. Comparing it to the different machine learning algorithms in Table 3, only the ensemble classifiers can match the speed of the FFBPANN in terms of training and prediction speed. Overall, in terms of classification accuracy and in terms of speed, the FFBPANN is the best classifier in discriminating the civet coffee from non-civet coffee using the NIRS absorbances.

# **5.**Conclusion and recommendations

This paper introduces a combined NIRS and ANN approach for civet coffee classification. Experimental results using ANN indicate that the approach is workable in classifying civet from noncivet coffee with a degree of accuracy from 95 to 100%. The results can be interpreted that there are different components in civet coffee that makes it different from non-civet coffee. This finding supports the previous work of other researchers [2, 7-10], that civet coffee is indeed different from non-civet coffee. This study proves that the fermentation in the gastric juices of the civet cat changes the properties of the coffee beans and that change in properties is observable using the NIRS approach. The advantage of the NIRS approach is that the samples are not destroyed by grinding and spectra can be generated in a quick manner. In this study, it was proven that the absorbance spectra values of 907nm, 1088nm, 1540nm and 1658nm can be used to discriminate civet coffee from non-civet coffee beans and their differences can be classified by the ANN. It can also be concluded that the most suitable classifier to use is the FFBPANN over other machine learning algorithms.

It is recommended that the training, test and validation samples be increased in number and tested multiple times to further validate the results in this study. For future work, the results gathered in this study can be used as a basis in developing a portable instrument that uses combined NIRS sensor and ANN for the authentication of civet coffee. It is recommended that the spectra of caged civet coffee and authentic free range civet coffee be compared using NIRS and ANN in the future. Another recommendation is that further studies be made in authentication of the caged civet coffee so that there will be a way to determine whether in bean form, roasted form or in ground form that it is a 100% caged civet coffee and not as claimed by businessmen as authentic free range civet coffee. In this way, production of caged civet coffee will be eliminated with proper legislation so that civet cats do not have to suffer in cages.

# Acknowledgment

The author would like to thank the faculty members, employees, students and administrators of Cavite State University especially Dr. Hernando Robles, President; Dr. Camilo A. Polinga, VPAA; Dr. Ruel Monica, VP-RECETS; Dr. Ma. Cynthia Dela Cruz, current Director of Research Center; Dr. Hosea Matel, former Director of Research Center and Dr. Marilyn M. Escobar, Dean, CEIT, for the support extended to him in the conduct of the study. This study is financially supported by the CvSU Faculty Research Grant. Also he would like to thank the Department of Science and Technology-Philippine Council for Industry, Energy and Emerging Technology Research and Development (DOST-PCIEERD) for financially International Journal of Advanced Computer Research, Vol 8(39)

supporting his training abroad on the Near Infrared Spectroscopy under the BCDA scholarship program in 2016. He would like to thank Phil Williams and Dianne Malley of PDK Grains in Nanaimo, Vancouver, Canada for teaching him the science of NIRS and for answering his questions while he is conducting this study. The author also would like to thank Prof. Agnes Francisco and Prof. Joether Francisco for the help extended to him in proofreading the manuscript.

#### **Conflicts of interest**

The authors has no conflicts of interest to declare.

#### References

- [1] Marcone MF. Composition and properties of Indonesian palm civet coffee (Kopi Luwak) and Ethiopian civet coffee. Food Research International. 2004; 37(9):901-12.
- [2] Ongo E, Falasconi M, Sberveglieri G, Antonelli A, Montevecchi G, Sberveglieri V, et al. Chemometric discrimination of Philippine civet coffee using electronic nose and gas chromatography mass spectrometry. Procedia Engineering. 2012; 47:977-80.
- [3] Williams P. A practical introduction to cereal chemistry. First Choice Books; 2011.
- [4] Kuncheva LI. Combining pattern classifiers: methods and algorithms. John Wiley & Sons; 2004.
- [5] Yegnanarayana B. Artificial neural networks. PHI Learning Pvt. Ltd.; 2009.
- [6] Muzaifa M, Hasni D, Patria A, Abubakar A. Sensory and microbial characteristics of civet coffee. International Journal on Advanced Science, Engineering and Information Technology. 2018; 8(1):165-71.
- [7] Suhandy D, Yulia M. Discrimination of several Indonesian specialty coffees using fluorescence spectroscopy combined with SIMCA method. In IOP conference series: materials science and engineering 2018 (pp. 1-7). IOP Publishing.
- [8] Yulia M, Suhandy D. Indonesian palm civet coffee discrimination using UV-visible spectroscopy and several chemometrics methods. In journal of physics: conference series 2017 (pp.1-6). IOP Publishing.
- [9] Lopetcharat K, Kulapichitr F, Suppavorasatit I, Chodjarusawad T, Phatthara-aneksin A, Pratontep S, et al. Relationship between overall difference decision and electronic tongue: discrimination of civet coffee. Journal of Food Engineering. 2016; 180:60-8.
- [10] Jumhawan U, Putri SP, Bamba T, Fukusaki E. Application of gas chromatography/flame ionization detector-based metabolite fingerprinting for authentication of Asian palm civet coffee (Kopi Luwak). Journal of Bioscience and Bioengineering. 2015; 120(5):555-61.
- [11] Williams P, Norris K. Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists, Inc.; 1987.
- [12] Sendin K, Manley M, Williams PJ. Classification of white maize defects with multispectral imaging. Food Chemistry. 2018; 243:311-8.
- [13] Islam MN, Nielsen G, Stærke S, Kjær A, Jørgensen B, Edelenbos M. Novel non-destructive quality

assessment techniques of onion bulbs: a comparative study. Journal of Food Science and Technology. 2018; 55(8):3314-24.

- [14] Hu J, Ma X, Liu L, Wu Y, Ouyang J. Rapid evaluation of the quality of chestnuts using near-infrared reflectance spectroscopy. Food Chemistry. 2017; 231:141-7.
- [15] Canneddu G, Junior LC, De Almeida Teixeira GH. Quality evaluation of shelled and unshelled macadamia nuts by means of near-infrared spectroscopy (NIR). Journal of Food Science. 2016; 81(7):1613-21.
- [16] Arendse E, Fawole OA, Magwaza LS, Nieuwoudt H, Opara UL. Evaluation of biochemical markers associated with the development of husk scald and the use of diffuse reflectance NIR spectroscopy to predict husk scald in pomegranate fruit. Scientia Horticulturae. 2018; 232:240-9.
- [17] Moscetti R, Raponi F, Ferri S, Colantoni A, Monarca D, Massantini R. Real-time monitoring of organic apple (var. Gala) during hot-air drying using near-infrared spectroscopy. Journal of Food Engineering. 2018; 222:139-50.
- [18] Xiao H, Li A, Li M, Sun Y, Tu K, Wang S, et al. Quality assessment and discrimination of intact white and red grapes from Vitis vinifera L. at five ripening stages by visible and near-infrared spectroscopy. Scientia Horticulturae. 2018; 233:99-107.
- [19] Song W, Wang H, Maguire P, Nibouche O. Differentiation of organic and non-organic apples using near infrared reflectance spectroscopy-a pattern recognition approach. In SENSORS, 2016 (pp. 1-3). IEEE.
- [20] Yang Y, Zhuang H, Yoon SC, Wang W, Jiang H, Jia B. Rapid classification of intact chicken breast fillets by predicting principal component score of quality traits with visible/near-infrared spectroscopy. Food Chemistry. 2018; 244:184-9.
- [21] Panagou EZ, Papadopoulou O, Carstensen JM, Nychas GJ. Potential of multispectral imaging technology for rapid and non-destructive determination of the microbiological quality of beef filets during aerobic storage. International Journal of Food Microbiology. 2014; 174:1-11.
- [22] Balage JM, E Silva SD, Gomide CA, De Nadai Bonin M, Figueira AC. Predicting pork quality using Vis/NIR spectroscopy. Meat Science. 2015; 108:37-43.
- [23] Shen F, Wu Q, Shao X, Zhang Q. Non-destructive and rapid evaluation of aflatoxins in brown rice by using near-infrared and mid-infrared spectroscopic techniques. Journal of Food Science and Technology. 2018; 55(3):1175-84.
- [24] Kosmowski F, Worku T. Evaluation of a miniaturized NIR spectrometer for cultivar identification: the case of barley, chickpea and sorghum in Ethiopia. PloS One. 2018; 13(3):1-17.
- [25] Basati Z, Jamshidi B, Rasekh M, Abbaspour-Gilandeh Y. Detection of sunn pest-damaged wheat samples using visible/near-infrared spectroscopy based on pattern recognition. Spectrochimica Acta Part A:

Molecular and Biomolecular Spectroscopy. 2018; 203:308-14.

- [26] Budiastra IW, Widyotomo S, Ayu PC. Prediction of caffeine content in java preanger coffee beans by NIR spectroscopy using PLS and MLR method. In IOP conference series: earth and environmental science 2018 (pp. 1-7). IOP Publishing.
- [27] Huck CW, Guggenbichler W, Bonn GK. Analysis of caffeine, the obromine and theophylline in coffee by near infrared spectroscopy (NIRS) compared to highperformance liquid chromatography (HPLC) coupled to mass spectrometry. Analytica Chimica Acta. 2005; 538(1-2):195-203.
- [28] Reh CT, Gerber A, Prodolliet J, Vuataz G. Water content determination in green coffee-method comparison to study specificity and accuracy. Food Chemistry. 2006; 96(3):423-30.
- [29] Morgano MA, Faria CG, Ferrão MF, Bragagnolo N, de Castro Ferreira MM. Determination of moisture in raw coffee using NIR spectroscopy and multivariate regression. Food Science and Technology. 2008; 28 (1): 12-7.
- [30] Arboleda ER, Fajardo AC, Medina RP. Classification of coffee bean species using image processing, artificial neural network and K nearest neighbors. In international conference on innovative research and development 2018 (pp. 1-5). IEEE.
- [31] Lal S, Behera SK, Sethy PK, Rath AK. Identification and counting of mature apple fruit based on BP feed forward neural network. In international conference on sensing, signal processing and security 2017 (pp. 361-8). IEEE.
- [32] Asokan A, Gunavathi K, Anitha R. Classification of Melakartha ragas using neural networks. In international conference on innovations in information, embedded and communication systems 2017 (pp. 1-6). IEEE.
- [33] Tuteja SK, Bogiri N. Email spam filtering using BPNN classification algorithm. In international conference on automatic control and dynamic optimization techniques 2016 (pp. 915-9). IEEE.
- [34] Ashok Kumar D, Anburajan M, Snekhalatha U. Evaluation of low bone mass and prediction of fracture risk using metacarpal radiogrammetry method: a comparative study with DXA and X-ray phantom. International Journal of Rheumatic Diseases. 2018; 21(7):1350-71.
- [35] Qasim RH, Al-Ani MS. Efficient approach of detection and visualization of the damaged tablets. Journal of Theoretical & Applied Information Technology. 2018; 96(3):643-56.
- [36] Sheba KU, Gladston Raj S. An approach for automatic lesion detection in mammograms. Cogent Engineering. 2018; 5(1):1-16.

- [37] Akkasaligar PT, Biradar S. Diagnosis of renal calculus disease in medical ultrasound images. In international conference on computational intelligence and computing research 2016 (pp. 1-5). IEEE.
- [38] Lekshmanan S, Paul V, Smitha P, Sujathan K. Classification of lung columnar cells using feed forward back propagation neural network. International Journal of Biomedical Engineering and Technology. 2016; 20(4):344-55.
- [39] Khobragade S, Tiwari A, Patil CY, Narke V. Automatic detection of major lung diseases using chest radiographs and classification by feed-forward artificial neural network. In international conference on power electronics, intelligent control and energy systems 2016 (pp. 1-5). IEEE.
- [40] Raghtate GS, Salankar SS. Automatic brain MRI classification using modified ant colony system and neural network classifier. In international conference on computational intelligence and communication networks 2015 (pp. 1241-6). IEEE.
- [41] http://www.coffeeresearch.org/agriculture/drying.htm. Accessed 20 June 2015.
- [42] Caporaso N, Whitworth MB, Grebby S, Fisk ID. Rapid prediction of single green coffee bean moisture and lipid content by hyperspectral imaging. Journal of Food Engineering. 2018; 227:18-29.
- [43] Armstrong PR, Tallada JG. Prediction of kernel density of corn using single-kernel near infrared spectroscopy. Applied Engineering in Agriculture. 2012; 28(4):569-74.
- [44] Williams P. Near-infrared technology: getting the best out of light. Value Added Wheat CRC, Limited; 2004.
- [45] Schmidhuber J. Deep learning in neural networks: an overview. Neural Networks. 2015; 61:85-117.



Edwin R. Arboleda was born in Indang, Cavite, Philippines on October 1979. He graduated from Polytechnic University of the Philippines with a Bachelor of Science degree in Electronics Engineering in 2000 and is a licensed Electronics Engineer. He obtained his Master of Engineering

degree from De La Salle University-Manila in 2009. He is currently pursuing Doctor of Engineering from the Technological Institute of the Philippines-Quezon City. He works as an Assistant Professor in the Department of Computer and Electronics Engineering at Cavite State University, Indang, Cavite, Philippines. His research interests include Artificial Intelligence, Machine Learning, Internet Security, Near Infrared Spectroscopy, and Data Mining.

Email: edwin.r.arboleda@cvsu.edu.ph