Computers and Electronics in Agriculture 172 (2020) 105327

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal

June Anne Caladcad[a,*], Shiela Cabahug[b], Mary Rose Catamco[a], Paul Elyson Villaceran[b], Leizel Cosgafa[a], Karl Norbert Cabizares[b], Marfe Hermosilla[a], Eduardo Jr. Piedad[c]

[a] Department of Industrial Engineering, University of San Jose - Recoletos, Cebu City 6000, Philippines
[b] Department of Computer Engineering, University of San Jose - Recoletos, Cebu City 6000, Philippines
[c] Department of Electrical Engineering, University of San Jose - Recoletos, Cebu City 6000, Philippines

ARTICLE INFO

Keywords:
Coconut
Machine learning
Fruit classification
Acoustic signal
Intelligent systems

ABSTRACT

Advanced intelligent systems are becoming significant to many sectors, including farming. In agriculture, the intelligent classification of post-harvested fruits seems to have a direct impact on farmers, mainly for export products. Unlike other popular fruits, coconuts tend to have limited studies due to its tropical nature grown in developing countries as well as its unique physical structure. In this study, a classification of real coconut datasets is performed based on acoustic signals acquired through a developed tapping system and learned by three widely-used machine learning techniques - artificial neural network (ANN), random forest (RF) and support vector machine (SVM). There are 129 coconuts samples, each classified into three maturity levels – pre-mature, mature, and over-mature. A three-second tapping system gathered from each sample a total of 132,300 data points, which underwent noise reduction and signal processing. Each machine learning model predicts the class of the fruit by learning the patterns of the transformed frequency spectrums of each sample signal. Based on ten times cross-validated results, the three machine learning algorithms satisfactorily predicted the maturity level of coconuts with at least 80% classification accuracy. All models correctly predicted over-mature coconuts but confused in classifying pre-mature with mature and mature with over-mature coconuts. RF model outperformed the other models with efficiencies of 90.98% and 83.48% accuracies for training and testing, respectively. The imbalance data for each coconut class can be addressed to give better results. Additionally, the prepared coconut dataset may use more advanced deep learning techniques.

1. Introduction

Cocos nucifera, commonly known as coconut, is one of the highly nutritious and most popular export fruits from numerous countries (Gatchalian et al., 1994; Jarimopas and Ruttanadat, 2007; Terdwongworakul et al., 2009). Specifically, it significantly contributes to the economic and livelihood development in the humid tropical regions of the world, particularly in the countries of the Indian Subcontinent, Africa, Central America, among others (Ashburner et al., 1997; Chan and Elevitch, 2006). The coconut's hard shell and husk provide exceptional protection to its meat and water making it convenient when being exported to distant countries (Jarimopas and Ruttanadat, 2007). As reported by the Philippine Coconut Authority (2017), the Philippines' government agency administering the country's coconut industry, over 15 billion of coconuts are produced in 2015, and in 2017, exports have reached over 131 million USD worth of coconuts. In addition, the Philippines supplies four-fourths of the world's coconut oil exports, which creates a dominant firm market power in the coconut oil export market (Buschena and Perloff, 1991). On the other hand, Godoy and Bennett (1991) validated the economic benefits of intercropping coconuts even to smallholders in Indonesia. Other than economic gains, coconuts can also be used for medicinal purposes. For instance, the coconut oil, as proven by Feranil et al. (2011), has positive effects on lipid profiles, and can also be utilized against cardiovascular disease risks.

On the other hand, Gatchalian et al. (1994) indicated that the external characteristics and food usage of coconuts vary accordingly to its maturity level. For instance, the pre-mature coconuts can be used as a refreshing drink, mature coconuts for salads or desserts, while over-mature (Gatchalian et al., 1994). Additionally, pre-mature coconuts have tender meat and water. Mature coconuts have thicker meat with lesser water than pre-mature coconuts. While over-mature coconuts

* Corresponding author.
E-mail address: juneannecaladcad@gmail.com (J.A. Caladcad).

https://doi.org/10.1016/j.compag.2020.105327
Received 29 November 2019; Received in revised form 27 February 2020; Accepted 28 February 2020
Available online 13 March 2020

have the hardest meat with lower to no coconut water and its mesocarp fiber turns brittle and dark (Philippine National Standards, 2007). Moreover, Tan et al. (2014) highlighted that its maturity level significantly influences the water composition and physicochemical properties of coconuts. Jackson et al. (2004) also pointed out that coconut water substantially decreases as age increases due to the formation of jelly-like material of coconut water on the inner portion of coconuts, hence the reduction in the volume of water. Also, as presented by Terdwongworakul et al. (2009), the density of coconuts decreases linearly as its maturity increases, and that the coconut's thickness uniformly increases as its maturity level increases.

Bhargava and Bansal (2018) emphasized that the export market and quality evaluation are greatly affected by the sorting systems of fruits and vegetables. In the process of classifying fruits, color has been the significant attribute that dictates the quality and value of many fruit products (Kang et al., 2008; Mohammadi et al., 2015). Thus, the traditional process of sorting fruits according to its maturity level heavily relies on the inspector's perception (Donis-Gonzalez et al., 2013; Mohammadi et al., 2015). In the case of classifying coconuts, Gatchalian et al. (1994) illustrated that farmers and traders use fingernail tapping sounds in traditionally predicting the coconuts' maturity. However, farms' practice of manual classification of crops pertains to losses in yield (Le et al., 2019). Humans performing the sorting process may be prone to inaccuracies and high variability as a result of the individual's subjectivity, fatigue, and lack of training (Brosnan and Sun, 2002; Donis-Gonzalez et al., 2013). It can also be inconsistent, time-consuming, expensive, and easily influenced by its surrounding environment (Bhargava and Bansal, 2018). These errors committed in the selection process would risk the quality of the coconuts to be traded (Gatchalian et al., 1994). In the mass export of fresh coconuts, identification of maturity level is even more urgent since volume requirements may still imply the need to consolidate the fruits before shipment (Gatchalian et al., 1994). Hence, there is a need to propose an automation process to improve the classification systems of fruits, and at the same time, increase the quality, economic growth and productivity in the agricultural sector (Bhargava and Bansal, 2018).

Hahn (2012) utilized an on-line detector in sorting out coconut water according to its corresponding maturity level. Two experiments were conducted; first, maintaining controlled conditions with bioreactor and second, quantifying dissolved oxygen and pH at the different stages of maturity. Hahn (2012) concluded that the use of water from coconuts could sort them to its maturity level with high accuracy. Both very tender and mature coconuts, as clustered by Hahn (2012), could be detected with an accuracy of 100%, and a 92% success rate when sorting tender coconuts. Despite these findings, the study is only limited in sorting out the water content of coconuts. On the other hand, Richardson et al. (2019) conducted a rapid quantification on the adulteration of fresh coconut water using dilution and sugars with the use of Raman spectroscopy and chemo-metrics. In this, Richardson et al. (2019) used partial least squares regression (PLSR) in acquiring quantitative measures and predictive ability. Similar to Hahn (2012), Richardson et al. (2019) also proved that coconut water varies across maturity levels, which could aid in classifying them. Both methods may be impractical as a proposed classification method, especially in exporting coconuts to distant countries.

On the other hand, Gatchalian et al. (1994) proved the scientific basis of the tapping sounds of coconuts as a measure of its maturity level. To produce the tapping sounds, Gatchalian et al. (1994) used a person's middle fingernail and a blunt end of a stainless-steel knife. Analog sound samples are converted to digital formats with the use of analog-to-digital conversion. Analysis of soundwaves showed significant differences across the maturity levels of coconuts. Nonetheless, this study only proved the scientific basis of the tapping sounds of coconuts. Consequently, Terdwongworakul et al. (2009) exploited these acoustic properties in sorting coconuts according to its maturity level, and explored the coconut's physical, mechanical, physiological and acoustic properties across maturity levels. Regardless of the acceptable results of the prediction ability of the non-destructive model that Terdwongworakul et al. (2009) used, it could still be explicitly improved specifically for prediction and classification systems in mass exportation of coconuts.

Related studies have used acoustics to interrogate food products. Pearson (2001) investigated the use of impact sound as a means to separate closed-shell from open-shell pistachio nuts. Acquiring sound upon impact, processing these data, and then dividing the pistachio nuts into either an open-shell or closed-shell product composes the designed sound sorting system (Pearson, 2001). Results showed that the sorting system classified 96.10% closed-shell correctly and 98.00% open-shell correctly. Similarly, Pearson et al. (2007) developed a non-destructive, real-time device that could detect insect damage, sprout damage, and scab damage in kernels of wheat. Kernels are collided to a steel plate, and generated acoustic signal was analyzed to identify if damages are present (Pearson et al., 2007). The designed system correctly classified 87% of insect-damaged kernels with exit holes and 98% of whole kernels. In addition, Onaran et al. (2006) developed a system in which the generated acoustic signal when dropping hazelnuts onto a steel plate is processed to detect underdeveloped nuts from fully developed ones. In this study, Onaran et al. (2006) explored various signal processing techniques such as line spectral frequencies, discrete Fourier transforms, and time domain parameters to determine the best method for classifying fully developed and underdeveloped hazelnuts. While the abovementioned studies have successfully developed detection and classifier systems using acoustic signals, all of these focused onlyon smaller food products like nuts and kernels, and making it difficult to apply on coconuts.

While a significant effort has been made in studying predictive opportunities in sorting coconuts based on their maturity levels, limited effort has been made in establishing a predictive and classification system for coconut fruits. Creating such a system can significantly benefit companies with mass exportation of coconuts, most notably in terms of time and money. Thus, this paper attempts to fill such gaps. In this work, the study attempts to analyze and investigate the maturity levels of coconut fruits in the Philippine setting. As mentioned earlier, the Philippines is one of the biggest suppliers of coconuts locally and internationally. However, the classification system in the country still heavily relies on manual classification with human interventions. Furthermore, this work utilizes the acoustic properties of coconuts, which is already proven by Gatchalian et al. (1994) and Terdwongworakul et al. (2009), to have a scientific basis and significant difference across maturity levels. Furthermore, this study used three prominent machine learning algorithms in determining the maturity levels of the coconut fruits; artificial neural network (ANN), support vector machine (SVM) and random forest (RF). ANN algorithm has been successfully used in identifying and learning correlated data patterns, much known to be adaptive, between input datasets and corresponding actual target values that allows accurate predictions (Piedad et al., 2018; De Smet and Scheeres, 2019; Dozic and Urosevic, 2019). The application of SVM in solving classification problems through finding decision functions using discriminative power and structural risk minimization principle has also been successful (Subasi et al., 2018; Richhariya and Tanveer, 2018; Fayed and Atiya, 2019). Lastly, RF has been applied as a feature selection system which can handle numerous input parameters without having to delete any settings to reduce dimensionality, thus, increasing its prediction abilities (Masetic and Subasi, 2016; Shaikhina et al., 2017; Partopour et al., 2018). This work contributes to providing a practical classification system model with an acceptable standard classification for coconut fruits. This paper is structured as follows: Section 2 presents the materials and methods used in the study. Section 3 discusses the data processes and learning algorithms used in the study, while Section 4 espouses the results, along with a comprehensive discussion of its implications. The paper ends with a conclusion and presentation of future work on Section 5.

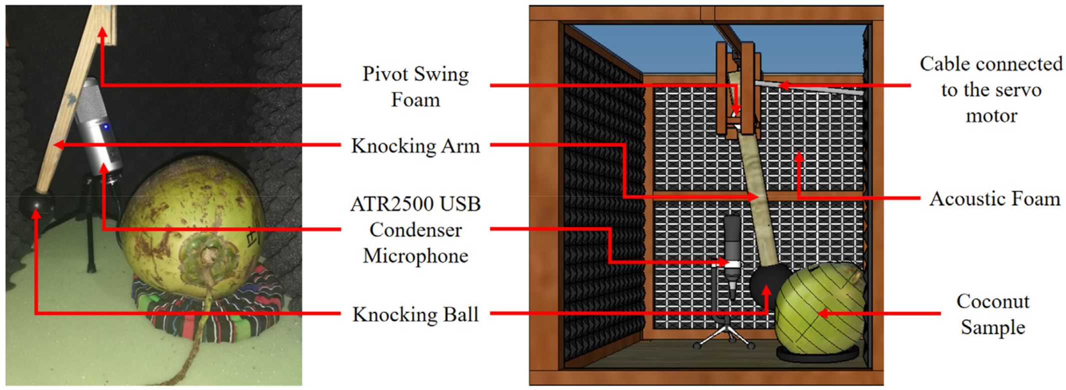

Fig. 1. A developed prototype of the coconut tapping system.

2. Materials and methods

This section presents the development of the tapping system, and discusses in detail both the hardware and software application developments. This is crucial for reliable data acquisition, which the machine learning section heavily relies from.

2.1. Hardware development

The physical design of the coconut tapping system was developed primarily for a reliable data acquisition system. There are two physical chambers – the coconut tapping chamber, and the tapping control and data acquisition chamber.

The first chamber is designed for actual physical tapping of coconut, as shown in Fig. 1. To minimize undesirable interference from the surrounding environment, interior walls were attached with acoustic foams. Also, a rigid hollow holder secures every coconut in a fixed and stable position during the tapping process. All coconuts are tapped uniformly at one of their side ridges instead of their top and bottom ridges, which tend to have uneven thickness among the samples (Gatchalian et al., 1994). Furthermore, an omnidirectional microphone is situated beside the coconut to record the sound of tapping. This noninvasive recording mimics the actual manual tapping process of coconut farmers to classify according to its coconut class. In the study of Terdwongworakul et al. (2009), an impact force is applied at the equator of all the three ridges of a coconut. Similarly, the generated sound also came from the collision of the mechanical tapper and the coconut. Moreover, the construction of a glass door gives access inside the tapping chamber for ease in replacing the coconuts, and conceal it from the surrounding environment every after tapping turn.

The second chamber, composed of servo motor, microcontroller, and computer, is built for controlling the tapping process and gathering data, as shown in Fig. 2. A rigid cable arm, which controls the tapping arm in the first chamber, is connected to the servo motor timed and regulated by the microcontroller unit (MCU). Both the microphone, which is located in the first chamber, and the MCU are connected to the computer. The analog-to-digital converter module of the MCU has 16-bit information length with a sampling frequency of 44.1 kHz following the Nyquist sampling theorem (Baker, 2003). The sampled frequency and amplitude in discrete-time signals underwent the quantization process, and the timing of the tapping and the computer through a developed software application synchronize recording.

2.2. Software development

The developed software application will facilitate data acquisition, as well as in controlling the tapping process. Through this, it provides the user an option to choose clean data and to troubleshoot possible problems encountered during the process, such as loss of connection or power. Note that *CocoKnock* refers to the developed software application.

CocoKnock application interface enables the user to collect information and visualize the acoustic signal recording after the tapping process, as shown in Fig. 3. The designed hardware is synchronized and connected with the software application through the microcontroller unit. It will enable the user to oversee the current status of the recording process, and identify that the recorded data is clean and complete for further data processing and learning.

Various functions and buttons located in the software will help the user in different ways. First, the "CHECK MICROPHONE" button will allow the user to identify if the microphone is functioning well or not. The bar located on its right side will detect sounds captured by the microphone. A green bar will move along it, depending on the captured audio. Located below the "CHECK MICROPHONE" button are the "Select Ridge Label" drag-down option and "START KNOCK" button. The drag-down option will allow the user to label which ridge side the designed hardware is knocking. On the other hand, the "START KNOCK" button will signal the designed device to proceed with the knocking and recording process. Since the acquisition of acoustic signals should be made during the knocking process, both the sound recording and the knocking process are synchronized and done at the same time.

Another feature of the *CocoKnock* application is to enable the user to oversee the current status of the entire process. Located at the leftmost side is an "AUTOMATIC RECORDING STATUS" column. The status column is composed of the following procedure: "SAVING LOCALLY," "KNOCKING," "GENERATING WAVEFORM," "CONVERTING TO CSV," "PLAYING RECORD" and "SAVING TO DATABASE." If a specific process is already done, a designated button for that process will be highlighted green.

On the other side of the software application, located at the rightmost, are buttons for further procedures, and will only be activated after the knocking process. "VIEW WAVEFORM" button would display the waveform of the recorded signals in the middle of the application. Instructions under "INFO" will help the user in manipulating this waveform data for further investigation, if the user wishes to do so. Furthermore, the "CONVERT WAV TO CSV" button will enable the user to convert the current waveform audio (WAV) format to a comma-separated value (CSV) format. "LISTEN TO RECORDING" button plays the currently acquired recording, and the "VIEW RIDGES COMPARISON" button will display, in a waveform, the acoustic signals of the recorded ridges. Lastly, the "START NEW RECORDING" button signals the software application for another knocking process.

Fig. 4 presents the flowchart of the application system process. The diagram visualizes the steps to undergo before the machine learning process. Before the actual knocking and recording process, checking of connection is essential to determine if the software application is successfully synchronized with the designed hardware before the actual

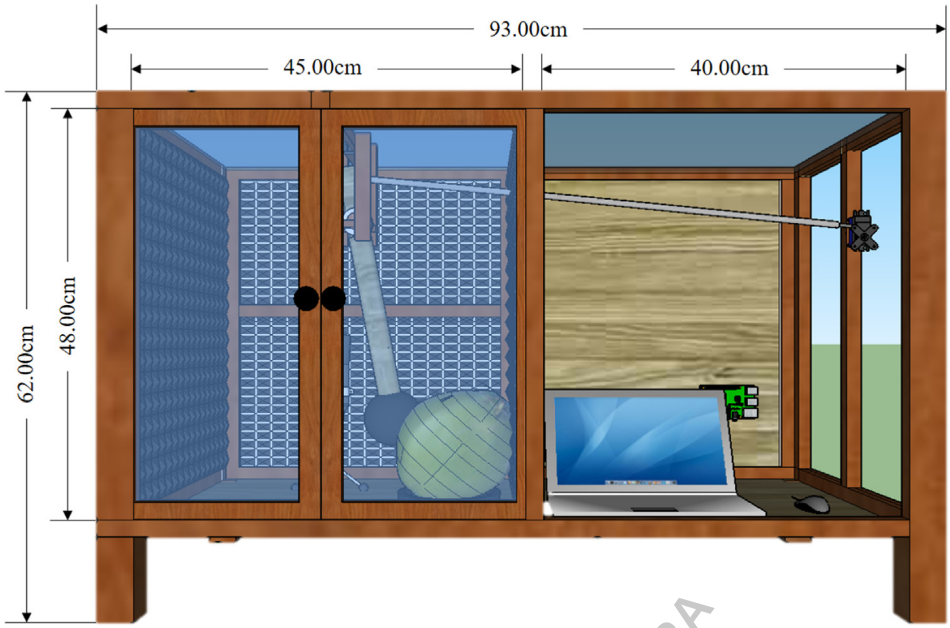

Fig. 2. The physical design of coconut tapping prototype.

knocking and recording process. Note that the checking of the connection between the designed device and the software application is only once, and not in every knocking process. After connection checking, ridge selection is also a crucial step and should be done before the actual knocking process to avoid duplication and overlapping of recordings. If the user did not select a ridge and instead, proceeds with the knocking process, the knocking process will not be carried out; at the same time, an error message will notify the user.

During the knocking process, a command is sent to the servo motor to rotate at 360° thrice, indicating three knocks on the coconut sample. After the knocking process, the user may view and investigate the waveform of the acquired acoustic signal, listen to the recording, and save the file to a designated database. Waveform audio files are imported and retrieved into one CSV file. This CSV file will be used for training and testing machine learning models.

3. Data processing and learning

This section presents and discusses the coconut samples, as well as the process in acquiring the data using the *CocoKnock* application. Also, this section describes the machine learning process of predicting the maturity levels of the coconut samples.

3.1. Coconut samples

Cocos nucifera L. or coconut fruits classified under the tall coconut variety are the subjects of the study. For practical applications, coconuts are classified individually into three maturity levels: "pre-mature," "mature" and "over-mature." As stated by Gatchalian et al. (1994), premature coconuts are the main ingredient for producing coconut drinks and tender coconut meat with coconut water, which possesses export qualities. Mature coconuts are the main ingredient for various delicacies and are highly consumed domestically and traded locally. Lastly, over-mature coconuts, which is also used for cooking, are usually processed for oil extraction and other uses (see Fig. 5).

During the fruit sample preparation, this study only considered coconuts on its post-harvest period. Since most markets need whole coconut fruits, removing some parts of the coconut would be impractical before classification and recording of acoustic signals. Local experts and farmers of the Philippine Coconut Authority, Cebu, Philippines pre-classified the harvested coconuts fruits. Similar to Hongwiangjan et al. (2015), a total of 129 pre-mature, mature, and over-mature coconuts are gathered randomly in which there are 8, 36, and 85 coconuts per maturity level, respectively. It is important to note the imbalanced number of coconuts per class, which serves as the limitation of this study. This uneven sample data is due to the unforeseen season cycle of coconuts when the coconuts already matured during the harvest. The coconut maturity classes are converted into its respective numerical value (0: "pre-mature", 1: "mature", 2: "over-mature").

3.2. Acoustic signal processing

Three acoustic signals are gathered, extracted from tapping three different side ridges from each of the 129 coconut samples, making a total of 387 signals acquired using the *CocoKnock* system. Each signal has 132,300 time-series data points, as shown in Fig. 6.

The acoustic features from these data points are used to determine the maturity level of the coconuts. The frequency domain is the most fundamental transformation to extract acoustic features from a time-series signal (Stankovic, 1994). Fast Fourier transform (FFT) is a well-known technique for digital frequency transformation. In a similar study by Gatchalian et al. (1994), FFT was also the technique used for measuring young coconut's maturity using sound waves. FFT is a simple signal processing technique that uses the discrete Fourier transform (DFT), as shown in Eq. (1).

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{\frac{-j2\pi nk}{N}} \tag{1}$$

Wherein $j = -1$, for every number of audio data k, frequencies are exponentially obtained from a data point $N = 0$ to data point limit $N -$ 1, and which either positive or negative frequency will be obtained. In which for this study, $N - 1 = 132{,}300$ data points (Fessler, 2004).

A python-based library, SciPy, is used to implement FFT using the coconuts' acoustic signals (Jones et al., 2019). Fig. 7 shows a sample frequency spectrum of an over-mature coconut using FFT. Note the presence of frequency modes in the lower frequency spectrums. The other frequencies with insignificant magnitudes in the higher spectrum are suspected to be undesirable noises. These are many causes for these noises, which are unavoidable by nature, such as sound leakage in the

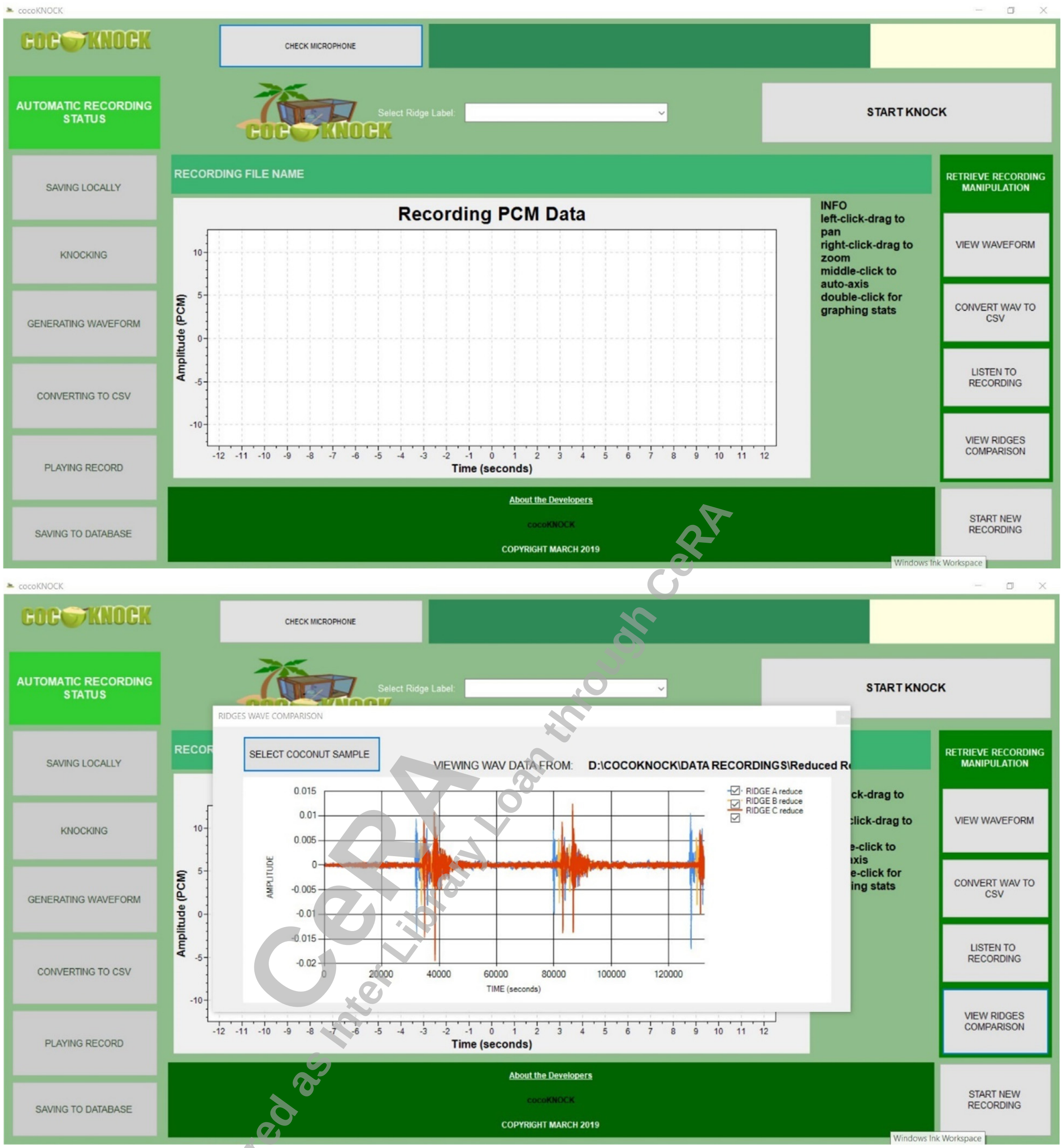

Fig. 3. *CocoKnock* application and interface developed to record coconut's acoustic signal.

first chamber, echoes, and harmonics of the generated sound and other variables intrinsic to *CocoKnock* application and FFT process. This is the second limitation of the study.

3.3. Machine learning

The data in the CSV files, which are imported from the *CocoKnock* application, are used to train and test the three machine learning classifiers. The training phase is when these algorithm classifiers are fed with a set of representative data and be able to learn patterns as the basis for the predictions. Similar to the study of Piedad et al. (2018), three popular machine learning models – ANN, RF, and SVM – are used in this study. Presented in Table 1 are the parameters used for each machine learning models.

In the testing phase, the learned algorithm classifiers discriminate between necessary and unnecessary data encountered during the training phase.

A machine learning pipeline is typically composed of two major phases – training and testing phases, as shown in Fig. 8. The k-fold cross-validation is applied wherein the same dataset is used for both training and validation. Specifically, this study performed a ten-fold cross-validation. The dataset was split into two groups, the training dataset and the testing dataset, wherein this study, the ratio of the division would be 70/30, respectively. The machine learning models are fitted during the training phase, and evaluated during the testing phase. The average of the model evaluation scores determined the estimated skill of the model. In evaluating the models used, there are three metrics used: accuracy, *F-score*, and confusion matrix. The classification

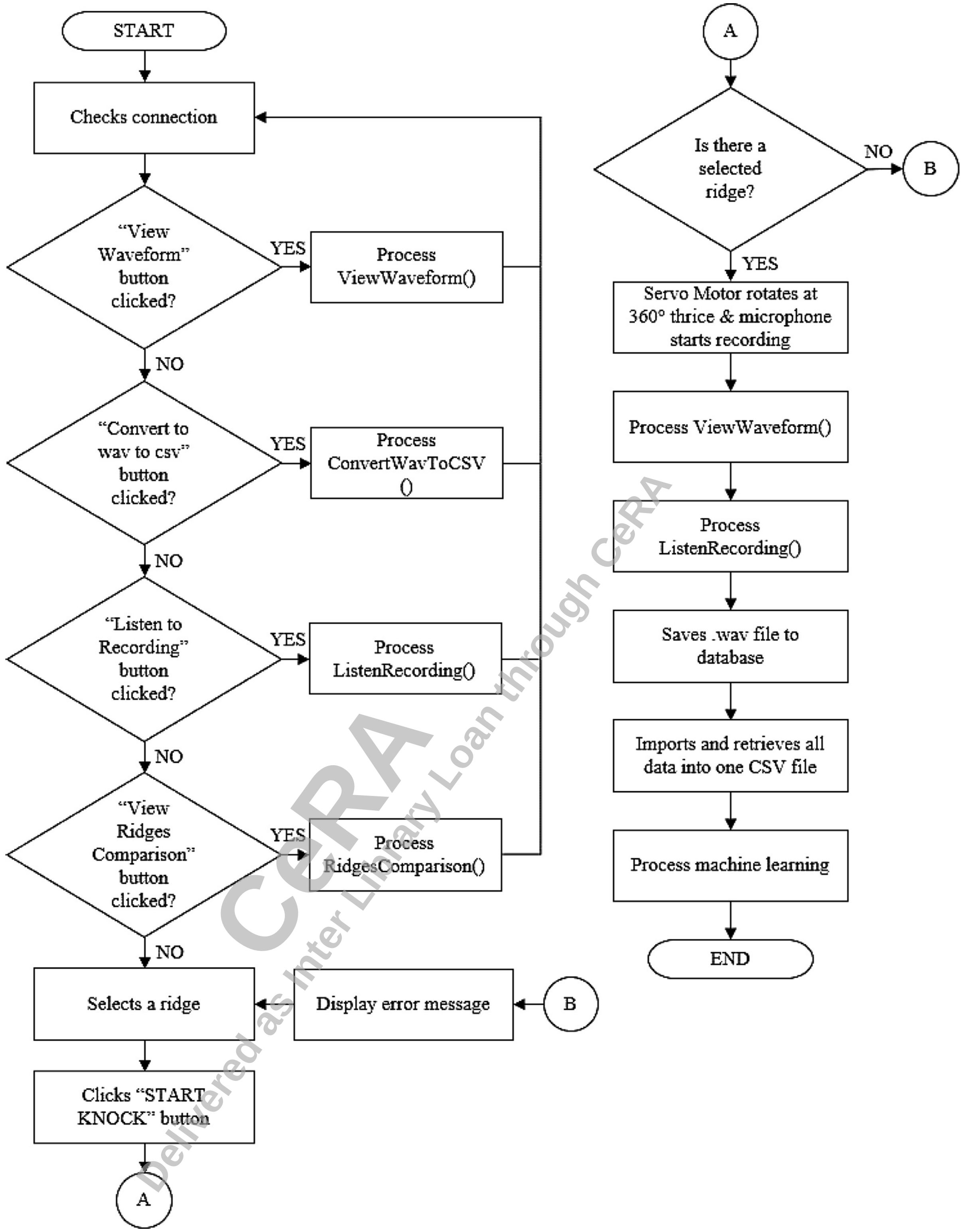

Fig. 4. *CocoKnock* application system process.

accuracy is equal to the number of correctly classified samples over the total number of coconut samples (Piedad et al., 2018; Pedregosa et al., 2011), as shown in Eq. (2).

$$Accuracy = \frac{number \;\; of \;\; correctly \;\; classified \;\; samples}{total \;\; number \;\; of \;\; samples}\% \tag{2}$$

The second measure of accuracy, *F-score*, determines the weighted average of the precision and recall wherein the *F-score* ranges from 0, worst value, to 1, best value (Piedad et al., 2018; Pedregosa et al., 2011).

$$F-score = 2\left[\frac{1}{\left(\frac{1}{Precision}\right)+\left(\frac{1}{Recall}\right)}\right] \tag{3}$$

Furthermore, a normalized confusion matrix will visualize the

(a) (b) (c)

Fig. 5. Harvested (a) pre-mature, (b) mature and (c) over-mature coconut samples.

performance of each classification model by showing the actual classification versus the predicted maturity class.

4. Results and discussions

After establishing the hardware and software prototype, three machine learning models are trained, tested, and evaluated based on the frequency data generated from tapping the coconut samples. Two FFT graphs from each coconut maturity class are displayed in Figs. 9–11, respectively. To compare and illustrate consistency, two sample frequency graphs for each maturity lever are shown.

Accordingly, pre-mature coconut samples tend to have a higher magnitude than the other maturity levels since these coconuts have thinner and softer fibrous husk. In addition, older coconuts seem to have a lower maximum frequency, also known as central/resonant frequency, than the younger ones. Among the three maturity levels, over-mature coconuts are suspected to have the thickest and hardest coconut meat. This development of coconut meat utilizes the inside volume, which also tends to vary the tapped sound of a coconut. Even, as the fruit matures, its water content starts to decrease, thus creating gaps inside, which allows sound waves to bounce, repetitively resulting in higher sound frequency. This variation in sound is the primary indicator utilized by farmers when identifying the maturity level of coconuts using a manual tapping system. The machine learning algorithms mimic this and automate the classification process based on the frequency.

Table 2 shows the comparison of the performance of the three machine learning algorithms. All the algorithms succeeded in getting at least 76% training and testing accuracies. After 10-times cross-validation, RF outperformed the other algorithms with training and testing accuracies of 90.98% and 83.48%, respectively. However, it may tend to overfit due to the 10% difference between its training and testing f-scores. The same is true for the SVM classifier, with almost 12%

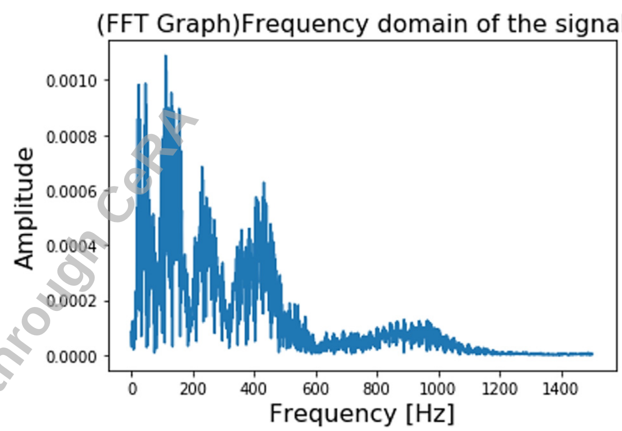

Fig. 7. A sample frequency spectrum of an over-mature coconut using FFT.

Table 1
Parameters used in the machine learning models.

Machine Learning Model	Parameters
Artificial Neural Network	hidden_layer_sizes[32,16,8], maximum iteration = 200
Support Vector Machine	kernel='rbf', C = 1, gamma = 0.5
Random Forest Classifier	n_estimators = 100, maximum depth = 3

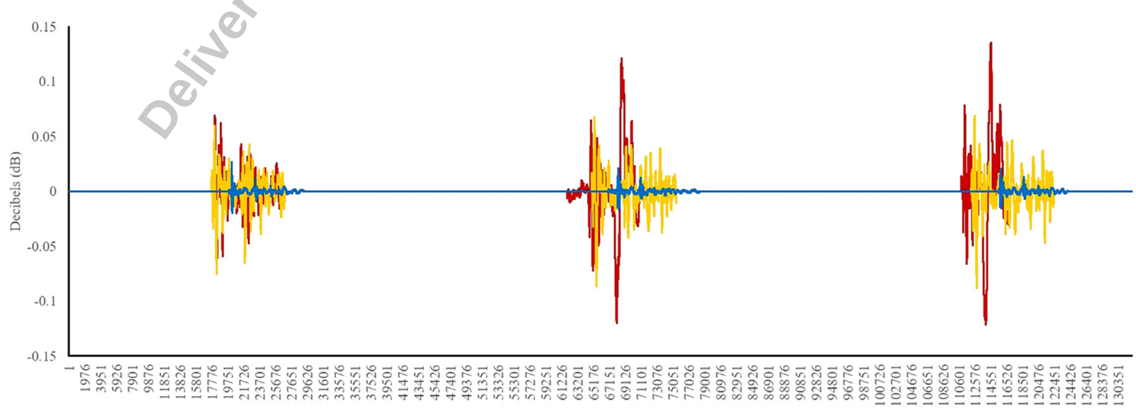

Fig. 6. Sample signals of three coconuts of different maturity levels.

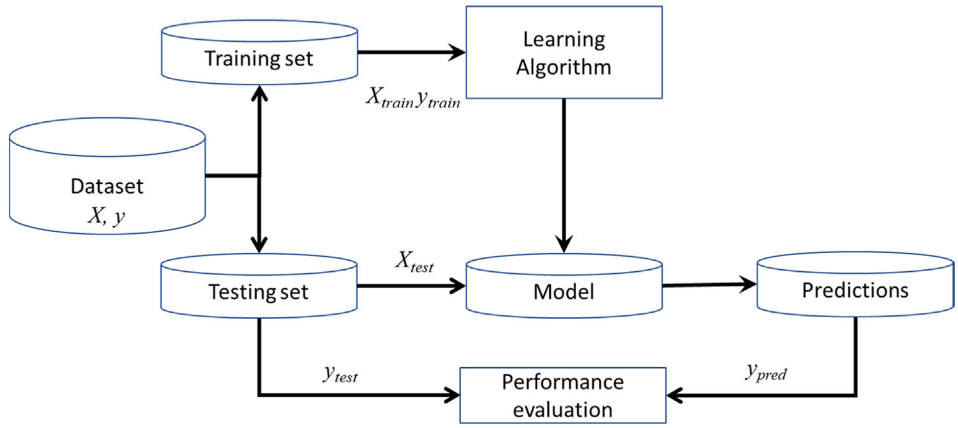

Fig. 8. Machine learning pipeline with training, testing and evaluation phases.

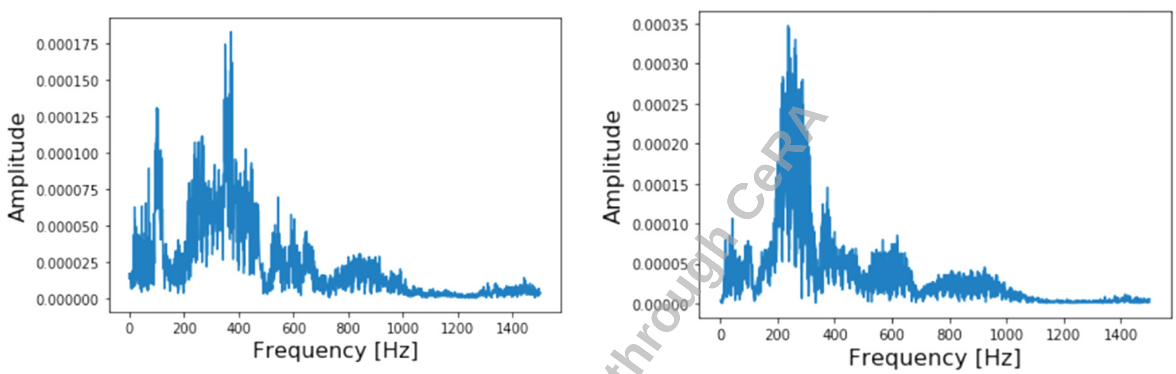

Fig. 9. The magnitude with respect to frequency of two pre-mature acoustic signal samples.

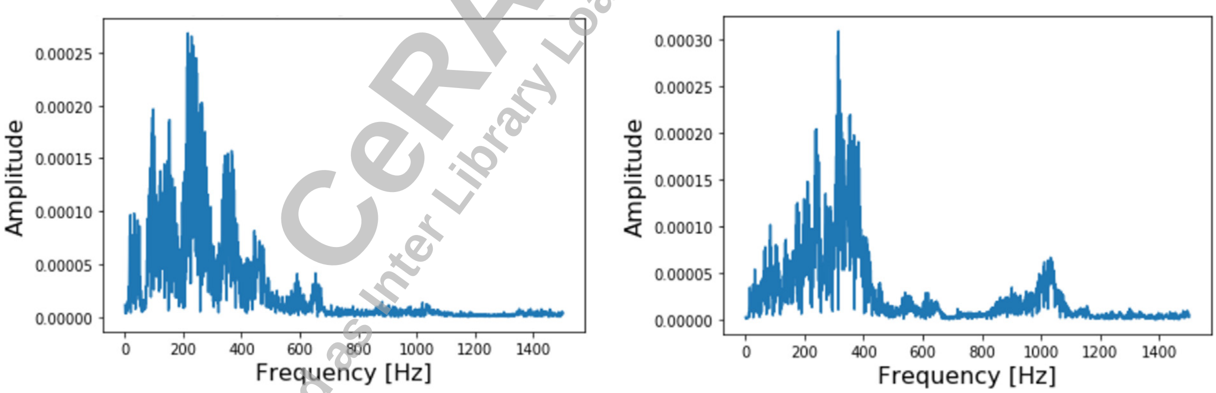

Fig. 10. The magnitude with respect to frequency of two mature acoustic signal samples.

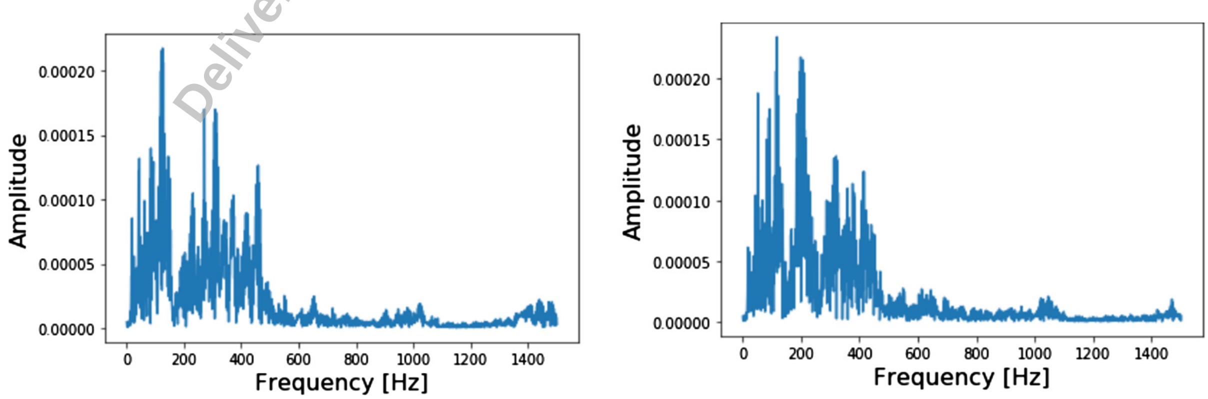

Fig. 11. The magnitude with respect to frequency of two over-mature acoustic signal samples.

Table 2
Performance comparison of the three machine learning classifiers.

Model	Classification Accuracy (%)		F- Score (%)	
	Train	Test	Train	Test
Artificial Neural Network	79.32	81.74	77.46	79.27
Random Forest	**90.98**	**83.48**	**91.41**	**81.35**
Support Vector Machine	88.35	80.00	88.79	76.67

difference in f-scores. On the other hand, ANN has the most reliable f-score with satisfactory classification performance.

The confusion matrices of Fig. 12(a)-(c) further elaborates the performance of the three models. The numbers located in the diagonal axis of each confusion matrix indicates the percentage of correctly classified samples, and the remaining are the percentages of incorrectly classified samples, both in decimal form. For instance, the ANN algorithm correctly classified 38% and 44% of the pre-mature and mature coconut samples, respectively. The remaining 62% of the pre-mature samples and 56% of the mature samples were classified incorrectly as mature and over-mature, respectively. On the other hand, RF classified correctly 25% and 59% of the pre-mature and mature samples, respectively. Seventy-five percent (75%) of the remaining pre-mature samples, and 41% of the remaining mature samples were classified incorrectly as mature and over-mature, respectively. Lastly, SVM categorized correctly 38% for both pre-mature and mature samples. The remaining 62% samples for both pre-mature and mature were classified incorrectly as mature and over-mature, respectively.

All the models perfectly classified over-mature coconuts. However, the models are confused in classifying mature coconuts with over-mature and in classifying pre-mature with mature coconuts. This similar performance among the models is due to the imbalance of data. Across the 129 coconut fruits, there are 255 over-mature coconut acoustic signals, which are more than the 24 pre-mature and 105 over-mature acoustic signals. It seems that all three algorithms have the difficulty of classifying the coconut samples. There are three possible reasons why. First, the number of coconut samples may be insufficient due to limited resources. In artificial intelligence, the higher the number of data samples, the better the learning performance is. Second, the first chamber of the *CocoKnock* prototype is not fully sound-proofed. Various noises may have significantly contributed to the acoustic signals, as shown in Figs. 9–11. Lastly, the three machine learning parameters are manually tuned and are not optimized. A more powerful algorithm, such as a more profound ANN architecture, may be able to discover hidden patterns in the acoustic signal.

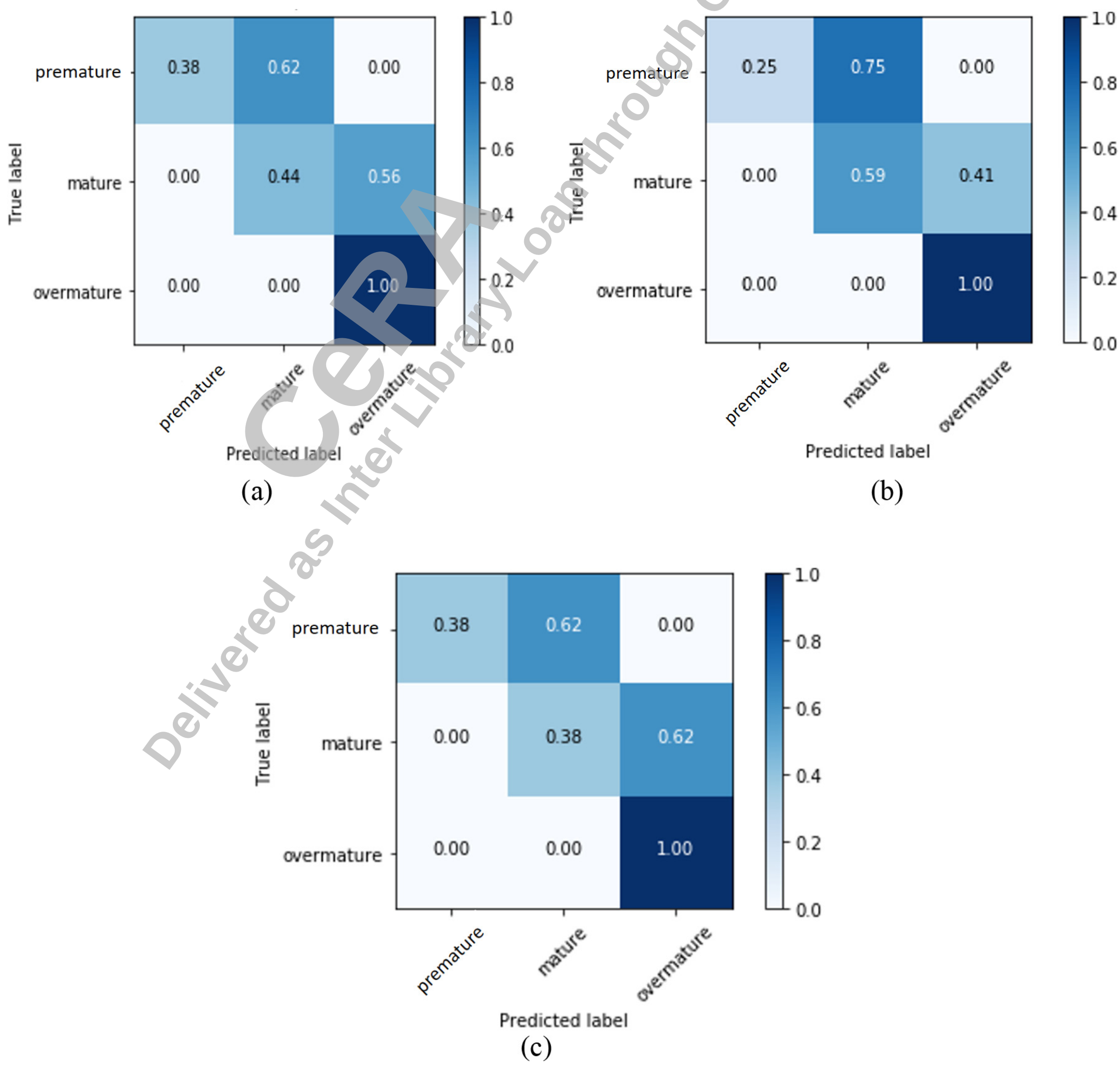

Fig. 12. Confusion matrices for (a) ANN, (b) RF, and (c) SVM classifiers.

5. Conclusion

This research study provides a solution to one of the significant challenges in post-harvest classification activity through the application of artificial intelligence, specifically in its sub-field, machine learning algorithms, as well as the implementation of digital-signal processing, specifically in its sub-field, audio signal processing. Furthermore, the study proposed the application of RF in classifying coconut maturity, which is the machine learning algorithm that performed best, outperforming the ANN and SVM. The findings of this study prove a useful application of digital signal processing and machine learning algorithms in the field of agriculture, specifically for determining the maturity level of coconuts. For future works, the study can be improved through the use of other statistical feature extraction and classifying models. Additional data samples could be collected to increase the accuracy of the algorithms and to avoid tendencies of over-fitting the model.

CRediT authorship contribution statement

June Anne Caladcad: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Shiela Cabahug:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft. **Mary Rose Catamco:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - original draft. **Paul Elyson Villaceran:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation. **Leizel Cosgafa:** Conceptualization, Methodology, Validation, Formal analysis, Investigation. **Karl Norbert Cabizares:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation. **Marfe Hermosilla:** Conceptualization, Methodology, Validation, Formal analysis, Investigation. **Eduardo Jr. Piedad:** Methodology, Validation, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ashburner, G.R., Thompson, W.K., Halloran, G.M., Foale, M.A., 1997. Fruit component analysis of South Pacific coconut palm populations. Genetic Resources Crop Eval. 44 (4), 327–335.

Baker, M., 2003. Demystifying Mixed Signal Test Methods. Elsevier Inc., s.l.

Bhargava, A., Bansal, A., 2018. Fruits and vegetables quality evaluation using computer vision: a review. J. King Saud Univ. – Comput. Inform. Sci. https://doi.org/10.1016/j.jksuci.2018.06.002.

Brosnan, T., Sun, D.-W., 2002. Inspection and grading of agricultural and food products by computer vision systems – a review. Comput. Electron. Agric. 36 (2–3), 193–213.

Buschena, D.E., Perloff, J.M., 1991. The creation of dominant firm market power in the coconut oil export market. Am. J. Agric. Econ. 73 (4), 1000–1008.

Chan, E., Elevitch, C.R., 2006. *Cocos nucifera (coconut)*, 2.1 ed. Permanent Agricultural Resources, Holualoa, Hawaii.

De Smet, S., Scheeres, D.J., 2019. Identifying heteroclinic connections using artificial neural networks. Acta Astronaut.

Donis-Gonzalez, I.R., Guyer, D.E., Leiva-Valenzuela, G.A., Burns, J., 2013. Assessment of chestnut (Castanea spp.) slice quality using color images. J. Food Eng. 115 (3), 407–414.

Dozic, D.J., Urosevic, G., 2019. Application of artificial neural networks for testing long-term energy policy targets. Energy.

Fayed, H.A., Atiya, A.F., 2019. Speed up grid-search for parameter selection of support vector machines. Appl. Soft Comput.

Feranil, A.B., Duazo, P.L., Kuzawa, C.E., Adair, L.S., 2011. Coconut oil predicts a beneficial lipid profile in pre-menopausal women in the Philippines. Asia Pacific J. Clin. Nutr. 20 (2), 190–195.

Fessler, J., 2004. The Discrete Fourier Transform. [Online] Available at: https://web.eecs.umich.edu/~fessler/course/451/l/pdf/c5.pdf [Accessed 5 February 2020].

Gatchalian, M.M., De Leon, S.Y., Yano, T., 1994. Measurement of Young Coconut (Cocos nucifera, L.) Maturity by Sound Waves. J. Food Eng. 23 (3), 253–276.

Godoy, R., Bennett, C.P.A., 1991. The economics of monocropping and intercropping by smallholders: the case of coconuts in Indonesia. Human Ecol. 19 (1), 83–98.

Hahn, F., 2012. An on-line detector for efficiently sorting coconut water at four stages of maturity. Biosyst. Eng. 111 (1), 49–56.

Hongwiangjan, J., Terdwongworakul, A., Krisanapook, K., 2015. Evaluation of pomelo maturity based on acoustic response and peel properties. Int. J. Food Sci. Technol. 50 (3), 782–789.

Jackson, J.C., et al., 2004. Changes in chemical composition of coconut (Cocos nucifera) water during maturation of the fruit. J. Sci. Food Agric. 84 (9), 1049–1052.

Jarimopas, B., Ruttanadat, N., 2007. Development of a young coconut fruit trimming machine. J. Food Process. 79 (3), 752–757.

Jones, E., Oliphant, T., Peterson, P., 2019. SciPy: Open source scientific tools for Python. [Online] Available at: http://www.scipy.org/.

Kang, S.P., East, A.R., Trujillo, F.J., 2008. Colour vision system evaluation of bicolour fruit: A case study with 'B74' mango. Postharvest Biol. Technol. 49 (1), 77–85.

Le, T.-T., Lin, C.-Y., Piedad, E.J., 2019. Deep learning for noninvasive classification of clustered horticultural crops–A case for banana fruit tiers. Postharvest Biol. Technol. 156, 110922.

Masetic, Z., Subasi, A., 2016. Congestive Heart Failure Detection using Random Forest Classifier. Comput. Methods Programs Biomed. 130, 54–64.

Mohammadi, V., Kheiralipour, K., Ghasemi-Varnamkhasti, M., 2015. Detecting maturity of persimmon fruit based on image processing technique. Sci. Hortic. 184, 123–128.

Onaran, I., Pearson, T.C., Yardimci, Y., Cetin, A.E., 2006. Detection of Underdeveloped Hazelnuts from Fully Developed Nuts By Impact Acoustics. Trans. ASABE 49 (6), 1971–1976.

Partopour, B., Paffenroth, R.C., Dixon, A.G., 2018. Random Forests for mapping and analysis of microkinetics models. Comput. Chem. Eng. 115, 286–294.

Pearson, T.C., 2001. Detection of pistachio nutes with closed shells using impact acoustics. Appl. Eng. Agric. 17 (2).

Pearson, T.C., Cetin, A.E., Tewfik, A.H., Haff, R.P., 2007. Feasibility of impact-acoustic emissions for detection of damaged wheat kernels. Digital Signal Process. 17 (3), 617–633.

Pedregosa, F., et al., 2011. Scikit-learn: Machine Learning in Python. J. Machine Learning Res. 12, 2825–2830.

Philippine Coconut Authority. 2017. Department of Agriculture - Philippine Coconut Authority. [Online] Available at: http://pca.gov.ph/industry/coconut-statistics/2-uncategorised/64-export-performance-of-coconut-products-2017 [Accessed 13 July 2019].

Philippine National Standards. 2007. Bureau of Agriculture and Fisheries Standards (BAFS). [Online] Available at: http://www.bafps.da.gov.ph/2017-10-12-00-46-55/standard-formulation/philippine-national-standards [Accessed 2019].

Piedad, E.J., Larada, J.I., Pojas, G.J., Ferrer, L.V.V., 2018. Postharvest classification of banana (Musa acuminata) using tier-based machine learning. Postharvest Biol. Technol. 145, 93–100.

Richardson, P.I.C., Muhamadali, H., Ellis, D.I., Goodacre, R., 2019. Rapid quantification of the adulteration of fresh coconut water by dilution and sugars using Raman spectroscopy and chemometrics. Food Chem. 272, 157–164.

Richhariya, B., Tanveer, M., 2018. EEG signal classification using universum support vector machine. Expert Syst. Appl. 106, 169–182.

Shaikhina, T., et al., 2017. Decision tree and random forest models for outcome prediction in antibody incompatible kidney transplantation. Biomed. Signal Process. Control.

Stankovic, L., 1994. A Method for Time-Frequency Analysis. IEEE Trans. Signal Process. 42 (1), 225–229.

Subasi, O., et al., 2018. Exploring the capabilities of support vector machines in detecting silent data corruptions. Sustain. Comput.: Inform. Syst. 19, 277–290. https://doi.org/10.1016/j.suscom.2018.01.004.

Tan, T.-C., et al., 2014. Composition, physicochemical properties and thermal inactivation kinetics of polyphenol oxidase and peroxidase from coconut (Cocos nucifera) water obtained from immature, mature and overly-mature coconut. Food Chem. 142, 121–128.

Terdwongworakul, A., Chaiyapong, S., Jarimopas, B., Meeklangsaen, W., 2009. Physical properties of fresh young Thai coconut for maturity sorting. Biosyst. Eng. 103 (2), 208–216.