# *Non-destructive Ripeness Determination of Watermelon through Mel Frequency Cepstral Co-efficients*


Khen Zeimmer C. Maranan
*College of Engineering*
*Batangas State University*
Batangas City, Philippines
khen.maranan@g.batstate-u.edu.ph

Chinie M. Dimacquiat
*College of Engineering*
*Batangas State University*
Batangas City, Philippines
chinie.dimacquiat@g.batstate-u.edu.ph

Nikki S. Villanueva
*College of Engineering*
*Batangas State University*
Batangas City, Philippines
nikki.villanueva@g.batstate-u.edu.ph



***Abstract*—The degree of ripeness of the food is an essential factor that determines its overall quality. While watermelon ripeness is a common aspect measured by the fruit's acoustic properties, this study presented a discussion that presents watermelon ripeness determination through the sound produced by the fruit upon tapping its rind. One hundred watermelons were sampled in this study where fifty (50) watermelons belong to freshly harvested, and the remaining is classified as over ripe watermelons as it were stored for 1 month after harvest. Both Convolutional Neural Network or CNN was used as well as Feedforward Neural Network or FNN. A Convolutional Neural Network, or CNN, were utilized in order to train the data that was acquired. CNN generated a model which accuracy amounts to 95.833% and a testing accuracy of 88.333. FNN which yielded in a better model has an accuracy of 92.50%, and a testing accuracy of 90.833% which effectively classifies ripe and over ripe watermelons based on their acoustic properties.**



***Keywords—non-destructive testing, acoustic properties, ripeness, machine learning***


## I. INTRODUCTION

Food is one of the most important things in people's lives. It has the nutrients that a body needs for a person to survive. The nutrients coming from the food the person eats are also based on its quality. The producers may determine its quality but after 1 week they may not know the fruits or vegetable quality in retailers or resellers. Thus, this problem is one of the retailers selling fresh fruits/vegetables. When it comes to fresh foods, ripeness, and quality is essential yet difficult to define and measure. This is one of the factors that affect retailers and consumers in the market.

Food quality is crucial but difficult to define and measure. This is one of the market factors that affect retailers and consumers. When the time for collection is limited, problems of this kind become considerably more challenging. Due to the unpredictability of the harvest, retailers frequently have a one-week timeframe for fruits and vegetables. The outcome of the fact that retail stores experienced problems with stocking their goods without ensuring that they were in their proper condition or that their quality was accurately measured, resulted in the waste of food [1].

Ripeness which was interchangeably used with freshness, is an important factor in determining fruit's quality and general consumer acceptance. The exterior of the fruit is not, however, a reliable indicator of its actual quality. Thus, consumers cannot discern the fruit's quality based on its exterior appearance [2]. Due to the watermelon's thick exterior, it is more difficult to determine its condition unless it is cut open. Moreover, post-harvesting is the final and most crucial aspect of agriculture that requires additional attention. After completing all processes beginning with yield estimation and concluding with harvesting, post-harvest neglect can reverse all the farmers' hard work and result in significant financial losses. Shelf-life of fruits and vegetables, post-harvest classification, and export may be considered at this stage [3].

The challenge of retaining the freshness of fruit was one of the factors that were examined in [4] on the topic of fruit freshness. The freshness of the fruit can only be maintained and established by doing thorough inspections consistently. They created an automated fruit recognition and freshness evaluation system with audio output by utilizing machine vision. This is the strategy that stops the fruit from going bad and minimizes the amount of food that is wasted.

Due to the subjective character of assessment and the destructive nature of measures, measuring internal and external attributes was difficult and time-consuming in the past. Initially, the majority of the evaluation has been based on the visual appeal and defects of the fruit, which has resulted in wrong inferences. Furthermore, manual sorting is extremely reliant on subjective human labor and time management. For example, the conventional process for selecting a high-quality watermelon includes tapping the fruit's skin to evaluate its sound. The sound is used to detect a fault in the hollow heart of the fruit. The search for yellow patches on the skin's surface is another typical strategy [5].

Numerous techniques exist for determining and categorizing the quality of fruits. Some methods require an external examination of fruits, whereas others require an in-depth analysis of the fruits' internal composition. Reference [6] evaluated watermelon quality determination with machine vision using images, which entails capturing the watermelon's physical appearance with a device and using color masks to evaluate the fruit's quality. Using the principle of sound absorption, [7] attempted to classify mature and unripe watermelons in their experimental setup. They utilized immature and ripe watermelons, as well as rotting fully ripe watermelons. Reference [8] developed a ripening detector to assess the ripeness of cantaloupe in 2021. In contrast, [9] utilized a volatile organic compounds (VOC) sensor to measure the freshness of fruits, including bananas, tomatoes, and yellow pepper.

Non-destructive testing includes more than just tools or machinery; it also covers agriculture, particularly the daily foods like fruits. Non-destructive testing, or NDT for short, refers to a variety of inspection techniques that allow inspectors to gather information and assess a system, component, or material without damaging it or permanently altering it. [10] Non-destructive testing is used on fruits to assess their quality using a variety of techniques without cutting or opening the fruit [11].

Furthermore, reference [12] presented a ripeness classification of watermelon fruits based on acoustic testing with a control unit. They defined ripe as watermelons with a shelf life within two weeks and on the other hand, watermelons with a shelf life with more than two weeks as over ripe.

Thumping is an effective method for detecting overripeness in round-shaped melons. Fruit thumping creates a metallic ringing sound indicating immaturity. Furthermore, a more muted or bland sound indicates maturity or overmaturity [13]. With experience, the thumping technique can be useful. However, while expert growers may be able to detect the ripeness of watermelons using the "thump test," the vast majority of people will struggle to distinguish between the sounds, especially consumers [14].

Given that watermelons can stay a month in stores, additional challenges develop for retailers and buyers who require greater skill to evaluate the watermelon's quality. Retailers are not farmers, and the majority of their things come from the same enormous wholesale market in the city, where quality does not differ [3].

Having information about the good characteristics helps in conducting proper storing and proper handling of produce. There are commonly used conventional non-destructive indicators to assess the maturity of watermelons. However, some of these are subjective, making maturity to be sometimes difficult to determine [13].

Furthermore, existing methods used in watermelon quality detection includes but not limited to near-infrared spectroscopy, acoustic methods, electrical and magnetic approaches among others.

As stated in [15] the research on watermelon quality with acoustic properties has improved significantly in recent years. The technologies of cutting-edge equipment and innovative methodologies existed. Other algorithms have now been used by researchers to create classification models to study the relationship between auditory characteristics and the quality of watermelons.

The acoustic property is recommended to determine the inner quality when it comes to its lightweight and handy parts which enable researchers and users to bring it everywhere.

Moreover in the analyzed study, by observing the effect of striking a ball and a fruit tray to the spectrum, it was discovered that there was a clear association between hardness and firmness indices. This led to the development of an acoustic instrument. As a result, people have a preference for instruments that are convenient to carry. It has an increasingly significant role and is just like mobile devices play in people's everyday lives.

Determining the watermelon's ripeness while knowing how long the fruit sits in a store will assist in assessing its marketability and shelf life. In addition, it will also help in better quality control and promote food waste reduction.

Among the methods used in watermelon quality detection which include near-infrared spectroscopy, acoustic methods, and electrical and magnetic approaches, acoustic signal detection offers the advantages of simple equipment and inexpensive cost compared to non-destructive approaches such as near-infrared spectroscopy and nuclear magnetic resonance. As a result, it is extensively utilized to detect watermelon internal quality information [16]. Furthermore, among different non-destructive technologies, it has garnered the most interest from academics and researchers since the vibration information obtained by this method was directly related to the mechanical and physical qualities of the fruit [17].

This study aims to determine the differences in sound quality that are produced by watermelons of different level of ripeness. The sound waves were gathered using a device that was built specifically for the purpose of tapping the rind of the fruit while simultaneously recording the sound that the tapping would produce.

## II. METHODS

Fig. 1 presents the flow chart for generating and evaluating the model which will classify the watermelon as ripe or over ripe. Data collection, data treatment, model generation, and model evaluation are all part of the classification process.

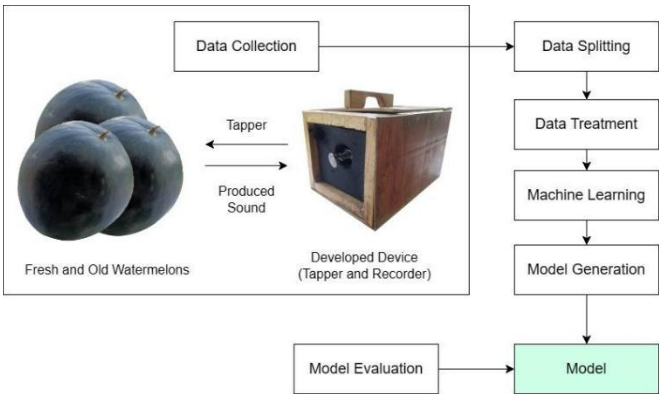



Fig. 1. Model generation and evaluation workflow

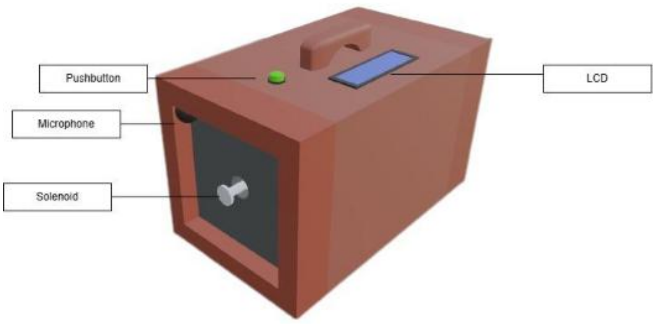
Following model evaluation, the generated were fed to the model in order to assess its performance real time. This also aided the selection of which generated model performs better.

### A. Data Collection

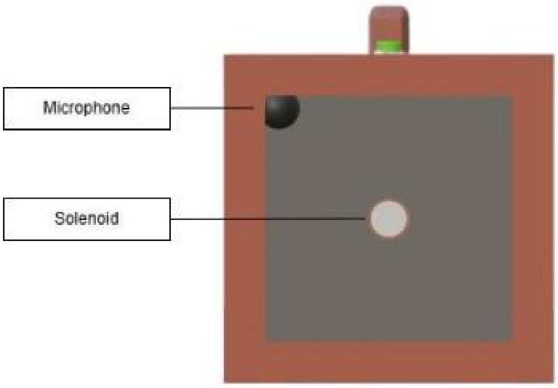
The conventional technique of using fingers to tap or thump a watermelon’s rind may influence the produced sound of the fruit upon tapping due to inconsistency of force, speed, and distance. With this, a device capable of tapping the fruit and recording the sound it will produce is developed to aid the data-gathering process. The two main processes of the device are the following: the tapping mechanism and the sound recording.

The developed device features a microprocessor, which is responsible for controlling the function of the device. It is also responsible for storing the sound files produced upon tapping. In addition, it also features a trigger which is through the push button and a display which allows the device to have a potential to be used as a portable device.

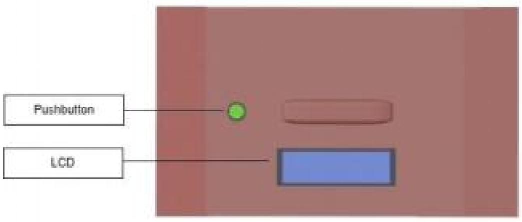
The figure below shows the design of the device in a 3D layout.




a.




b.




c.

Fig. 2. Design of the device in (a.) isometric view, (b.) front view and (c.) top view

The tapping mechanism is composed of a push-pull solenoid where a stainless steel disk is attached to its end. It has a maximum output force of 5N and an impact distance of 1 cm. On the other hand, the sound recording is done with an omnidirectional USB microphone, having a frequency range of 20 to 20KHz, an impedance of 480 kΩ, and a sensitivity of -47dB±3dB.

Fig. 2 features the front of the developed device with its two main components. The tapping mechanism is placed in the middle while the microphone is placed on the upper left side of the device. The processes where these components are involved run simultaneously.

Fig. 3. Front of the developed device used in data collection

The watermelon samples used in this study were purchased in two batches with each batch for each class. The first batch consisted of over ripe watermelons which were one month old at the store while the second batch consisted of ripe watermelons and were freshly harvested. Both batches were bought from the same vendor and farm in Nueva Ecija.

Fig. 4. Watermelons used for data collection

Watermelon, also known as Citrullus-lanatus, has different varieties in the Philippines including red delight, orange delight, sweet 16, sweet 18, and sugar baby [18]. Sugar baby is common in the Philippines, it is described as round watermelons that are sometimes referred to as "picnic" or "icebox" watermelons. These melons are ideal for small families and, true to their name, are compact enough to be stored in an icebox. They have a circumference between 7 and 8 inches (18-20 cm) and weigh between 4 and 5 kilograms (8-10 pounds).

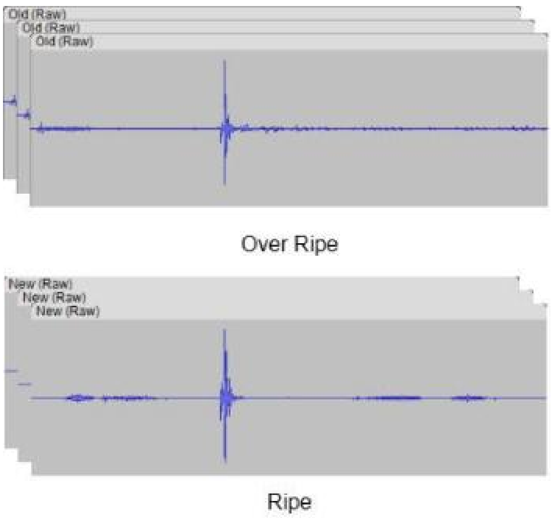
In gathering the sound produced by watermelons upon thumping or tapping, sugar baby watermelons were used. The watermelon is placed in a secured spot and ensured that it will not move from its position. Fig. 4 presents the setup used for data gathering. The data collection took place in an uncontrolled environment with minimal to no noise tho environmental noise is still present.

Fig. 5. Data collection setup

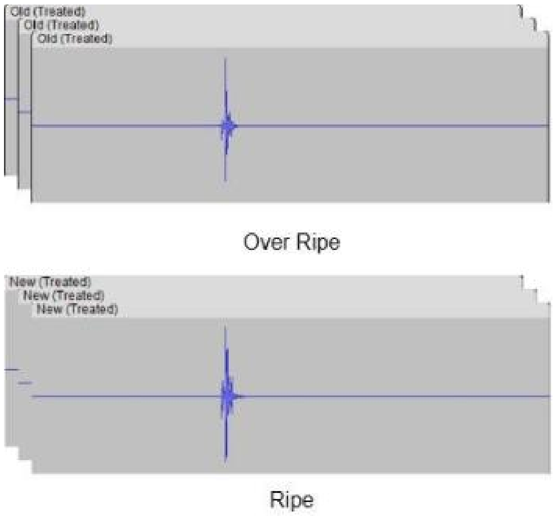
A total of 100 watermelons weighing 3 to 4 kilograms were used in this study. Fifty (50) of the watermelons are freshly harvested watermelons which were classified as ripe while the remaining half are watermelons that were harvested a month before and were classified as over ripe. Each watermelon was tapped 6 times on different sections producing 600 sounds with 300 sounds for each class. The gathered dataset are all saved as a wav file.

*B. Data Treatment*

After collecting the acoustic responses of the fruit, the data is subjected to data treatment. Audacity is used in removing noise on the recorded sounds for training, with its noise reduction feature. All of the sound recordings have a duration of 2 seconds featuring a single tap of the watermelon. The overall duration of the audio is 3 seconds. Fig. 5 shows the sound waves of the raw sound produced by tapping the watermelon with the help of the developed device. It can be seen that the raw sound waves consist of ambient noise.




**(a)**




**(b)**

Fig. 6. Sound waves of (a) raw dataset, and (b) treated dataset

The dataset was split with 80% of it used in training while the remaining 20% is used for testing. A total of 600 sound files were used for generating the model while 120 sound files were used for testing the effectiveness of the generated model. The overall dataset contains of both filtered and unfiltered sound.

*C. Tensorflow & Librosa*

The two main libraries utilized in this paper are Tensorflow and Librosa. Tensorflow is commonly used in artificial intelligence and machine learning while Librosa is known for working with audio and for performing audio analysis. Tensorflow has been incorporated into the code in order to import numerous modules and functions derived from a variety of external packages in order to construct a deep learning model utilizing the Keras API.

Keras is a high-level neural network application programming interface (API) that was created in Python. It is compatible with operating on top of CNTK, Theano, or TensorFlow. The model is defined in Keras by making use of the Sequential model type, which makes it possible to construct a layer stack in a linear fashion. The code imports Sequence from the Keras library in order to make it possible to develop Python generators that can create data batches for the purpose of training deep learning models using big datasets [19].

In addition, to_categorical is imported to convert a vector of class labels into a one-hot encoded matrix, and LabelEncoder is used to convert category labels into numerical values. Both of these functions are described in more detail below. The code also imports a number of different Keras modules, including as losses, models, and optimizers, in order to supply a number of different loss functions, model topologies, and optimization methods for the purpose of training deep learning models.

A deep neural network model is created by importing several different layers from the Keras library, including Dense, Dropout, Input, Activation, and Flatten, among others. In the final step, an evaluation of the effectiveness of a classification model is carried out by importing the confusion_matrix file from the scikit-learn package.

In general, computations are being done with TensorFlow using Keras because Keras can be configured to use TensorFlow as its backend. This is how TensorFlow is being used within the code [19].

While Librosa, on the other hand, was used within the code to extract the MFCC features which are used as the basis of the dataset that needed to be compared with one another, within the training model. It is   to extract also used in various ways to extract features such as Mel-Spectrogram or amplitudes which is used as another way for comparison for the difference f frequencies of ripe and over ripe watermelon audios [20].

*D. CNN Architecture*

The first architecture for identifying and predicting the classes within the study is Convolutional Neural Network (CNN). Convolutional Neural Networks also known as CNN are a specific sort of deep neural network that is frequently utilized for the process of image and video recognition. The primary purpose of CNNs is to automatically learn relevant features from the input images through convolutional layers, which are then followed by pooling layers which is responsible for reducing the dimensionality of the feature maps, and lastly fully connected layers, which conduct classification [21].

The architecture of a CNN is comparable to that of the connectivity pattern of neurons in the human brain, which was the source of inspiration for the architecture of the CNN. The term "receptive field" refers to a specific portion of the visual field that is the sole place where individual neurons will respond to inputs. The entire visible region is covered by a cluster of fields that overlap one another [21].

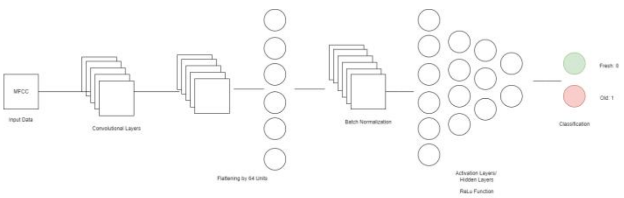



Fig. 7. CNN Algorithm Used within the Training

The above shows how the main process of CNN is used within the training and testing of the dataset created. First and foremost, the gathered data were listed in a csv file alongside their appropriate labels such as filename, filepath, classification, and lastly MFCC. CNN used the csv file as a way to direct the location of each file in order to be accessed. Each audio file is converted into an MFCC file with the usage of the librosa python library. The resulting data are then processed by the CNN algorithm to train and test the model created. As images are just plotted numbers, the usage of MFCC and CNN was decided to be the best choice of model to be used within the study.

The generated MFCCs were initially processed by a number of Convolutional Layers in order to extract specific features and Feature maps for each of the audio files included in the dataset. The first stage is the activation function. In this stage, the feature map is run through an activation function (such as ReLU) to introduce non-linearity into the model and improve its capability of modeling complex interactions between the input and output.

The output of the activation function is sent through a pooling layer after being processed by the activation function, which downsamples the feature map in order to make it smaller. Because of this, the computational complexity of the model is simplified, and it also becomes more resilient to changes in the input.

The process of applying convolutional layers, activation functions, and pooling layers is repeated a number of times, with each layer learning increasingly complicated representations of the input data as it progresses through the process. Once the feature maps have been generated by the convolutional layers, they are then "flattened" into a single vector before being passed through one or more fully connected layers responsible for the final classification process. The output of the final fully connected layer is then sent into a softmax activation function, which generates a probability distribution over the many classes to which the input image could potentially belong [22].

*E. FNN Architecture*

In generating the second model for the target classification system. Feedforward Neural Network (FNN) was used. FNN is a tipe of artificial neural network that does not have looping connections between its nodes.

Because all information only flows forward in feedforward neural networks, they are frequently referred to as a multi-layered network of neurons.

After passing through the hidden layers and entering the input nodes, the data eventually leaves the output nodes. There are no links in the network that would allow data leaving the output node to be transmitted back into it. Furthermore, feedforward neural networks are designed to approximate functions.

A feedforward neural network (FNN) is an artificial neural network that allows information to move in a single direction, starting from the input layer and passing through hidden layers before reaching the output layer. FNNs are widely used in classification, regression, and pattern recognition tasks due to their simplicity and popularity in the field of neural networks. Thus, the process of FNN started with an input layer that received the initial data or features then after the input layer, there are one or more hidden layers in the FNN architecture. These hidden layers are positioned between the input layer and the output layer. Each hidden layer consists of multiple neurons or nodes, which process and transform the information from the previous layer [23].

The specific model or training that the model has taken is specifically called a Sequential Model. It is the simplest and most common type of model architecture used in deep learning frameworks like Keras. In a sequential model, each layer is connected to the previous and next layers, forming a sequence or chain of layers.

With the main process of exttracting audio features using librosa as te main dependancy with the implemntation of MFCC all of the extracted data are then saved within an array of data, which will then be trained immediately through the generated sequential model training. As the main difference between FNN and CNN is that CNN uses a convulutional network in training while FNN does not apply it. It shows that the extracted MFCC features must be plotted and processed repeatedly in CNN while FNN uses the generated audio features within an array to be trained directly.

The training begins by initializing the model and subsequently adds a series of dense layers (total of 3 layers) with varying numbers of units. Each dense layer is followed by a rectified linear unit (ReLU) activation function, which introduces non-linearity, and a dropout layer that randomly sets a fraction of input units to zero during training to prevent overfitting. The final dense layer has a number of units equal to the number of output labels, and it is followed by a softmax activation function that converts the output values into probabilities representing class probabilities. Overall, this model architecture enables the neural network to learn hierarchical representations of the input data and make predictions based on the computed probabilities.

The model is then trained with an overall 800 epoch and a batch size of 12. Within the training a ModelCheckpoint callback in TensorFlow Keras is implmented. This callback allows for the automatic saving of the model's weights during training, ensuring that only the best-performing model based on the validation loss is saved.

### F. Model Evaluation and Inference

After data treatment, the dataset is subjected to data training with both CNN and FNN which generates models whose performance will be evaluated. These models were calibrated with batch size and a number of epochs.

After the generation of the training model and Training and Testing Numpy files, which followed the main training of the dataset/model, it is then used for the calculation of the accuracy, loss, val_loss, and aval_accuracy of the generated model. Model evaluation is done by using the generated model to predict the classes of the sound files from the testing dataset. Its accuracy on the other hand is calculated with PyCM which is also responsible for visualizing the model performance through confusion matrix.

In the analysis of the generated models, PyCM was used. PyCM is an appropriate tool for post-classification model evaluation that supports most classes and overall statistics parameters, as well as a multi-class confusion matrix library that accepts both input data vectors and direct matrices [24]. Such parameters include accuracy and area under the ROC curve (AUC). Accuracy (ACC) refers to the number of correct predictions from all predictions made (1), where TP refers to true positive, TN as true negative, P as condition positive, and N as condition negative. AUC on the other hand, corresponds to the arithmetic mean of sensitivity and specificity values of each class (2), where TNR refers to true negative rate and TPR to true positive rate.

$$ACC = \frac{TP+TN}{P+N} \tag{1}$$

$$AGF = \sqrt{F_2 \times InvF_{0.5}} \tag{2}$$

$$AUC = \frac{TNR+TPR}{2} \tag{3}$$

## III. Results and Discussions

This section provides a brief discussion of the results of the training, validation, and testing.

### A. Data Collection

The dataset consists of 600 sound files produced by watermelons upon tapping using the developed device. Each class consists of 300 sound files sampled at 48kHz in .wav file extension. In data training, Python 3.9.16 was utilized as Google Colaboratory Pro is used. A total of 600 sound files are used for the data training.

The following figures show some of the gathered sound dataset for ripe and over ripe watermelons.

Fig. 8. Sound files gathered from ripe watermelons

Fig. 9. Sound files gathered from ripe watermelons

These sound files were first stored in the microprocessor of the developed device and were then retrieved to undergo data treatment.

### B. Model Results

For the model generated by the use of CNN architecture, after training the training dataset with a batch size of 32 and epoch of 150, the accuracy and loss of the training and validation stage are listed in Table I.

TABLE I. MODEL1: ACCURACY AND LOSS FOR TRAINING AND VALIDATION

| Stage | Accuracy | Loss |
|---|---|---|
| Training | 1.0000 | 1.1362e-05 |
| Validation | 0.9583 | 0.1514 |

On the other hand, the accuracy and loss of the training and validation stage of the model generated with the use of FNN architecure specifically sequential model is shown in Table II.

TABLE II. MODEL2: ACCURACY AND LOSS FOR TRAINING AND VALIDATION

| Stage | Accuracy | Loss |
|---|---|---|
| Training | 0.9750 | 0.0684 |
| Validation | 0.9250 | 0.2101 |

Both of the generated models showed a high model accuracy with a training accuracy of 1.000 and a validation accuracy of 0.9583 for the first model and a training accuracy of 0.9750 and a validation accuracy of 0.9250 for the second model.

Confusion matrices were generated to visualize the model performance of both models. To plot the confusion matrix, Matplotlib, a Python plotting library was used. Fig. 7 shows that out of 60 ripe watermelons, 55 watermelons are predicted by the model as ripe or 0, while all over ripe watermelons are correctly classified as old which is represented by 1.

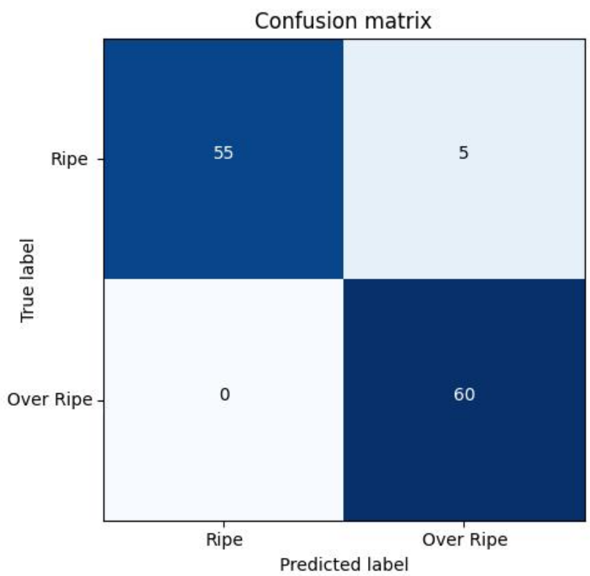



Fig. 10. Confusion Matrix for Model Evaluation of Model 1

Fig. 8 on the other hand shows that 2 ripe watermelon sounds were incorrectly classified as over ripe having 58 correctly classified ripe watermelons. Seven (7) out of 60 over ripe watermelons were also incorrectly classified for the over ripe class making only 53 correct predictions for the said class.

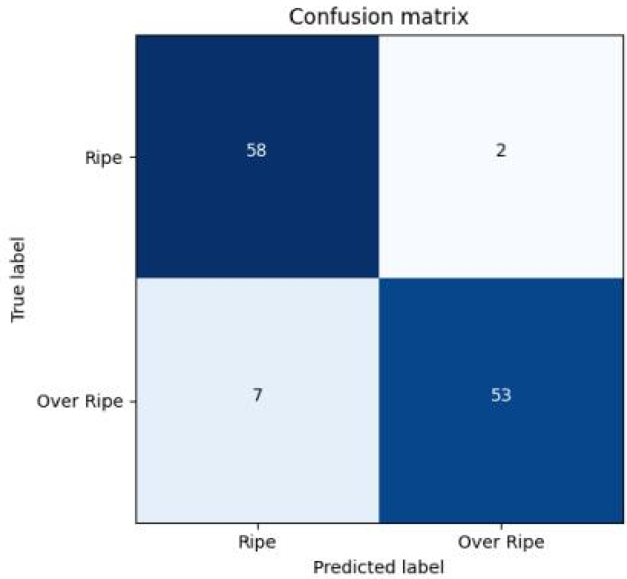



Fig. 11. Confusion Matrix for Model Evaluation of Model 2

The x-axis of the confusion matrix represents the predicted classes done by the model while the y-axis refers to the actual classes of the sound file.

### C. *Model Evaluation*

The generated models were evaluated with PyCM's statistics. Among the list of statistics, accuracy, AUC, and AUC value interpretation for both classes are listed in Table III and Table IV. It can be seen that the generated model successfully classifies watermelon as fresh or old watermelon depending on the sound waves it produces upon tapping. It can be seen that the AUC value interpretation mentions Excellent for both classes.

TABLE III.: MODEL 1: CLASS STATISTICS FOR MODEL EVALUATION

| Class Statistic | Table Column Head | |
|---|---|---|
| | ***Ripe (0)*** | ***Over Ripe (1)*** |
| Accuracy (ACC) | 0.9833 | 0.9833 |
| Adjusted F-1 Score | 0.93485 | 0.98287 |
| Area Under the ROC Curve (AUC) | 0.9833 | 0.9833 |
| AUC Value Interpretation | Excellent | Excellent |

Table III shows that the accuracy for both ripe and over ripe class amounted to 98.33%. On the other hand, the F-1 Score for the ripe class is 93.485% and 0.98287 for over ripe. The Area Under the ROC Curve (AUC) for both classes is 98.33% and has a corresponding AUC value interpretation of *Excellent.*

TABLE IV: CLASS STATISTICS

| Class Statistic | Table Column Head | |
|---|---|---|
| | ***Ripe (0)*** | ***Over Ripe (1)*** |
| Accuracy (ACC) | 0.9250 | 0.9250 |
| Adjusted F-1 Score | 0.94862 | 0.90227 |
| Area Under the ROC Curve | 0.9250 | 0.9250 |
| AUC Value Interpretation | Excellent | Excellent |

Table IV on the other hand shows that the accuracy for both ripe and over ripe class is 92.50%. Furthermore, the F-1 Score for the two classes are 94.862% and 90.227% respectively. The Area Under the ROC Curve (AUC) for both classes is 92.50% and has a corresponding AUC value interpretation of *Excellent.*

*D. Model Inference*

After model evaluation, both models were feeded to the device as both models have an AUC value interpretation of Excellent. Watermelons were used for evaluating the performance of the developed device. These watermelons consists of freshly harvested watermelons for the ripe class and watermelons a month after harvest for the class of over ripe.

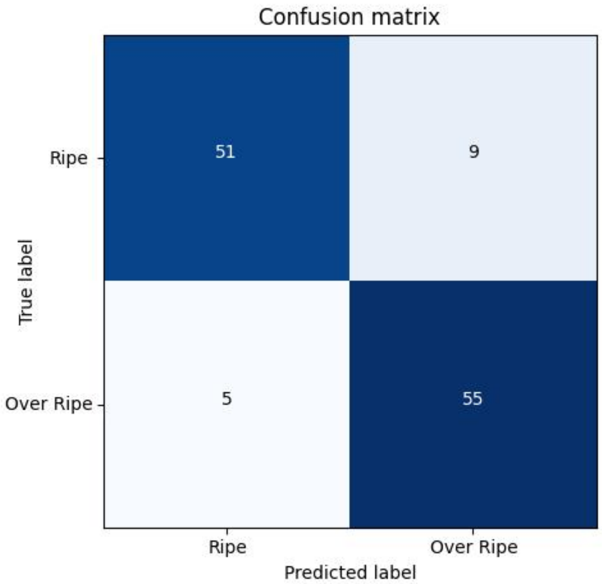



Fig. 12. Confusion Matrix for Model 1 Inference

Fifty-one taps were correctly classified as ripe watermelons out of 60 taps. On the other hand, 55 were correctly classified as over ripe.

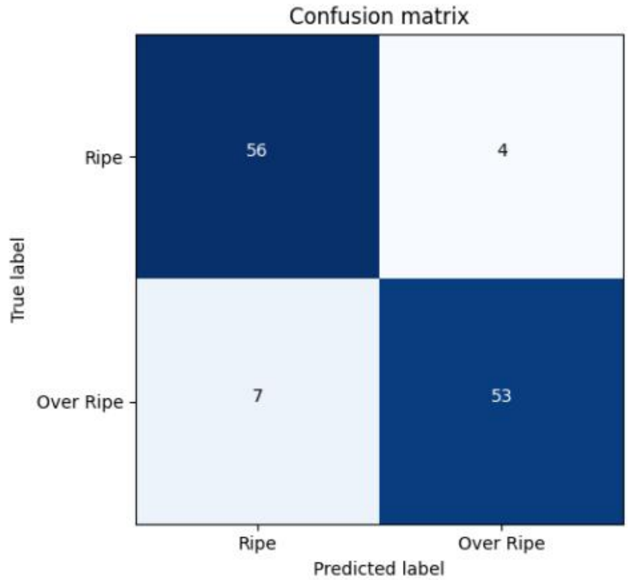



Fig. 13. Confusion Matrix for Model 2 Inference

For the model inferencing in the second model which was generated with FNN, 56 were correctly classified as ripe while 53 were classified as over ripe.

Along with the initial model inferencing conducted, several testing were also conducted to further assess the performance of the device. A total of two additional testings were conducted which both took place in an uncontrolled environment.

The first testing, which was conducted in a setup with a minimal to no environmental noise yielded the following results which was listed in Table V. On watermelons that were both ripe and overripe, 100 taps were made. On the ripe watermelon, 92 out of 100 taps indicated that the fruit was in the projected stage of ripeness. On the other side, 91 taps indicated that the overripe class was correctly predicted.

TABLE V: MODEL INFERENCING RESULT CONDUCTED IN QUIET ENVIRONMENT.

| Class | Predicted Ripe | Predicted Over Ripe |
|---|---|---|
| Actual Ripe | 92 | 8 |
| Actual Over Ripe | 9 | 91 |

The setting for the second test was carried out in the same room as the test for the calm environment. However, this time a loud noise was part of the setup, which was produced by playing a background noise within the testing room, in contrast to the previous testing where little to no noise was present. The ripe and over ripe watermelons were tapped 100 times each, just like in the previous simulation. However, the findings revealed that out of the 200 attempts at tapping, every single one was identified as ripe watermelon.

TABLE VI: MODEL INFERENCING RESULT CONDUCTED IN NOISY ENVIRONMENT.

| Class | Predicted Ripe | Predicted Over Ripe |
|---|---|---|
| Actual Ripe | 100 | 0 |
| Actual Over Ripe | 100 | 0 |

Several conclusions were drawn from the outcomes of the development of the device that could assess the ripeness of watermelon using machine learning.

As the microphone created static sounds when positioned close to other components, such the solenoid, the placement of the parts had an impact on how well the system worked.

Different architectures were used to implement MFCC for machine learning in this study. Depending on the architecture used during the procedure, the produced model's performance changed.

## IV. CONCLUSION

It is important to determine how fresh a good is to avoid consuming spoiled products which may pose health risks. For watermelon, where the internal condition is not discernable based only on the outside appearance, a ripeness determination model is generated based on the sound produced by the watermelon upon thumping.

The generated audio from the device is then saved into a folder within the device which is then processed. Model 2 which was generated with the use of FNN performed better when model evaluation or model inferencing was conducted, with a performance accuracy of 90.833%, which is higher compared to the testing accuracy of the first model with a performance accuracy of 88.333%.

Model 1 on the other hand, which was generated with the use of CNN, performed better when model evaluation or model inferencing was conducted, with a performance accuracy of 92.50%.

The setting in which the item was tested had an impact on its performance. In contrast, when tested in a noisy environment, the device failed to accurately forecast when a watermelon would be ripe.

This implies that FNN is an effective algorithm to be used for determining the freshness of watermelon. It also shows that within the various usage of Librosa, MFCC is a powerful tool to use for determining the features.

For future works, the researchers recommend to use watermelons with different characteristics such as variant and size for the dataset. It is also recommended to improve the device by considering an implementation of a controlled environment

## ACKNOWLEDGMENT

The researchers gratefully acknowledge their friends and families who showed their utmost support. They also acknowledge the Computer Engineering Faculty of Batangas State University for giving advices and ideas for the improvement of the study, and most importantly, the institute they belong to.

## REFERENCES

[1] Äijälä, E., Managing the intricacies of fresh fruits and vegetables in grocery retail, RELEX Solutions, 2023. Available at: https://www.relexsolutions.com/resources/managing-the-intricacies-of-fresh-fruits-and-vegetables-in-grocery-retail/ (Accessed: February 26, 2023).

[2] J. Bai, "Melons as Lemons: Asymmetric Information, Consumer Learning and Quality Provision," Wp, no. July, 2018, [Online]. Available: https://drive.google.com/file/d/0B52sohAPtnAWYVhBYm11cDBrSmM/view.

[3] V. Meshram, K. Patil, V. Meshram, D. Hanchate, and S. D. Ramkteke, "Machine learning in agriculture domain: A state-of-art survey," Artif. Intell. Life Sci., vol. 1, no. October, p. 100010, 2021, doi: 10.1016/j.ailsci.2021.100010.

[4] A. B. Kuriakose, "Automatic Fruit Classification and Freshness Detection," vol. 9, no. 13, pp. 52–55, 2021, [Online]. Available: www.ijert.org.

[5] El-Mesery, H. S., Mao, H. and Abomohra, A. E. F. (2019) 'Applications of non-destructive technologies for agricultural and food products quality inspection', Sensors (Switzerland), 19(4), pp. 1–23. doi: 10.3390/s19040846.

[6] G. Vorobyev, D. A. Subbotin, and E. Losevskaya, "Watermelon Quality Determining from a Photo Using Machine Vision," Proc. 2021 IEEE Conf. Russ. Young Res. Electr. Electron. Eng. ElConRus 2021, pp. 2286–2289, 2021, doi: 10.1109/ElConRus51938.2021.9396492.

[7] W. Daosawang, K. Wongkalasin, and N. Katewongsa, "A Study Sound Absorption for Ripeness and Unripe Classification of Watermelon," 2020 8th Int. Electr. Eng. Congr. iEECON 2020, pp. 10–13, 2020, doi: 10.1109/iEECON48109.2020.229521.

[8] W. A. Pamungkas and N. Bintoro, "Evaluation of watermelon ripeness using self-developed ripening detector," IOP Conf. Ser. Earth Environ. Sci., vol. 653, no. 1, 2021, doi: 10.1088/1755-1315/653/1/012020.

[9] R. Nino-Esparza, T. Cameron, J. Sampson, M. Thibodeau, T. Cayaoyao, and A. Ahmadi, "The use of a VOC Sensor to Measure Freshness of Fruits," pp. 1–4, 2019.

[10] Fujifilm (2021) 'What Is Non-Destructive Testing (NDT)? | Fujifilm NDT Guide'. Available at: https://ndtblog-us.fujifilm.com/blog/what-is-non-destructive-testing/.

[11] Bratu, A. M. et al. (2021) 'Non-destructive methods for fruit quality evaluation', Scientific Reports, 11(1), pp. 1–15. doi: 10.1038/s41598-021-87530-2.

[12] Albert-Weiß, D. et al. (2020) 'CNN for Ripeness Classification of Watermelon Fruits Based on Acoustic Testing', Proceedings of the Virtual 3rd International Symposium on Structural Health Monitoring and Nondestructive Testing (SHM-NDT 2020), pp. 25–26. Available at: http://www.ndt.net/?id=25569.

[13] J. Shrefler, L. Brandenberger, E. Rebek, J. Damicone, and M. Taylor, "Watermelon Production | Oklahoma State University," 2017. https://extension.okstate.edu/fact-sheets/watermelon-production-2.html (accessed May 22, 2022).

[14] S. Olsen, D. Drost, and T. Hunsaker, "Harvest and Storage of Vegetables and Fruits | USU." https://extension.usu.edu/yardandgarden/research/harvest-and-storage-of-vegetables-and-fruits (accessed May 26, 2022).

[15] D. Jie, W. Zhou, and X. Wei, "Nondestructive detection of maturity of watermelon by spectral characteristic using NIR diffuse transmittance technique," Sci. Hortic. (Amsterdam)., vol. 257, no. May, p. 108718, 2019, doi: 10.1016/j.scienta.2019.108718.

[16] Z. Xiaobo, Z. Junjun, H. Xiaowei, Z. Kaiyi, and W. Shengbin, "Distinguishing watermelon maturity based on acoustic characteristics and near infrared spectroscopy fusion technology," Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric. Eng., vol. 35, no. 9, pp. 301–307, 2019, doi: 10.11975/j.issn.1002-6819.2019.09.036.

[17] C. Ding, Z. Feng, D. Wang, D. Cui, and W. Li, "Acoustic vibration technology: Toward a promising fruit quality detection method," Compr. Rev. Food Sci. Food Saf., vol. 20, no. 2, pp. 1655–1680, 2021, doi: 10.1111/1541-4337.12722.

[18] J. Rudich and E. Zamski, "Citrullus Lanatus," CRC Handbook of Flowering. pp. 272–274, 2020, doi: 10.1201/9781351072540-37.

[19] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016, [Online]. Available: http://arxiv.org/abs/1603.04467.

[20] B. McFee et al., "librosa: Audio and Music Signal Analysis in Python," Proc. 14th Python Sci. Conf., no. Scipy, pp. 18–24, 2015, doi: 10.25080/majora-7b98e3ed-003.

[21] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way | by Sumit Saha | Towards Data Science," Towards data Science. pp. 1–12, 2018, [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.

[22] M. Mandal, "CNN for Deep Learning | Convolutional Neural Networks," Analytics Vidhya. 2022, [Online]. Available: https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/.

[23] Prashant, S. (2022) 'Feedforward Neural Network: Its Layers, Functions, and Importance', Analytics Vidhya. Available at: https://www.analyticsvidhya.com/.

[24] S. Haghighi, M. Jasemi, S. Hessabi, and A. Zolanvari, "PyCM: Multiclass confusion matrix library in Python," J. Open Source Softw., vol. 3, no. 25, p. 729, 2018, doi: 10.21105/joss.00729