International Journal of Applied Science and Engineering

Enhanced non-destructive of degree of pineapple juiciness using ensemble learning model based on tapping sound sensing

Suphachai Phawiakkharakun, Sunee Pongpinigpinyo*

Department of Computing, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand

ABSTRACT

This research proposes to enhance the non-destructive method for classifying the juiciness of pineapples using tapping sound sensing. Ten statistical features were extracted from the waveform signals by waveform analysis. These features were then separated into a spectral feature set (3 features) and a temporal feature set (7 features). Each feature set was calculated with the weight of important features and selected features for 15 training datasets using 10 machine learning classifiers. Ten machine learning classifiers were Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Machine (GBM), Extreme Gradient Boosting (XGB), K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Ensemble Voting, Adaboost, Ensemble Bagging, and Ensemble Stacking. The classifiers were evaluated with accuracy and the kappa coefficient. Grid search was used to determine various important hyperparameters for Machine Learning classifiers. The experiment results showed that the Ensemble Voting (soft), Ensemble Stacking, and MLP outperformed other classifiers. They can obtain an accuracy of 92.08%, and kappa coefficients are 0.8811, 0.8808 and 0.8808, respectively.

Keywords: Tapping sound sensing, Ensemble learning, Pineapple juiciness, Waveform signal, Non-destructive quality.

OPEN ACCESS

Received: September 20, 2023
Revised: October 25, 2023
Accepted: November 8, 2023

Corresponding Author:
Sunee Pongpinigpinyo
pongpinigpinyo_s@su.ac.th

Publisher:
Chaoyang University of Technology
ISSN: 1727-2394 (Print)
ISSN: 1727-7841 (Online)

1. INTRODUCTION

The pineapple (Anamas comosus) is one of the most well-known tropical fruits. Pineapple is classified as a non-climacteric fruit. It does not continue ripening after being harvested. As a result, juiciness is an essential characteristic of edible pineapple and is useful for exporters and consumers who want to sort pineapples based on their ripeness and juiciness. Unfortunately, it is difficult to ascertain the degree of juiciness in pineapple using traditional non-destructive detection techniques.

Traditionally, farmers and sellers used various non-destructive techniques to determine the juiciness degree of pineapple. Pineapple farmers or merchants typically employ a traditional technique of tapping the fruit skin with a rubber-tipped stick or their middle fingernails using force impulse techniques to hear the sound to judge pineapple juiciness. However, the conventional classification method calls for years of training or practice. Due to personal judgment, classification errors can happen quickly using traditional approaches, resulting in low classification accuracy. Notably, a human's ability to detect cannot accurately determine the juiciness level of pineapple.

Various image processing techniques have been extensively employed in previous evaluation studies. They have been employed in the fields of medical for the critical classification of lymphoblastic cancer (Saeed et al., 2023) as well as for the detection of objects in remote sensing imaging (Wang et al., 2022). Furthermore, the agricultural sector has applied these techniques to assess the quality of fruits (Azman and Ismail, 2017; Chaikaew et al., 2019). However, it is important to note that assessing the external appearance of fruit may be inaccurate due to potential damage from the external environment.

Acoustic signals have been the subject of numerous researches that have attempted to classify the quality and maturity of various fruits such as durian (Kharamat et al., 2020), coconut (Rahmawati et al., 2019; Caladcad et al., 2020; Fadchar and Cruz, 2020), cacao (Bueno et al., 2020), watermelon (Chawgien and Kiattisin, 2021), pistachio nut (Hosseinpour et al., 2022), and pineapple (Huang et al., 2022; Phawiakkharakun et al., 2022).

Additionally, several studies applied the acoustic signal to assess the internal quality of fruits like apples (Lashgari et al., 2020; Ekramirad et al., 2021; Zhao et al., 2021), pears (Zhang et al., 2021), and wheat (Yang et al., 2021). Furthermore, researchers have adapted the machine learning using the statistical features to parameters of 11 features of time domain, 7 features of the frequency domain, and 18 features of the combined feature set to identify early core browning in pear fruit. In each feature set, the minimal number of features was determined using the distance evaluation approach. As a result, the browning classifier achieved an accuracy of 93.90% using only three time-domain features (shape factor, kurtosis and square root amplitude value) and one frequency domain feature (variance). On the other hand, the classifier for slight browning achieved an accuracy of 86.40% overall with two time-domain features (shape factor and clearance factor) and one frequency-domain feature (mean square) (Zhang et al., 2021). When applying digital signal processing, natural frequency, spectrum entropy, and zero-crossing were the three most effective components of the carob moth diagnostic method in pomegranate fruit. A classification accuracy of 97.55% was achieved using these three features. (Janati et al., 2022). To enable real-time evaluation of kiwifruit firmness, 10 features from the frequency domain data were extracted using statistical characteristics. The root mean square, energy, means, spectral centroid of magnitude, and reference firmness indices all showed strong relationships ($|R| > 0.7$). The CARS-PLS model's prediction accuracy for fresh firmness, stiffness, and skin firmness in external cross-validation sets yielded $R^2cv = 0.96$, 0.95 and 0.93, respectively (Tian et al., 2022). In addition, the industrial sector has also applied statistical features for machinery fault diagnosis (Hui et al., 2017; Lei et al., 2017), and utilizing machine learning methods for the purpose of diagnosing, treating, and overseeing cognitive rehabilitation in individuals with neurological disorders is a primary focus. EEGs are employed to monitor and analyze brain activity (Das et al., 2023).

Machine learning methods can be used to evaluate quality by analyzing data collected from sensing devices. This is a useful tool since it can be used to find flaws in products and raise their quality. (Nturambirwe and Opara, 2020). Therefore, high-dimensional data in machine learning issues is problematic, especially with numerous characteristics and extracting feature significance from these variables and data with a high dimension. The redundant data and noise in the dataset were removed using statistical methods. This is important because feature selection for models is crucial for classifying the phenotypes of colorectal cancer cases. Many different methods have been proposed for selecting the most important features in a dataset. (Cenggoro et al., 2019). The result showed that random forests (RF) have the highest performance in techniques for selecting features from RFs and extra trees for malware detection in ensemble classification (Gbenga et al., 2021). Additionally, it used RF, Boruta, and Recursive Feature Elimination (RFE) selection methods to select essential features and compare different machine learning for classification analysis. In all experiment groups, the RF algorithm outperformed other algorithms in terms of performance (Chen et al., 2020).

However, no prior research works have been proposed to perform the statistical features extracted from the waveform signals and use feature selection for pineapple juiciness classification on Machine Learning methods. Signal processing might be a more suitable choice compared to image processing due to the potential for errors resulting from adverse external factors. These external conditions can lead to inaccuracies in the assessment process. Therefore, this paper proposes a statical feature of acoustic sensing and 10 machine learning classifiers to classify the juiciness level of pineapple. The juiciness level is divided into three classes: Juiciness 1 is defined as the flattening sound that is produced when a pineapple is particularly juicy, sweet, and slightly acidic in flavor. Juiciness 2 is the dullness that results from fruit that is just a little bit juicy, sweet, and sour in flavor. Juiciness 3 is defined as the echo transmitted as a tympany sound when the pineapple is slightly less juicy and sweet and has a more acidic flavor. (Phawiakkharakun et al., 2022).

The primary aim of this research is to address the challenge of assessing the juiciness level in pineapples using non-invasive methods, particularly by analyzing audio waveforms generated through the detection of tapping sounds on pineapples. Traditionally, fruit quality assessment has relied on the subjective personal experience judgments, which can be time inefficiencies and lead to inconsistencies and inaccuracies. Our research proposes a modern approach that utilizes machine learning techniques to provide an objective and efficient solution.

The research makes several valuable contributions: 1) The study utilizes machine learning models to assess the juiciness of pineapples, demonstrating the practical application of these models in assessing fruit quality. 2) It highlights the importance of feature selection by highlighting key attributes that significantly affect classification accuracy, such as the form factor, crest factor, and root mean square. 3) The study demonstrates notable gains in accuracy when analyzing the effects of combining features from several datasets when compared to reference datasets. This suggests that model performance may be enhanced through data integration. 4) The investigation employs ensemble learning techniques, including ensemble layering and ensemble voting, to enhance classification accuracy effectively, improving the efficiency of machine learning models. 5) The study provides a comparative analysis of various machine learning models, demonstrating variations in their accuracy when classifying pineapple juiciness. 6) Applying machine learning to assess fruit quality, particularly pineapples, has practical implications for the fruit production and quality control industries. 7) Unlike prior research that relied on Convolutional Neural Networks (CNNs), this approach prioritizes accessibility and the ability to conclude the significance of feature engineering with ensemble learning, making it suitable for practical applications.

These collective contributions enhance the understanding and real-world utilization of machine learning for assessing fruit quality, with a specific focus on classifying the juiciness of pineapples. The structure of this study is as follows:

Section 2, we present a summary of pertinent research related to the assessment of fruit quality, the utilization of acoustic signals, and the utilization of machine learning approaches. We delve into foundational investigations explore related applications and highlight the gaps in the literature.

Section 3, we introduce the datasets and machine learning models used in our study. We elaborate on the features extracted from audio waveforms and the essential pre-processing steps required for effective model training.

Section 4, we showcase the results of our experiments, illustrating how different machine learning models perform on a range of datasets. We underscore the significance of feature selection, feature integration, and the utilization of ensemble learning to improve the accuracy of classification.

Section 5, we delve into a thorough examination of our results, emphasizing the importance of distinct attributes like crest factor, shape factor, and root mean square in the classification process. Additionally, we explore the advantages of utilizing ensemble classification techniques to enhance the precision of our classifications.

Section 6 encapsulates the findings of our research and their significance in the context of quality assessment. We emphasize the disruptive nature of our methodology and its potential for wider adoption within the agricultural post-harvest industry.

2. MATERIALS AND METHODS

The experimental process for classifying the juiciness of Sriracha pineapples. The process begins with Step 1, which involves acquiring data from 30 pineapple samples using a mobile phone. In Step 2, 1,200 audio waveforms are input and divided into three classes based on juiciness levels (Juiciness 1, Juiciness 2 and Juiciness 3). Step 3 involves audio preprocessing, where the spectral centroid, spectral bandwidth, spectral roll-off, crest factor, skewness, zero-crossing rate, root mean square, impulse factor, shape factor, and kurtosis are extracted as features from the waveform.

In Step 4, a score of feature importance values is calculated, and the dataset is prepared by selecting features based on their rank of importance. Step 6 employs an evaluation model comprising a SVM, RF, GBM, XGB, KNN, Multi-Layer Perceptron (MLP), Ensemble Voting, Adaboost, Ensemble Bagging, and Ensemble Stacking, with accuracy and kappa coefficient serving as the metrics. Finally, in step 7, the classification results are visualized and compared among the different models as shown in Fig. 1.

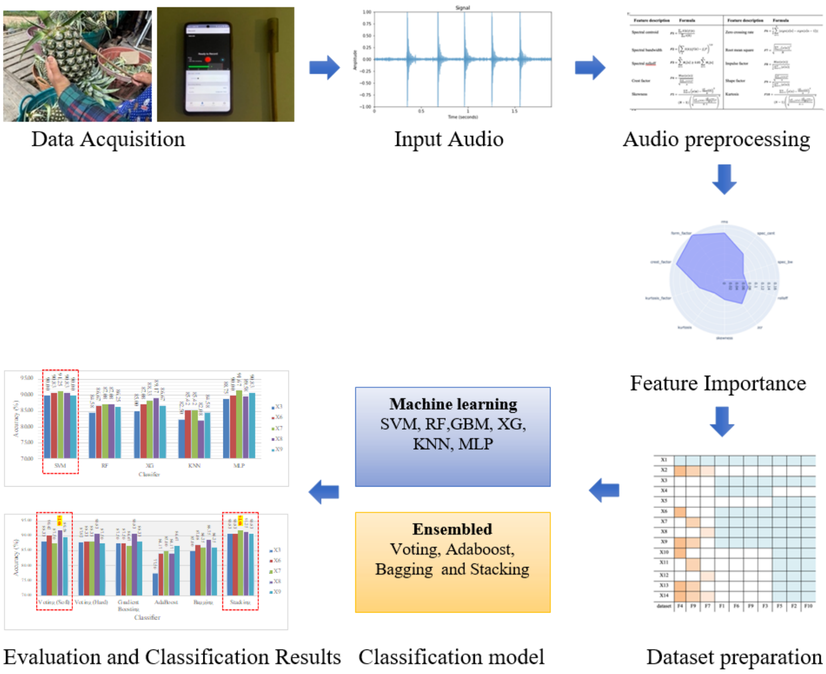
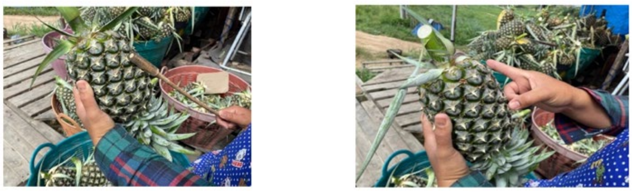


Fig. 1. Sriracha pineapple juiciness classification experimental process

2.1. Dataset Preparation

Thirty Sriracha pineapples were collected and tapped with a rubber-tipped stick. The tapping sound was recorded by using Motiv audio software in real environments (Phawiakkharakun et al., 2022). This non-destructive recording approach is based on the technique used by pineapple sellers and pineapple farmers who tap on pineapple samples to identify their juiciness. There are two methods of tapping: (a) employing a stick with a rubber tip and (b) utilizing the middle fingernail of an individual as shown in Fig. 2.



(a) using a rubber-tipped stick (b) using a person's middle fingernail

Fig. 2. The process of assessing the level of pineapple's juiciness

Samples of pineapple were divided by pineapple connoisseurs or pineapple farmers. Juiciness 1 was separated into three groups of identical size, followed by Juiciness 2 and Juiciness 3 (Fig. 3(a)). Consequently, ten pineapples of each level of juiciness are included in each dataset. Each level of the pineapple's juiciness is shown in Fig. 3(b).

Juiciness 1 Juiciness 2 Juiciness 3

(a) Pineapples samples

Juiciness 1 Juiciness 2 Juiciness 3

(b) juiciness level

Fig. 3. The pre-classified the harvested pineapple fruits

The impact response technique, the rubber-tipped stick, is used to process the acoustic data. A sampling rate of 44,100 Hz and a bit-depth of 16 bits per sample were obtained for the mono audio. Each pineapple was tested by tapping it five times, resulting in a total of 40 audio waveform files. In total, 1,200 audio waveform files were created, which were categorized into three groups: Juiciness 1 and Juiciness 2 each have 400 audio waveform files, and Juiciness 3 has 400 audio waveform files.

2.2. Data Exploration

Our dataset consists of 1,200 audio waveform files, out of which 1,098 files (equivalent to 91.5% of the total) have a length of less than 3 seconds. The remaining 102 audio waveform files have a period of time longer than three seconds, constituting 8.5% of the total files and the density of the data set varies within the range of 1.5–3.0 seconds.

The audio signal contains noise, such as variations in the duration of the first tap produced by a rubber-tipped stick hitting a pineapple. Additionally, ambient noise from human, animal, machine, and engine sources. These factors contribute to the presence of noise or errors in each audio signal, resulting from data acquisition in an uncontrolled environment.

2.3. Audio Pre-processing

Each audio signal was used to extract features. The spectral features consist of three representative statistical attributes, namely, the spectral centroid (F1), the spectral bandwidth (F2), and the spectral roll-off (F3). The Librosa library (McFee, 2015) was used in our research to extract the features from the spectral features.

In addition to these, it lists seven statistical features for the temporal features, namely, crest factor (F4), skewness (F5), zero-crossing rate (F6), root mean square (F7), impulse factor (F8), shape factor (F9) and kurtosis (F10) of the audio signal. The identification of juiciness classes was carried out using both spectral and temporal feature datasets as shown in Table 1.

An excessive amount of information and irrelevant data can lead to bias, which can affect the accuracy of machine learning outcomes. Therefore, it is essential to focus on important features that have a significant impact on machine learning. Feature importance approaches are employed for the computation of a score across all input features in a machine learning model. The scores provide an indication of the "significance" of each feature, where a higher score implies that the feature will exert a more substantial influence on the model's parameters.

Our research analyzes features to identify which ones are effective in classifying data. RFs were used in these experiments. The top three features identified were crest factor, shape factor, and root mean square, with obtained feature importance values of 0.1796, 0.1661, and 0.1341, respectively. The feature importance values for spectral centroid, crest factor, impulse factor, spectral roll-off, skewness, spectral bandwidth, and kurtosis were 0.0924, 0.0881, 0.0862, 0.0832, 0.0633, 0.0547 and 0.0522, respectively.

The dataset comprises 15 sets of experiments. X1 includes all features (F1-F10), X2 includes crest factor (F4), shape factor (F9), and root mean square (F7) (X1 is a dataset with feature importance values greater than 0.1). X3 includes spectral centroid (F1), crest factor (F6), shape factor (F9), spectral roll-off (F3), skewness (F5), spectral bandwidth (F2) and kurtosis (F10) (X3 is a dataset with

Table 1. Features description and formula of spectral and temporal features

Feature description	Formula	Feature description	Formula
Spectral centroid[1]	$\boldsymbol{F1} = \frac{\sum_k S(k)f(k)}{\sum_k s(k)}$	Zero crossing rate[3]	$\boldsymbol{F6} = \frac{1}{2}\sum_{n=1}^{N} \lvert sign(x[n] - sign(x[n-1]) \rvert$
Spectral bandwidth1	$\boldsymbol{F2} = (\sum_k S(k)(f(k) - f_c)^p)^{1/p}$	Root mean square[3]	$\boldsymbol{F7} = \sqrt{\frac{\sum_{n=1}^{N}(x(n))^2}{N}}$
Spectral rolloff[2]	$\boldsymbol{F3} = \sum_{n=1}^{R_t} M_t[n] \geq 0.85 \sum_{n=1}^{N} M_t[n]$	Impulse factor[3]	$\boldsymbol{F8} = \frac{Max\lvert x(n)\rvert}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$
Crest factor[3]	$\boldsymbol{F4} = \frac{Max\lvert x(n)\rvert}{\sqrt{\frac{\sum_{n=1}^{N}(x(n))^2}{N}}}$	Shape factor[3]	$\boldsymbol{F9} = \frac{\sqrt{\frac{\sum_{n=1}^{N}(x(n))^2}{N}}}{\frac{1}{N}\sum_{n=1}^{N}\lvert x(n)\rvert}$
Skewness[3]	$\boldsymbol{F5} = \frac{\sum_{n=1}^{N}\left(x(n) - \frac{\sum_{n=1}^{N} x(n)}{N}\right)^3}{(N-1)\left(\sqrt{\frac{\sum_{n=1}^{N}(x(n) - \frac{\sum_{n=1}^{N} x(n)}{N-1})^2}{N-1}}\right)^3}$	Kurtosis[3]	$\boldsymbol{F10} = \frac{\sum_{n=1}^{N}\left(x(n) - \frac{\sum_{n=1}^{N} x(n)}{N}\right)^4}{(N-1)\left(\sqrt{\frac{\sum_{n=1}^{N}(x(n) - \frac{\sum_{n=1}^{N} x(n)}{N-1})^2}{N-1}}\right)^4}$

[1] where the spectral magnitude at frequency bin k is denoted as S(k), the frequency at bin k is represented by f(k), and the spectral centroid is given by fc.. [2]where R_t is the roll-off frequency, and M_t is the magnitude of the n-th frequency component of the spectrum. [3]where x(n) is a signal series for n = 1,2, … N, and N is the number of data point

feature importance values less than 0.1). X4 includes spectral centroid (F1), crest factor (F6), shape factor (F9), and spectral roll-off (F3) (X4 is a dataset with feature importance values between 0.08 and 0.1). X5 includes skewness (F5), spectral bandwidth (F2), and kurtosis (F10) (X5 is a dataset with feature importance values between 0.6 and 0.8). Another dataset (X6-X14) combines the top three features (crest factor, shape factor, and root mean square) with X3 and X5. X15 comprises a combination of X2 and X4, eliminating the skewness, spectral bandwidth, and kurtosis features. To conduct the experiment, all datasets were split randomly into two groups, with 80% used to train the model and 20% used to test the performance of the model on unseen data, as shown in Fig. 4.

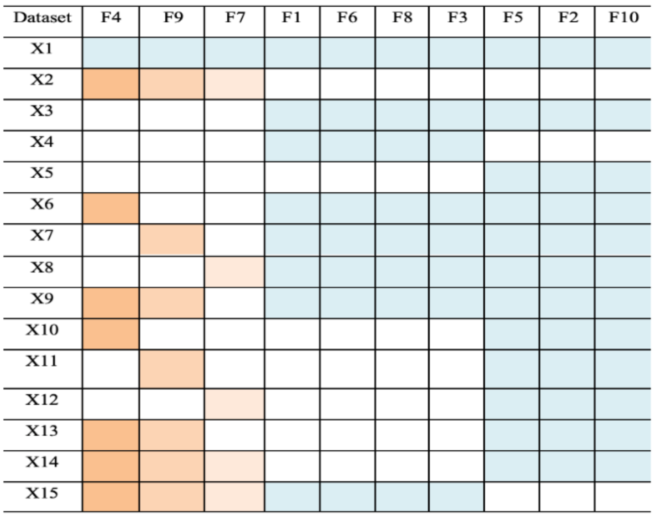


Fig. 4. The dataset features an experiment for input into the classifier

2.4 Ensemble Learning Model and Fine-Tune Training Configuration

The examined features for classification, including the 15 datasets and the target output (Juiciness class), were applied to the development of a variety of machine learning models. Some of the implemented machine learning models were SVM, RF, GBM, XGB, KNN, MLP, Ensemble Bagging, Adaboost, Ensemble Voting, and Ensemble Stacking. To carry out an optimized analysis, each ML model underwent hyperparameter fine-tuning using the tuning procedure.

We utilized grid search to determine various essential hyperparameters. Grid search is a way to find the best hyperparameters for a model by trying out all possible combinations of hyperparameter values within a specified range and then applying them to the learning process. The experimental results for the best parameter of each model consisted of the classification accuracy achieved when tuning hyperparameters via SVM, RF, GBM, XGB, KNN, MLP, bagging, Adaboost, and the optimal hyperparameters selected via grid search, presented in Table 2.

Ensemble voting implements both "hard" and "soft" voting. When using "hard" voting, it relies on predicted class labels to determine the majority vote. On the other hand, the class label is predicted by “soft” voting based on the argmax of the sums of the expected probability. The estimators used in ensemble voting are SVM, RF, GBM, XGB, KNN, MLP, Bagging, and Adaboost.

Ensemble stacking, on the other hand, utilizes a meta-learning algorithm that decides how to combine predictions from two or more fundamental machine learning techniques most effectively. Like ensemble voting, it also uses the estimators SVM, RF, GBM, XGB, KNN, MLP, Bagging, and Adaboost. However, the final estimator is logistic regression. Overall, ensemble methods are powerful techniques that can significantly enhance the robustness and accuracy of machine learning models. Derived from the study details in the research, the pseudocode framework for an Ensembled learning classification algorithm is presented in Algorithm 1.

Table 2. The selected optimal parameters using grid search

Model	Optimal parameters	Accuracy (%)
SVM	c = 50 gamma = 0.1 kernel = rbf max_depth = 5	91.56
RF	max_leaf_nodes = 30 n_estimators = 50	86.88
GBM	max_depth = 4 subsample = 0.5 learning_rate = 0.02 n_estimators = 500	88.25
XG	gamma = 0.0 learning_rate =. 0.05 max_depth = 5	88.75
KNN	n_neighbors = 4	86.35
MLP	activation = tanh alpha = 0.0001 hdden_layer = (10,30,10)	89.48
BAGGING	base_estimator = Decision Tree Classifier max_dept = 8 max_sample = 0.5 N_estimators = 50	86.88
ADABOOST	base_estimator = Decision Tree Classifier max_depth = 2 nearning_rate = 0.1 n_estimator = 100	85.94

Algorithm 1: Pseudocode of the ML algorithms

Input: Training and testing dataset
```
        X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size=0.2,
        random_state=1)
```
Output: Trained ML Classifier

BEGIN

Step 1: Initialize the ML estimator.

 # Initialize the machine learning
```
   [mlp,'MLPClassifier']
   [ensemble_voting1,'Ensembled VotingClassifier
(Soft) ']
   :
   :
   [stack,Ensembled Stacking  Classifier']
```
Step 2: Machine Learning and Ensemble Learning Classification

 # Initialize a list of models and their names
```
   algo = [
          [mlp,'MLPClassifier'],
          [ensemble_voting1,'Ensembled
VotingClassifier (Soft)'],
          :
          :
          [stack.'Ensembled StackingClassifier']
          ]
   model_scores = []
```
 # Define target names for classification report
```
   target_names = ['Juiciness 1', 'Juiciness 2', 'Juiciness
3']
```
 # Loop through the models and evaluate their performance
```
   for a in algo:
          model = a[0]  # Get the model
          model.fit(X_train, y_train)  # Train the model
          score = model.score(X_test, y_test)  #
Calculate accuracy
          y_pred = model.predict(X_test)  # Make
predictions
          cm= ConfusionMatrix(actual_vector=y_test,
predict_vector=y_pred)
```
 # Store the model's performance metrics in a list
```
   model_scores.append([score, cm.Overall_ACC,
cm.TPR_Macro, cm.TNR_Macro,
   cm.PPV_Macro, cm.F1_Macro, cm.Kappa, a[1]])
```
Step 3: Print parameter values and evaluation results.
```
    print (score) # Print the accuracy score.
    print (confusion_matrix(y_test, y_pred))    # Print the
confusion matrix
    print (classification_report(y_test, y_pred,
target_names=target_names))
```
END

2.5. Evaluation

The optimal network with the best hyperparameters is selected and applied to the dataset, which is randomly split into 80% for the training dataset and 20% for the validation dataset. The validation precision and loss (error) are both recorded. The accuracy and Cohen's kappa are compared to the results of each model on the test dataset using the confusion matrix and classification report to determine the model's performance. As shown in Equation (1), accuracy denotes the ratio of accurately categorized pineapple samples to the overall count of pineapple samples.

$$\text{Accuracy} = \frac{\text{number of correctly classified samples}}{\text{total number of samples}} \times 100 \quad (1)$$

Cohen's kappa is a widely used statistical measure that determines the level of agreement between two important factors. It is frequently employed to assess how well a classification model performs the equation for computing

Cohen's kappa as shown in Equation (2).

$$K = \frac{p_o - p_e}{1 - p_e} \quad (2)$$

The kappa coefficient value is denoted by K, where po represents the overall accuracy of the model, and pe represents the degree of concurrence between the predictions of the model and the authentic class values that could happen randomly. The degree of agreement can be categorized based on the kappa coefficient value as follows: slight agreement if the value is less than or equal to 0.20. Fair agreement if the value is between 0.41 and 0.60. There is moderate agreement if the value is between 0.61 and 0.80. There is a strong and almost perfect agreement if the value is between 0.81 and 1.0.

3. RESULTS AND DISCUSSION

In this paper, the important scores for all input features were calculated. The analysis revealed that the crest factor had the highest feature importance value of 0.1796, followed by the shape factor with a value of 0.1661, and the root mean square with a value of 0.1341.

Notably, the crest factor, shape factor, and root mean square obtained feature importance value greater than 0.1. In the next step, we partitioned the features into 15 datasets, applied grid search to tune hyperparameters, and showed the best parameters for each of the 15 datasets (X1-X15). The experiments conducted in this paper demonstrate the performance of the machine learning model for classifying pineapple juiciness, as shown below:

1. The performance of the ML model for pineapple juiciness classification shows that the SVM and MLP achieve greater accuracy than other models, with the MLP model achieving the highest accuracy of 92.28% and a kappa coefficient of 0.8808 in the X15 dataset. The SVM model had a maximum accuracy of 91.25% and a kappa coefficient of 0.8681 in the X7 dataset, as shown in Table 3.
2. The ensemble learning model for pineapple juiciness classification shows that the ensembled voting (soft) and ensemble stacking methods achieve higher accuracy than other methods. Specifically, the ensembled voting (soft) model achieves the highest accuracy of 92.08% and a kappa coefficient of 0.8811 in the X8 dataset. Meanwhile, in the X7 dataset, the ensemble stacking model achieves the best accuracy of 92.08% and a kappa coefficient of 0.8808, as shown in Table 4.
3. Comparing the baseline datasets with the combined datasets shows that better accuracy values can be obtained. For instance, the X3 dataset comprises spectral centroid (F1), zero-crossing rate (F6), shape factor (F9), spectral roll-off (F3), skewness (F5), spectral bandwidth (F2), and kurtosis (F10), achieves the best accuracy value of 90.00% with the SVM model. The X6 dataset combines crest factor with a feature in X3; the X7 dataset combines shape factor with a feature in X3; the X8 dataset combines root mean square with a feature in X3; and the X9 dataset combines crest factor and shape factor with a feature in X3, achieve accuracies of 90.83%, 91.25%, and 90.83%, respectively. Moreover, other ML models can achieve higher accuracy than the baseline dataset (X3), as shown in Fig. 5.
4. The results indicate that both ensemble voting (soft) and ensemble stacking methods can achieve a high accuracy of 92.08% in the X8 and X7 datasets, as demonstrated in Fig. 6.

Table 3. The performance of ML models for pineapple juiciness classification

Classifier/ Dataset	SVM		RF		GBM		XGB		KNN		MLP	
	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa
X1	90.42	0.8559	85.83	0.7865	88.75	0.8308	89.58	0.8432	84.58	0.7676	91.67	0.8746
X2	83.33	0.7485	82.92	0.7424	85.00	0.7737	84.58	0.7676	82.08	0.7308	84.58	0.7674
X3	90.00	0.8498	84.58	0.7689	87.50	0.8123	85.00	0.7746	82.50	0.7362	88.75	0.8309
X4	88.33	0.8248	86.67	0.7997	87.50	0.8123	87.08	0.8060	86.25	0.7926	87.92	0.8181
X5	78.33	0.6747	77.50	0.6613	76.67	0.6491	75.42	0.6299	72.50	0.5844	76.25	0.6432
X6	90.83	0.8621	86.67	0.8000	87.50	0.8122	87.08	0.8058	85.42	0.7804	90.00	0.8496
X7	91.25	0.8681	87.08	0.8057	86.67	0.7992	88.33	0.8244	85.42	0.7803	91.67	0.8745
X8	90.83	0.8621	87.08	0.8061	90.83	0.8623	89.17	0.8374	82.08	0.7298	89.58	0.8435
X9	90.00	0.8493	86.25	0.7926	88.33	0.8246	86.67	0.7994	84.58	0.7676	90.83	0.8621
X10	86.25	0.7934	82.08	0.7308	82.50	0.7364	82.08	0.7302	78.75	0.6787	87.08	0.8054
X11	85.00	0.7738	85.42	0.7802	87.08	0.8053	85.42	0.7801	84.17	0.7611	83.33	0.7490
X12	86.67	0.8006	84.58	0.7688	86.25	0.7933	86.25	0.7935	77.50	0.6610	85.83	0.7874
X13	87.08	0.8051	85.83	0.7864	86.25	0.7926	85.42	0.7803	83.75	0.7544	86.25	0.7927
X14	87.92	0.8178	85.00	0.7739	87.92	0.8179	87.50	0.8115	84.58	0.7675	89.58	0.8428
X15	90.00	0.8496	87.50	0.8117	88.33	0.8244	88.75	0.8309	89.17	0.8369	92.08	0.8808

An orange highlight is the best performance of each input dataset compared with all models. (SVM, RF, GBM XGB, KNN, MLP, ensemble Voting (soft), ensemble Voting (hard), Adaboost, ensemble bagging, and ensemble stacking)
A blue highlight the performance of the model obtaining an accuracy greater than 90% of each dataset.

Table 4. The performance of Ensembled learning models for pineapple juiciness classification

Classifier/ Dataset	Voting (Soft)		Voting (Hard)		AdaBoost		Bagging		Stacking	
	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa	Accuracy (%)	Kappa
X1	90.00	0.8495	88.75	0.8307	87.50	0.8118	86.25	0.7927	91.67	0.8747
X2	84.17	0.7613	84.17	0.7610	82.92	0.7417	82.50	0.7362	84.58	0.7674
X3	88.33	0.8247	87.92	0.8183	77.50	0.6626	85.00	0.7746	90.83	0.8623
X4	90.42	0.8560	89.17	0.8369	79.58	0.6953	86.67	0.7992	88.75	0.8310
X5	76.25	0.6425	77.92	0.6678	75.83	0.6370	78.33	0.6733	76.67	0.6492
X6	90.42	0.8560	88.33	0.8243	84.17	0.7628	87.08	0.8055	90.83	0.8622
X7	87.50	0.8118	88.33	0.8241	85.00	0.7741	86.25	0.7926	92.08	0.8808
X8	92.08	0.8811	90.83	0.8622	84.17	0.7638	88.75	0.8307	91.25	0.8686
X9	89.58	0.8431	87.50	0.8117	86.67	0.7990	86.25	0.7926	90.83	0.8621
X10	84.58	0.7680	82.92	0.7428	80.83	0.7121	81.25	0.7177	85.42	0.7803
X11	86.67	0.7989	86.25	0.7925	83.75	0.7553	85.83	0.7863	84.17	0.7614
X12	87.50	0.8125	87.08	0.8060	78.75	0.6822	84.58	0.7684	87.92	0.8187
X13	85.83	0.7865	86.25	0.7925	87.08	0.8054	86.25	0.7924	85.83	0.7861
X14	87.92	0.8180	87.08	0.8051	87.08	0.8052	86.67	0.7987	86.67	0.7994
X15	90.42	0.8560	89.58	0.8431	85.42	0.7803	88.75	0.8302	90.42	0.8560

An orange highlight is the best performance of each input dataset compared with all models. (SVM, RF, GBM XGB, KNN, MLP, ensemble Voting (soft), ensemble Voting (hard), Adaboost, ensemble bagging, and ensemble stacking)
A blue highlight is the performance of the model obtaining an accuracy greater than 90% of each dataset.

5. The X5 dataset comprises skewness (F5), spectral bandwidth (F2), and kurtosis (F10), combined with the top three features (i.e., crest factor, shape factor, and root mean square) with a score of feature importance greater than 0.1. The X10 dataset includes the crest factor combined with a feature from X5, the X11 dataset includes the shape factor combined with a feature from X5, and the X12 dataset includes root mean square combined with a feature from X5. The X13 dataset comprises the crest factor and the shape factor combined with a feature from X5, and the X14 dataset comprises the crest factor, the shape factor, and root mean square combined with a feature from X15. The MLP model can achieve an accuracy of 76.25% in the baseline dataset (X5) and 89.58% in the combined dataset (X14), indicating an improvement of approximately 17.48% from the baseline dataset (X5). Other models can also outperform the baseline dataset, as shown in Fig. 7.
6. The performance of the ensemble learning model shows that ensemble stacking can achieve an accuracy of 76.67% in the X5 dataset and an accuracy of 87.92% in the X12 dataset, resulting in an accuracy improvement of approximately 14.67% from the baseline dataset (X5). Additionally, other models can achieve higher accuracy levels than the baseline dataset (X5), as demonstrated in Fig. 8.
7. The features of skewness, spectral bandwidth, and kurtosis were removed from the X15 dataset because their feature importance values were less than 0.07. As a result, the X15 dataset now includes the crest factor, the shape factor, root mean square, the spectral cent, the zero-crossing rate, the impulse factor, and the roll-off. When compared with the X1 dataset, which consists of all 10 features, the MLP model achieved an accuracy of 92.08% with the X15 dataset, while the X1 dataset could only obtain an accuracy of 91.67%, as shown in Fig. 9

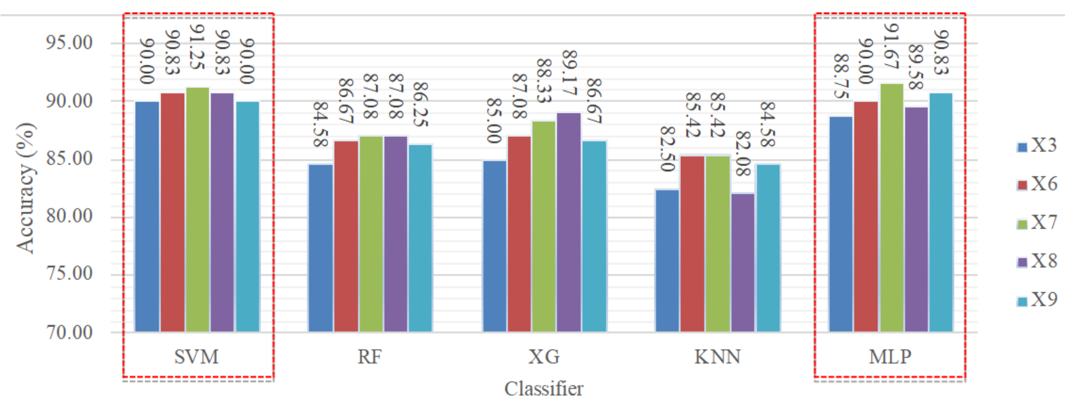


Fig. 5. The performance of the ML model on the baseline dataset and each combined dataset (X6–X9)

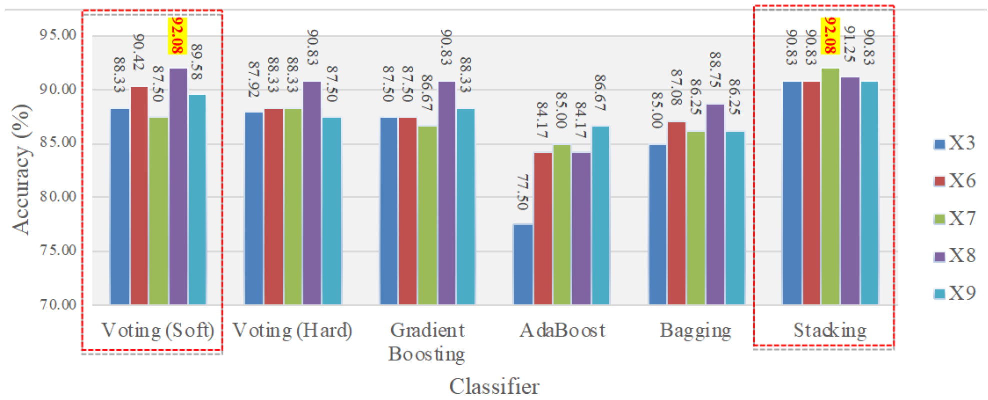


Fig. 6. The performance of the ensembled learning model on the baseline dataset (X3) and each combined dataset (X6–X9)

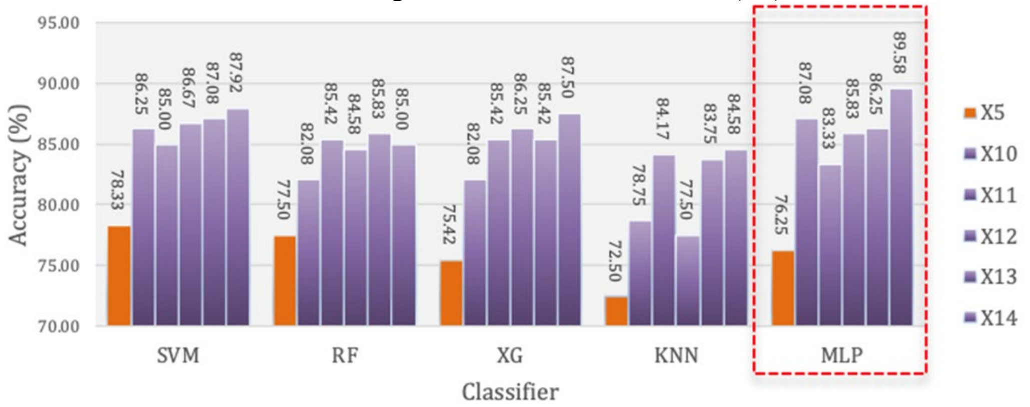


Fig. 7. The performance of the ML model on the baseline dataset (X5) and each combined dataset (X10–X14)

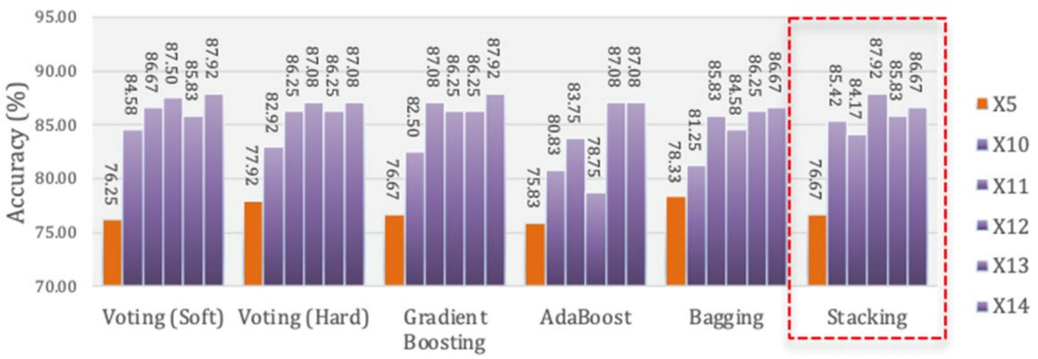


Fig. 8. The performance of the Ensembled learning model on the baseline dataset (X5) and each combined dataset (X10–X14)

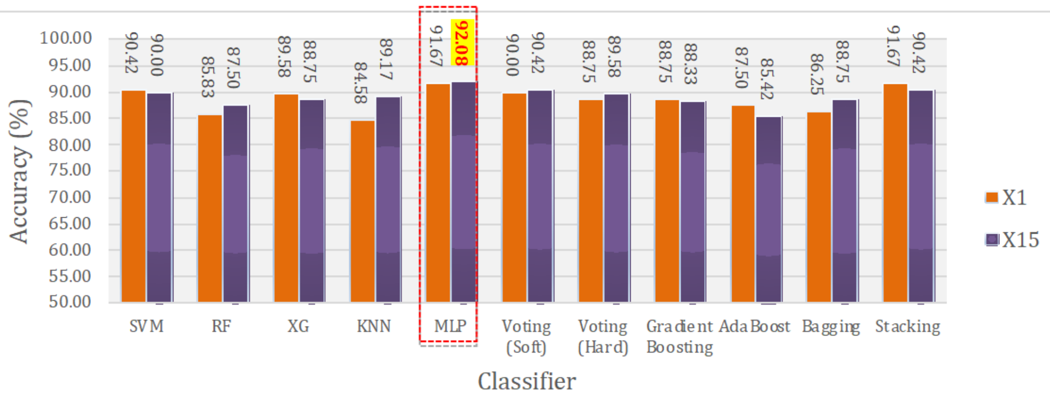


Fig. 9. The performance of each model on the base dataset (X1) and combined dataset (X15)

The primary objective of the study was to apply machine learning models to classify the degree of juiciness in pineapples. The study included processes such as feature selection, hyperparameter tuning, and a comprehensive evaluation of multiple models. This study presents the following significant findings and results:

1. Feature Importance Analysis: The calculation of feature importance scores included all input features, leading to the identification of three features with the most significant importance values: crest factor, form factor, and root mean square. The features mentioned above showed significance values exceeding 0.1, indicating their substantial contribution to the classification activity.
2. Impact of Combined Datasets: Integrating features from various datasets significantly improved accuracy compared to the reference datasets. The X3 dataset achieved the highest accuracy at 90.00% when used with the SVM model. Additionally, incorporating features from X3 into other datasets further enhanced accuracy. The MLP model showed a significant performance boost when applied to the combined dataset X14, reaching an accuracy of 89.58%. This represents a significant improvement compared to the baseline dataset X5.
3. Feature Selection: To improve the X15 dataset, features with significance values below 0.07 were specifically dropped. The reduced dataset, denoted as X15, exhibited a superior accuracy rate of 92.08% when used with the MLP model, compared to the original dataset, X1, which contained all ten features and achieved an accuracy rating of 91.67%.

In this discussion, the study highlights the significance of feature selection and combination, hyperparameter tuning, and the utilization of ensemble learning methods to enhance the classification accuracy of machine learning models for determining the degree of pineapple juiciness. The findings indicate that specific features, such as crest factor, shape factor, and root mean square, play a crucial role in the classification task. When combined with other relevant features, they result in significant improvements in accuracy. Additionally, ensemble methods like ensemble stacking and ensemble voting (soft) prove effective in enhancing classification accuracy. In comparison to the previous work by Phawiakkharakun et al. (2022) as shown in Table 5.

Our previous research (Phawiakkharakun et al., 2022) utilized a CNN model to evaluate the juiciness level of Sriracha pineapple. This was done by comparing the performance of two different feature extraction methods, namely Mel Frequency Cepstral Coefficient (MFCC) and Mel-Spectrogram, utilizing acoustic sensors in conjunction with CNN. The results from the experiments revealed that both CNN and MFCC outperformed other approaches, achieving an impressive accuracy of 96.67 percent. In contrast, the present study employs statistical features (such as spectral centroid, spectral bandwidth, spectral roll-off, crest factor, skewness, zero-crossing rate, root mean square, impulse factor, shape factor and kurtosis) extracted from audio recordings as input variables for machine learning in the task of categorizing juiciness. Even if this approach's accuracy was lower than that of the preceding work at 92.08 percent, it is crucial to note that the strategy utilized in the earlier work may have problems with feature interpretability. As a result, the current research approach holds distinct advantages in terms of accessibility and the potential for drawing meaningful inferences about the significance of feature engineering when combined with ensemble learning. This makes it a valuable choice for applying machine learning models.

Table 5. The comparison for pineapple juiciness classification

Model	Feature	Accuracy
Deep Learning (CNN)	- Mel Frequency Cepstral Coefficient (MFCC) - Mel-Spectrogram	MFCC combined with CNN performed the best, with an accuracy of 96.67%
Purpose Method (Ensemble Learning)	Ten statistical features including -spectral centroid (F1) -spectral bandwidth (F2) -spectral roll-off (F3) -crest factor (F4) -skewness (F5) -zero-crossing rate (F6) -root mean square (F7) -impulse factor (F8) -shape factor (F9) -kurtosis (F10)	Ensemble Voting (soft), Ensemble Stacking, and MLP outperformed other classifiers. They can obtain an accuracy of 92.08%

4. CONCLUSIONS

This research presents a non-invasive approach to assess the quality of Sriracha pineapples based on the audio waveforms sound. The method classifies the juiciness of the pineapple using ten features extracted from spectral and temporal analyses. These features were divided into 15 datasets based on their importance scores and fed into machine learning and ensemble learning models. In this study, we employed grid search to optimize the hyperparameters of nine machine learning models. Our experiments revealed that the ensemble voting (soft) method performed the best, achieving an accuracy of 92.08% and a Kappa coefficient of 0.8811 in the X8 dataset. The ensemble stacking and MLP methods both achieved an accuracy of 92.08%, with Kappa coefficients of 0.8088 in the X7 and X15 datasets, respectively.

There are multiple avenues to explore for future research. 1) Develop mobile applications that utilize this method to assess the pineapple juiciness in real-time, making it applicable to business applications. 2) Increasing the dataset with additional pineapple samples from various sources and environments to enhance the generalizability of the model.

3) Investigate feature engineering methods to find additional features or combinations that can increase accuracy. 4) Integrating acoustic sensing devices to automate the data collection process and provide immediate feedback to users. 5) Applying the methodology to assess the quality of other fruits or agricultural products, thereby expanding its potential impact on the agriculture and culinary sectors.

Future research has the potential to expand on the groundwork laid in this study and drive the domain of non-invasive quality evaluation through acoustic signals by tackling these challenges.

ACKNOWLEDGMENT

This research was supported by Scholarship of the Celebrations on the Auspicious Occasion of His Majesty the King's 70 the Birthday Anniversary of Ph.D. Degree Scholarship Project Doctoral Studies in Thailand.

REFERENCES

Azman, A.A., Ismail, F.S. 2017. Convolutional neural network for optimal pineapple harvesting. Journal of Electrical Engineering, 16, 1–4.

Bueno, G.E., Valenzuela, K.A., Arboleda, E.R. 2020. Maturity classification of cacao through spectrogram and convolutional neural network. Jurnal Teknologi dan Sistem Komputer, 8, 228–233.

Caladcad, J. A., Cabahug, S., Catamco, M. R., Villaceran, P. E., Cosgafa, L., Cabizares, K. N., Piedad, E., Jr. 2020. Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal. Computers and Electronics in Agriculture, 172, 105327.

Cenggoro, T.W., Mahesworo, B., Budiarto, A., Baurley, J., Suparyanto, T., Pardamean, B. 2019. Features importance in classification models for colorectal cancer cases phenotype in Indonesia. Procedia Computer Science, 157, 313–320.

Chaikaew, A., Thanavanich, T., Duangtang, P., Sriwanna, K., Jaikhang, W. 2019. Convolutional neural network for pineapple ripeness classification machine. Proceeding of the 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 16, 373–376.

Chawgien, K., Kiattisin, S. 2021. Machine learning techniques for classifying the sweetness of watermelon using acoustic signal and image processing. Computers and Electronics in Agriculture, 181, 105938.

Chen, R.-C., Dewi, C., Huang, S.-W. and Caraka, R.E. 2020. Selecting critical features for data classification based on machine learning methods. Journal of Big Data, 7, 52.

Das, S., Adhikary, A., Laghari, A.A., Mitra, S. 2023. Eldo-Care: EEG with kinect sensor based telehealthcare for the disabled and the elderly. Neuroscience Informatics, 100130.

Ekramirad, N., Khaled, A.Y., Parrish, C.A., Donohue, K.D., Villanueva, R.T., Adedeji, A. 2021. Development of pattern recognition and classification models for the detection of vibro-acoustic emissions from codling moth infested apples. Postharvest Biology and Technology, 181, 111633.

Fadchar, N.A., and Cruz, J.C.D. 2020. A non-destructive approach of young coconut maturity detection using acoustic vibration and neural network. Proceedings of 16th IEEE International Colloquium on Signal Processing and Its Applications (CSPA), 136–140.

Gbenga, F., Adetunmbi, A., Elohor, O. 2021. Towards optimization of malware detection using extra-tree and random forest feature selections on ensemble classifiers. International Journal of Recent Technology and Engineering, 9, 223–232.

Hosseinpour-Zarnaq, M., Omid, M., Taheri-Garavand, A., Nasiri, A., Mahmoudi, A. 2022. Acoustic signal-based deep learning approach for smart sorting of pistachio nuts. Postharvest Biology and Technology, 185, 111778.

Huang, T.-W., Bhat, S. A., Huang, N.-F., Chang, C.-Y., Chan, P.-C. and Elepano, A.R., 2022. Artificial intelligence-based real-time pineapple quality classification using acoustic spectroscopy. Agriculture, 12, 1–17.

Hui, K.H., Ooi, C.S., Lim, M.H., Leong, M.S., Al-Obaidi, S.M. 2017. An improved wrapper-based feature selection method for machinery fault diagnosis. PLOS ONE, 12, 1–10.

Janati, S., Mehdizadeh, S.A., Heydari, M. 2022. Designing, manufacturing, and evaluating the diagnostic system of carob moth in pomegranate fruit using digital signal processing. Computers and Electronics in Agriculture, 192, 106564.

Kharamat, W., Wongsaisuwan, M., Wattanamongkhol, N. 2020. Durian ripeness classification from the knocking sounds using convolutional neural network. Proceedings of the 8th International Electrical Engineering Congress (iEECON), 1–4.

Lashgari, M., Imanmehr, A., Tavakoli, H. 2020. Fusion of acoustic sensing and deep learning techniques for apple mealiness detection. Journal of Food Science and Technology, 57, 2233–2240.

Lei, Y., He, Z., Zi, Y., Hu, Q. 2007. Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs. Mechanical Systems and Signal Processing, 21, 2280–2294.

McFee, B., Raffel, C., Liang, D., Ellis, D.P., McVicar, M., Battenberg, E., Nieto, O. 2015. Librosa: Audio and music signal analysis in python. Proceedings of the 14th Python in Science Conference, 8, 18–25.

Nturambirwe, J.F.I., Opara, U.L. 2020. Machine learning applications to non-destructive defect detection in horticultural products. Biosystems Engineering, 189, 60–83.

Phawiakkharakun, S., Taeprasartsit, P., Pongpinigpinyo, S. 2022. Acoustic sensing for quality edible evaluation of

sriracha pineapple using convolutional neural network. Current Applied Science and Technology, 22, 1–15.
Priasni, T., Oswari, T. 2021. Comparative study of standalone classifier and ensemble classifier. TELKOMNIKA (Telecommunication Computing Electronics and Control), 19, 1747–1754.
Rahmawati, D., Haryanto, H., Sakariya, F. 2019. The design of coconut maturity prediction. device with acoustic frequency detection using naive bayes method based microcontroller. Journal of Electrical Engineering, Mechatronic and Computer Science, 2, 15–20.
Saeed, U., Kumar, K., Khuhro, M.A., Laghari, A.A., Shaikh, A.A., Rai, A. 2023. DeepLeukNet—A CNN based microscopy adaptation model for acute lymphoblastic leukemia classification. Multimedia Tools and Applications, 1–25.
Tian, S., Wang, J., Xu, H., 2022. Firmness measurement of kiwifruit using a self-designed device based on acoustic vibration technology. Postharvest Biology and Technology, 187, 111851.
Wang, L., Shoulin, Y., Alyami, H., Laghari, A.A., Rashid, M., Almotiri, J., Alyamani, H.J., Alturise, F. 2022. A novel deep learning-based single shot multibox detector model for object detection in optical remote sensing images, 1–15.
Yang, X., Guo, M., Lyu, Q., Ma, M. 2021. Detection and classification of damaged wheat kernels based on progressive neural architecture search. Biosystems Engineering, 208, 176–185.
Zhang, H., Zha, Z., Kulasiri, D., Wu, J. 2021. Detection of early core browning in pears based on statistical features in vibro-acoustic signals. Food and Bioprocess Technology, 14, 887–897.
Zhao, K., Zha, Z., Li, H., Wu, J. 2021. Early detection of moldy apple core based on time-frequency images of vibro-acoustic signals. Postharvest Biology and Technology, 179, 111589.