Volume 14, Issue 3, Page 192-201, 2025 https://doi.org/10.46810/tdfd.1652908 Research Article

Türk Doğa ve Fen Dergisi
Turkish Journal of Nature and Science
www.dergipark.gov.tr/tdfd

Non-Destructive Prediction of Maturity from the Sound of Hand Hitting a Watermelon Using Machine Learning

Savaş KOÇ[1*], Ferit AKBALIK[2]

[1] Batman University, Faculty of Engineering and Architecture, Department of Mechanical Engineering, Batman, Türkiye
[2] Batman University, Vocational School of Social Sciences, Batman, Türkiye
Savaş KOÇ ORCID No: 0000-0002-5257-3287
Ferit AKBALIK ORCID No: 0000-0002-9791-1809

**Corresponding author: savas.koc @batman.edu.tr*

(Received: 06.03.2025, Accepted: 12.09.2023, Online Publication: 26.09.2025)

Keywords
Acoustic feature extraction,
Agricultural signal processing,
Maturity prediction,
Non-invasive fruit analysis,
Sound classification,
Supervised learning models

Abstract: Traditional methods for assessing the quality, taste, and maturity of fruits and vegetables without cutting rely on external attributes such as color, shape, surface patterns, and acoustic responses. In this study, maturity levels of watermelons were validated through physical inspection after cutting, and the corresponding acoustic data were analyzed using spectrograms, extracting 120 features each sample. Several machine learning algorithms, including Support Vector Classifier, Decision Tree Classifier, Random Forest Classifier (RFC), Multi-Layer Perceptron Classifier, and K-Nearest Neighbors Classifier (KNC), were evaluated. Among them, the KNC model achieved the highest accuracy (96.04%), followed closely by the RFC model (95.47%). Specifically, the RFC model classified mature watermelons with 98.2% accuracy, while the KNC model effectively distinguished overmature and immature samples with accuracies of 96.3% and 96.2%, respectively. Robustness against background noise was also demonstrated despite the naturally recorded dataset. The findings were compared with studies on acoustic pattern recognition in animals, environmental acoustics, and healthcare, highlighting the potential of machine learning-based approaches as a reliable, non-invasive method for maturity and taste assessment with practical applications in agriculture and the food industry.

Makine Öğrenimi Kullanarak Elle Vurulan Karpuzun Sesinden Tahribatsız Olgunluk Tahmini

Anahtar Kelimeler
Akustik özellik çıkarımı,
Gözetimli öğrenme modelleri,
Olgunluk tahmini,
Ses sınıflandırma,
Tahribatsız meyve analizi,
Tarımsal sinyal işleme

Öz: Meyve ve sebzelerin kesilmeden kalite, tat ve olgunluklarının değerlendirilmesinde geleneksel yöntemler genellikle renk, şekil, yüzey desenleri ve akustik tepkiler gibi dış özelliklere dayanır. Bu çalışmada, karpuzların olgunluk seviyeleri kesim sonrası fiziksel inceleme ile doğrulanmış, elde edilen akustik veriler spektrogramlarla analiz edilerek her örnek için 120 öznitelik çıkarılmıştır. Destek Vektör Sınıflandırıcısı (SVC), Karar Ağacı Sınıflandırıcısı, Rastgele Orman Sınıflandırıcısı (RFC), Çok Katmanlı Algılayıcı Sınıflandırıcısı (MLP) ve K-En Yakın Komşu Sınıflandırıcısı (KNC) algoritmaları değerlendirilmiştir. Bunlar arasında en yüksek doğruluğu %96,04 ile KNC modeli elde etmiş, onu %95,47 ile RFC modeli izlemiştir. Özellikle, RFC modeli olgun karpuzları %98,2 doğrulukla sınıflandırırken, KNC modeli aşırı olgun ve az olgun örnekleri sırasıyla %96,3 ve %96,2 doğrulukla ayırt etmiştir. Doğal ortamda kaydedilen veri setine rağmen arka plan gürültüsüne karşı da sağlamlık gösterilmiştir. Bulgular, hayvanlarda akustik desen tanıma, çevresel akustik ve sağlık alanındaki çalışmalarla karşılaştırılmış olup, makine öğrenmesine dayalı yaklaşımların olgunluk ve tat değerlendirmesinde güvenilir ve invazif olmayan bir yöntem sunduğu, tarım ve gıda endüstrisinde pratik uygulama potansiyeline sahip olduğu ortaya konmuştur.

1. INTRODUCTION

Agriculture is a cornerstone of global economic development and serves as a primary source of employment for impoverished populations worldwide [1]. The emergence of Smart Agriculture has introduced a transformative paradigm that utilizes Information and Communication Technologies (ICT) to address critical global challenges such as hunger, malnutrition, and food security. By incorporating advanced technologies, Smart Agriculture seeks to enhance crop yields, optimize cultivation practices, improve productivity, and ensure higher product quality [2]. Increasing consumer expectations for efficiency, reduced physical labor, and rapid outcomes have further accelerated the adoption of digital farming systems [3]. These innovations not only lower the dependency on traditional manual practices but also improve operational efficiency and generate new employment opportunities in the agricultural sector [4].

Watermelon (Citrullus lanatus), prized for its sweetness and high-water content, is among the most widely consumed fruits worldwide, primarily enjoyed fresh [5]. According to the Food and Agriculture Organization of the United Nations (FAO), nearly 70 million tons of watermelon are produced annually across approximately 100 countries and regions [6]. The quality of watermelon is influenced by both external characteristics, such as shape and size, and internal attributes, including maturity, pulp firmness, sugar concentration, water content, and internal defects. Among these, sweetness remains the most decisive factor shaping consumer preference and internal quality [7–8].

1.1. Non-Destructive Assessment of Watermelon Maturity

Traditionally, watermelon maturity has been evaluated using empirical methods such as tapping sounds, rind pattern inspection, and weight measurement. However, these approaches are often subjective and lack precision. Recent advancements in non-destructive technologies have introduced more reliable alternatives, including X-ray Computed Tomography (CT) [9], acoustic analysis [10], machine vision [11], dielectric property measurement [12], Hyperspectral Imaging (HSI) [13], and Surface-Enhanced Raman Spectroscopy (SERS) [14]. These methods support the development of regression and classification models for maturity prediction, leveraging acoustic and image processing for precise evaluation [6,15].

Non-destructive approaches have successfully assessed watermelon hardness and sweetness by integrating data from tapping sounds, rind features, and weights. Acoustic impulse methods, combined with linear models and Artificial Neural Networks (ANN), have achieved promising accuracy in classifying watermelon maturity [16]. Moreover, fusion techniques that combine acoustic and visual information have reported classification accuracies as high as 92% using gradient tree-based models [17]. These findings demonstrate the potential of integrating traditional practices with advanced technologies for improved agricultural outcomes.

The rapid progress of Artificial Intelligence (AI) has further transformed agricultural practices. AI methods such as neural networks, fuzzy logic, genetic algorithms, and machine learning are increasingly applied across domains including agriculture, healthcare, and robotics [18–20]. In sound-based analysis, approaches originally developed for medical applications such as heartbeat and respiratory monitoring are now being adapted to agriculture. These applications include classifying animal sounds, environmental noises, and fruit maturity [21–23]. For audio signal processing, representations such as spectrograms, Mel-Frequency Cepstral Coefficients (MFCC), and Zero-Crossing Rate (ZCR) are commonly employed. Advanced deep learning techniques, particularly Convolutional Neural Networks (CNN), enhance feature extraction from spectrograms and waveforms, significantly improving classification accuracy. Architectures such as VGGish and bottleneck CNN further strengthen feature representation through 1D and 2D convolutional layers, offering robust solutions for sound-based classification tasks [24–26].

Table 1. Non-destructive quality assessment methods

Studies	Technique	Advantages	Disadvantages
[27]	NIR Spectroscopy	Non-destructive, fast, accurate (>88%), can be portable	Expensive, sensitive to lighting/surface, needs calibration
[28]	Image Processing	Simple, low-cost, reasonably accurate (76–89%)	Surface-only info, lighting/angle dependent, rind pattern affects results
[29]	Acoustic Response	Easy, non-destructive, high accuracy (>90%)	Sensitive to tapping, noisy environments, some fruits must be cut
[30]	Microwave Imaging	Visualizes internal structure, high-resolution	Expensive, complex setup, limited field use
[31]	Portable NIR Spectrometer	Field-ready, practical accuracy	Requires per-variety calibration, sunlight may affect results, costly
[32]	Traditional/Visual Methods	Very simple, no equipment needed	Subjective, low accuracy, causes fruit waste

Recent studies have explored a variety of non-destructive techniques for assessing agricultural product quality, each with distinct advantages and limitations, as summarized in Table 1. Near-Infrared (NIR) Spectroscopy has been shown to provide fast, non-destructive, and accurate measurements (>88%), with potential for portable applications, though it remains costly and sensitive to lighting and surface conditions [27]. Image processing methods, including color and texture analysis, offer a simple and low-cost approach with reasonable accuracy (76–89%), yet they provide only surface-level information and are affected by lighting, camera angle, and rind patterns [28]. Acoustic response techniques, such as hand-tapping combined with MFCC analysis, are easy to implement, non-destructive, and achieve high accuracy (>90%), but their performance is influenced by tapping force, environmental noise, and occasionally requires cutting fruits for labeling [29]. Microwave imaging enables visualization of internal structure with high resolution, but its complexity, high cost, and limited field applicability restrict widespread use [30]. Portable NIR spectrometers offer practical, field-ready solutions, although calibration per variety and sunlight interference remain challenges [31]. Traditional visual methods, such as tapping and rind inspection, are simple and require no equipment, but they are subjective, less accurate, and can lead to fruit waste [32]. Despite these advances, scalable, non-destructive, and cost-effective quality evaluation methods remain a critical need. Combining traditional evaluation practices with modern machine learning algorithms holds promise for developing more accurate and objective assessment systems.

The primary aim of this study is to evaluate watermelon maturity in a non-destructive manner using machine learning. Unlike traditional methods that rely solely on tapping sounds or rind patterns, this study integrates traditional indicators with advanced acoustic signal processing and image analysis. The collected data are used to train classification models, including Multi-Layer Perceptron (MLP), Support Vector Classifier (SVC), Random Forest Classifier (RFC), and K-Nearest Neighbors Classifier (KNC), to identify the most accurate prediction approach.

Feature extraction focuses on acoustic representations such as mel-spectrograms, MFCC, and other sound-based attributes, which are subsequently analyzed using machine learning techniques. By bridging traditional practices with modern data-driven methods, this research aims to provide an efficient, scalable, and objective framework for agricultural quality assessment.

2. MATERIAL AND METHOD

The shape of a watermelon has traditionally been an important criterion for selection, particularly in markets and greengrocers. Experienced farmers often assess quality by tapping the watermelon and interpreting the resulting acoustic signals. To evaluate watermelon taste through these sounds, as illustrated in Figure 1, systematic sound recordings are collected during the data acquisition process. The tapping-generated acoustic signals from the outer surfaces of selected watermelons are recorded, after which the watermelons are cut and categorized into three groups: immature, mature, and overmature.

The recorded signals are then analyzed, and their features are extracted to serve as inputs for classification models. These models are trained on the extracted features to predict watermelon maturity and quality. The results highlight the correlation between tapping sounds and internal taste attributes, confirming the potential of sound-based methods as effective tools for agricultural quality assessment.

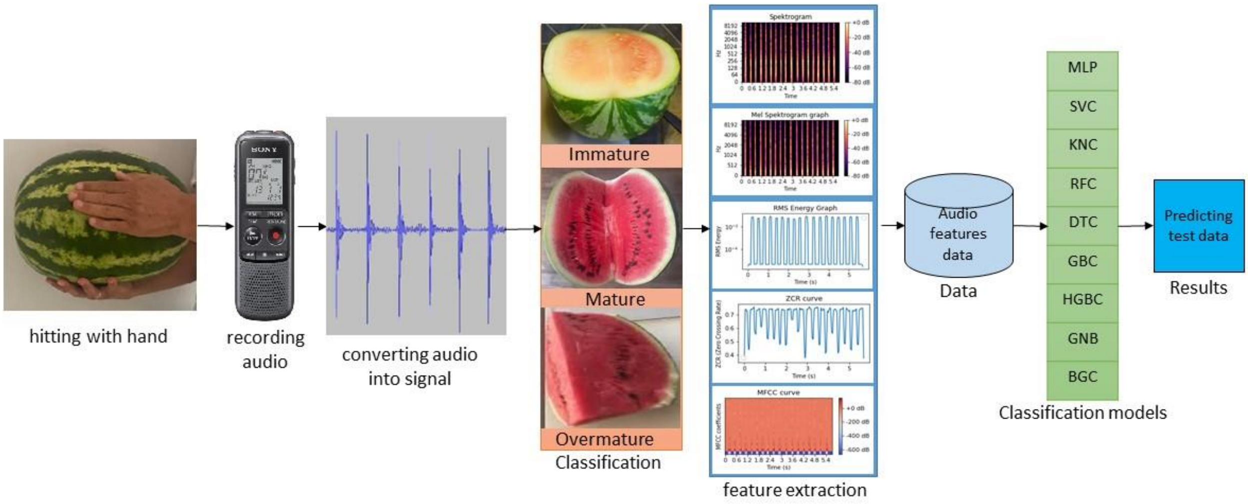

Figure 1. An overview of this study

2.1. Data Acquisition

This study was conducted on watermelons cultivated in the South-Eastern region of Turkey during the summer season, where two primary cultivation practices are common: irrigated and non-irrigated farming. Regional consumers generally prefer non-irrigated watermelons, which are perceived to offer superior taste and quality. Accordingly, for this study, non-irrigated watermelons weighing between 4 and 12 kilograms were selected as the main data source.

The samples were collected both directly from fields and from various points of sale, including local markets and roadside stands. During data acquisition, acoustic signals were recorded by tapping the outer surface of each watermelon with a seller’s hand. Before cutting, the watermelons were classified into three maturity groups: immature, mature, and overmature. The recorded signals were then used for feature extraction and classification analyses, allowing a systematic investigation of the relationship between acoustic properties and watermelon maturity.

2.2. Pre-processing Data

The acoustic signals generated by tapping the watermelons were recorded using a high-quality sound device and processed with Audacity software for detailed analysis. To ensure data integrity and clarity, background noise between taps—arising from natural environmental conditions—was carefully removed. The cleaned audio files were then systematically categorized according to watermelon maturity and organized into separate class folders, as shown in Figure 2. This structured procedure ensured accurate labeling and reliable preparation of the dataset for subsequent feature extraction, analysis, and classification.

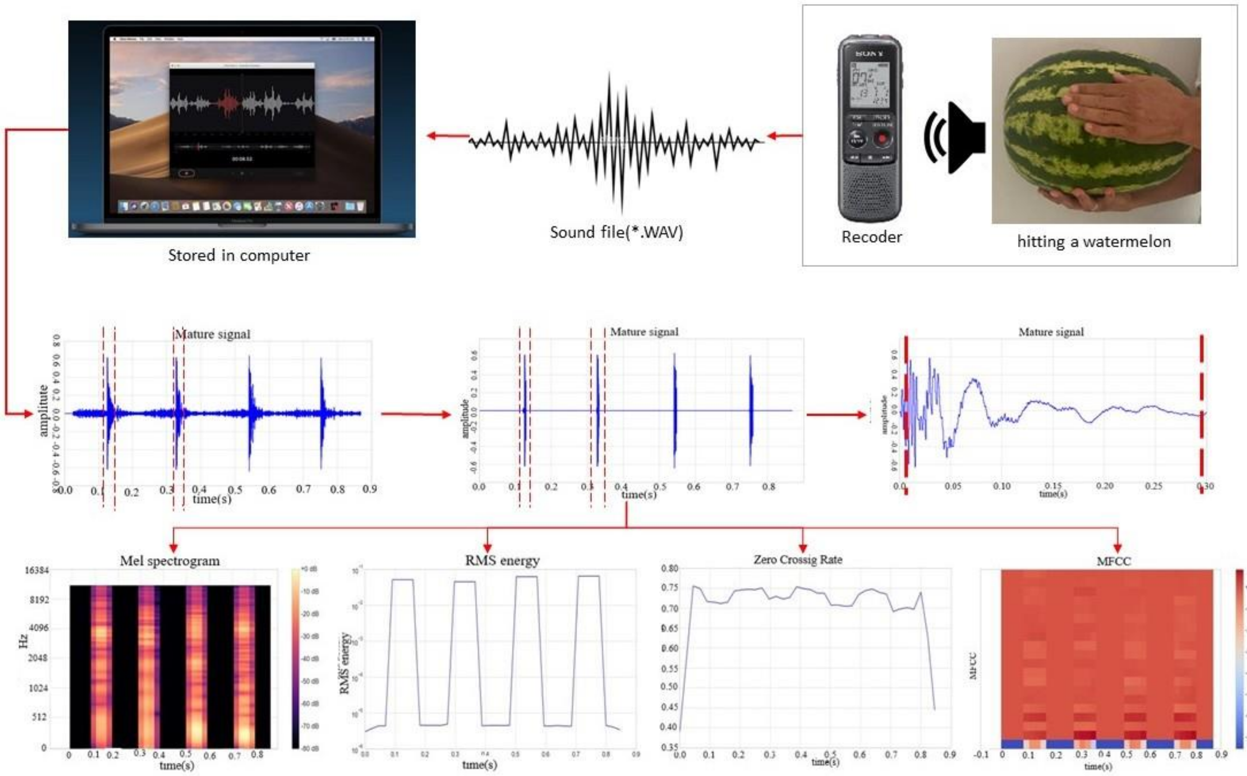

Figure 2. Processing and schematic of sound files

2.3. Sound Analysis and Feature Extraction

The Librosa library is widely used for sound and music analysis, providing essential tools for building audio-focused information systems. It supports various audio formats, including WAV and MP3, and enables spectrogram analysis as well as the extraction of relevant acoustic features. In addition to feature extraction, Librosa offers functionalities for audio manipulation, playback, and visualization. Key methods for generating instantaneous frequency spectrograms include spectral bandwidth, Short-Time Fourier Transform (STFT), and spectral flatness, with outputs presented graphically [33].

In this study, sound feature data were obtained through spectrograms, mel-spectrograms, MFCC images, Root Mean Square (RMS) energy curves, and ZCR curves. Spectrograms visually represent sound characteristics by showing signal intensity across different frequencies. The mel-spectrogram applies a perceptual scale based on human hearing (Eq. 3), while MFCC images capture frequency perception dynamics. ZCR indicates how often the signal crosses the zero amplitude line, reflecting the frequency of signal changes (Eq. 1), and RMS curves represent the overall energy distribution of the signal (Eq. 2). Using these analytical and visual techniques, a total of 120 features were extracted from each sound sample.

$$ZCR = \frac{1}{2N}\sum_{n=1}^{N} |sign(x[n]) - sign(x[n-1])| \quad (1)$$

$$RMS\ Energy = \sqrt{\frac{1}{N}\sum_{n=1}^{N} |x(n)|^2} \quad (2)$$

$$Mel(f) = 2595 \log\left(1 + \frac{f}{700}\right) \quad (3)$$

In this study, several libraries were employed to read sound files from directories and extract relevant features. The os library facilitated folder access and management, while Librosa was used for audio processing. NumPy and Pandas supported data manipulation and organization, and Matplotlib and Seaborn were used for visualization.

The scikit-learn (sklearn) library was applied to split the dataset into training and testing sets, develop classification models, and evaluate their performance. Key audio features were extracted and converted into a structured format suitable for analysis. The resulting feature dataset was then divided into training and testing subsets and applied to various classification techniques, allowing the identification of the most effective prediction models using machine learning methods.

2.4. Machine Learning Classification Models

In this study, various classification models were applied to the dataset, including SVC, Simple Linear Regression (SLR), Polynomial Regression, Decision Tree Classifier (DTC), RFC, MLP, KNC, Gradient Boosting Classifier (GBC), Gaussian Naive Bayes Classifier (GNB), and Bagging Classifier (BGC). The primary models used for classification were MLP, SVC, KNC, GBC, GNB, DTC, RFC, and BGC, as shown in Figure 3. SVC models employ linear and non-linear decision boundaries to maximize the margin between classes by identifying support vectors. Equation 4 defines linear support vectors, while Equation 5 provides the formulation for non-linear support vectors [34]. For the SVC, kernel="poly", degree=2, and gamma="auto" were used.

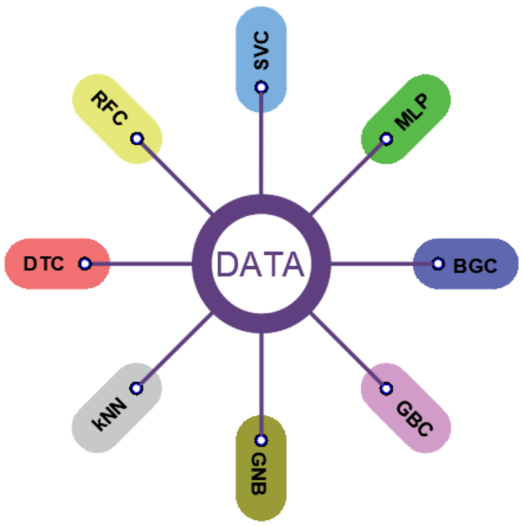

Fig. 3: Machine learning classification models

$$\sum_{i=1}^{N}(\alpha_i - \alpha_1^*).\langle x_i, x\rangle + b \qquad (4)$$

$$\sum_{i=1}^{N}(\alpha_i - \alpha_1^*).\langle \varphi(x_i), \varphi(x)\rangle + b \qquad (5)$$

The DTC algorithm partitions input data, which may include varying numbers of nodes and branches, into leaves and branches based on a function determined during training. It recursively splits the data space, integrating simple prediction models within each partition to construct a graphical decision tree. The primary goal is to generate an optimal decision tree from the dataset [35]. For the DTC, criterion="entropy" and splitter="best" were used.

The RFC creates multiple decision trees using different subsets of the input data to form an ensemble model for classification. Each tree is built from independently sampled data, and final predictions are obtained by aggregating outputs from all trees. Increasing the number of trees generally reduces generalization error, enhancing predictive accuracy, internal correlation, and error characteristics [36]. The KNC is widely used for its simplicity and robust performance. Selecting the optimal parameter k is critical, as both accuracy and training efficiency depend on it. KNC classifies data points by measuring their distance to the nearest neighbors in the training set, making decisions based on proximity within the feature space [37]. For the KNC, n_neighbors=3 was used.

ANN, inspired by the human nervous system, are applied to tasks such as dynamic system identification, pattern classification, and function approximation. In ANN, input data are processed by multiplying connection weights and applying activation functions across hidden layers. MLP models, a type of ANN, consist of an input layer, one or more hidden layers, and an output layer, and use a non-linear, supervised training algorithm to learn complex patterns [38]. For the MLP, hidden layers = (1024, 1024, 1024, 1024) and max_iter=1000 were used. The performance of the classification models was evaluated using accuracy, precision, recall, and F1 score, calculated according to Equations 6–9.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (6)$$

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \qquad (7)$$

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (8)$$

$$F1 = 2\,\frac{\mathrm{Precision*Recall}}{\mathrm{Precision+Recal}} \qquad (9)$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives.

3. RESULTS

In this study, watermelons were sourced from grocery stores and directly from fields. After being tapped by hand, the resulting sounds were recorded. Classification was determined post-cut based on internal color and taste: red and sweet watermelons were labeled mature, deteriorated or degraded watermelons as overmature, and white or pink interiors as immature, resulting in three classes. The dataset comprised 1,672 samples: 802 mature, 500 overmature, and 370 immature. Features were extracted using spectrograms, mel-spectrograms, MFCC images, RMS energy, and ZCR curves via Librosa, totaling 120 features each sample.

The dataset was split into 80% training and 20% test data. Training data were applied to classification models including MLP, KNC, SVC, RFC, DTC, GBC, HGBC, GNB, and BGC. Model performance was evaluated using precision, recall, and F1 scores. As shown in Table 2, KNC achieved the highest accuracy at 96.66%, while GNB had the lowest at 51%. KNC attained the highest precision (98.77%) for the mature class, whereas HGBC scored highest for immature (100%) and overmature (96.34%). Regarding recall, MLP scored highest for overmature (98.15%), and KNC performed best for mature (98.78%) and immature (96.15%). These results demonstrate that sound-based classification provides a reliable, non-destructive method for assessing watermelon maturity.

For F1 scores, the KNC model achieved the highest score of 97.3% for the mature class, while HGBC scored highest for overmature (95.93%) and immature (96.61%). The DTC, GNB, and BGC models showed lower F1 scores. Overall, precision, recall, and F1 analyses indicate that KNC and HGBC outperformed the others, particularly for the mature and overmature classes, whereas GNB and BGC underperformed. Classification training results are presented in Figure 4, highlighting KNC and HGBC as the most effective models for predicting watermelon maturity.

Confusion matrices generated from test data are shown in Figure 5. These matrices compare actual versus predicted classes to determine prediction success. For MLP, 159 of 167 mature samples were correctly classified (95.2%), 96 of 108 overmature samples were correctly predicted (88.9%), and 73 of 78 immature samples were correctly identified (93.6%). Misclassifications occurred among the remaining samples across the classes.

Table 2. Classification metric results

Model	Class	Accuracy (%)	Precision (%)	Recall (%)	f1 (%)
MLP	mature	92.92	97.39	89.22	93.13
	overmature		86.18	**98.15**	91.78
	immature		90.91	89.75	90.33
RFC	mature	95.46	93.72	98.2	95.91
	overmature		96.23	94.45	95.33
	immature		98.61	91.03	94.67
DTC	mature	70.26	78.95	71.86	75.24
	overmature		65.22	69.45	67.27
	immature		62.8	69.23	65.86
HGBC	mature	96.34	94.12	97.56	95.81
	overmature		**96.34**	95.5	**95.93**
	immature		**100**	93.59	**96.61**
GBC	mature	93.21	91.48	96.41	93.88
	overmature		94.34	92.6	93.46
	immature		95.78	87.18	91.28
SVC	mature	91.5	91.13	92.22	91.67
	overmature		94.4	93.52	93.96
	immature		88.32	87.18	87.75
GNB	mature	51	53.61	71.26	61.19
	overmature		55	30.56	39.29
	immature		39.44	35.9	37.59
KNC	mature	**96.66**	**98.77**	**98.78**	**97.3**
	overmature		95.42	93.7	95.86
	immature		91.47	**96.15**	96.16
BGC	mature	87.82	85.03	95.21	89.84
	overmature		90.91	83.34	86.96
	immature		91.05	78.21	84.14

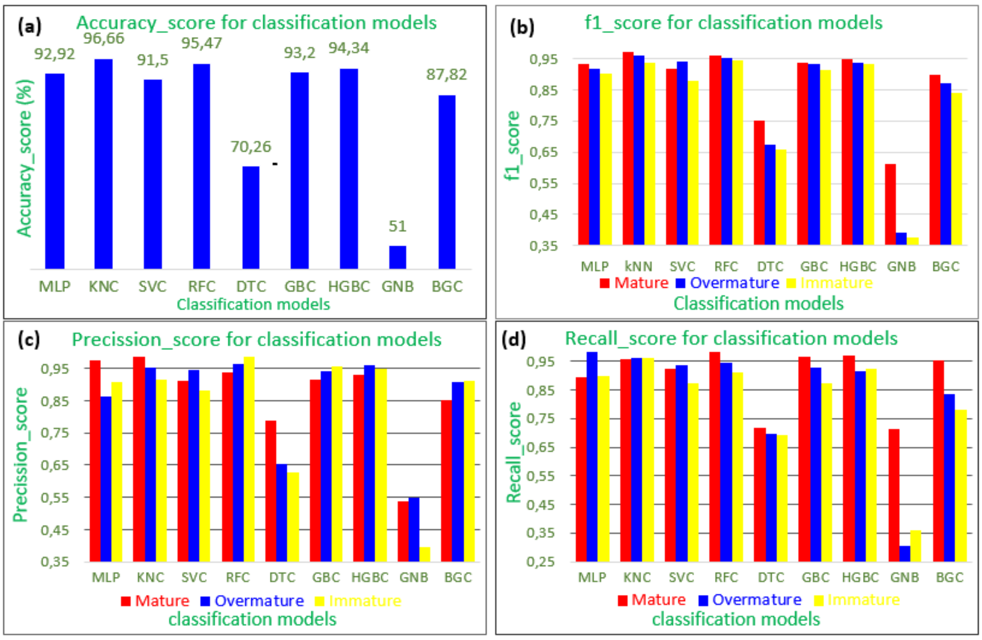

Figure 4. Results of training a) Accuracy_score b) f1_score c) Precission_score d) Recall_score

These confusion matrices illustrate model prediction accuracy across watermelon maturity classes. High success rates demonstrate that the classification models generally perform well, though some variation exists between classes. The RFC model achieved the highest accuracy for the mature class, while KNC performed best for the immature and overmature classes. In contrast, GNB exhibited the lowest accuracy across all classes. For the mature class, the top-performing model correctly predicted 164 of 167 test samples (98.2%), whereas the lowest-performing model correctly predicted only 119 samples (71.3%). For overmature, the best model achieved 104 correct predictions out of 108 (96.3%), while the lowest model predicted 33 correctly (30.6%). Similarly, for immature, the highest-performing model achieved 75 correct predictions out of 78 (96.2%), compared to 28 correct predictions (35.9%) by the lowest-performing model.

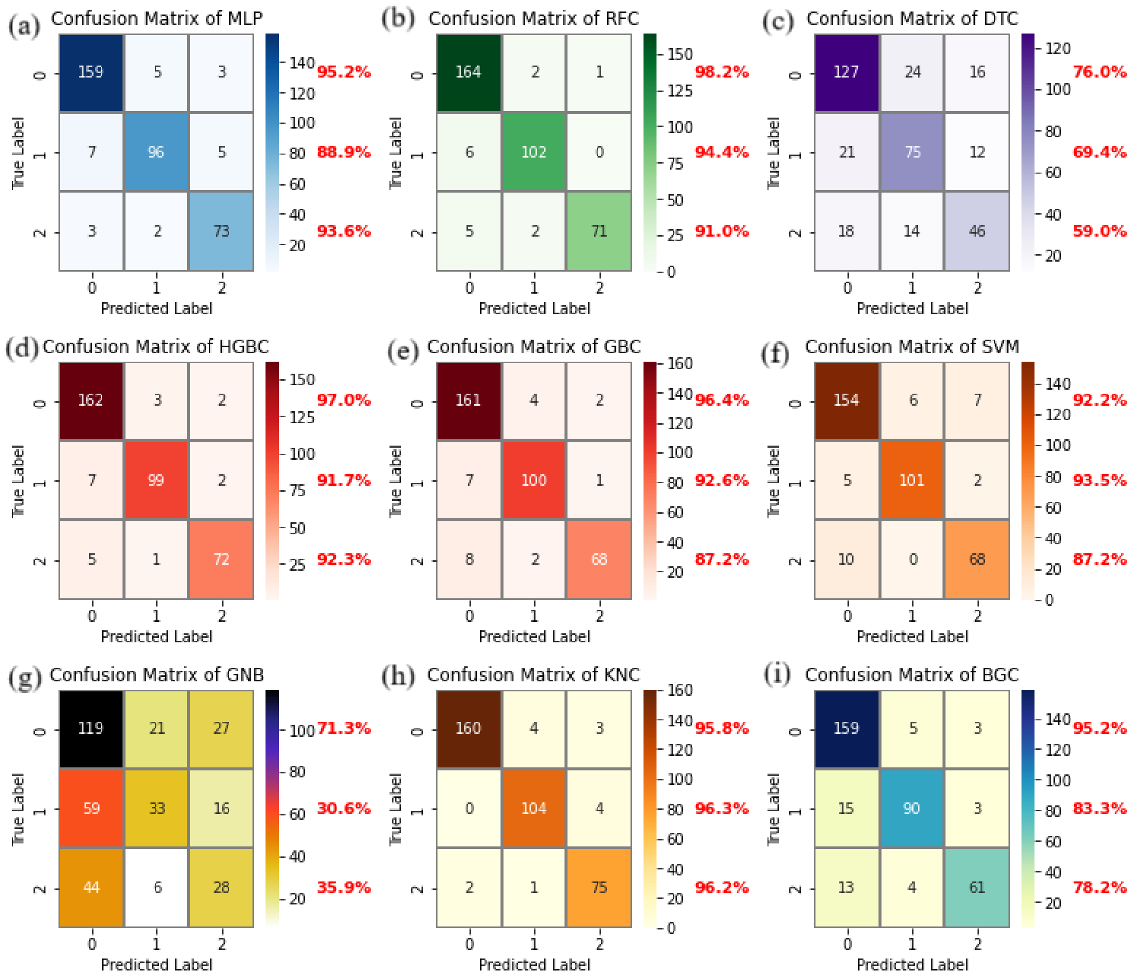

Figure 5. Predicting classification for machine learning models

Success rates, along with visualizations of true and false predictions for each class, are presented in Figure 6. DTC and GNB showed comparatively lower accuracy, whereas RFC and KNC demonstrated strong overall performance.

These findings indicate that machine learning models, particularly RFC and KNC, can effectively predict watermelon maturity based on hand-tapping sounds.

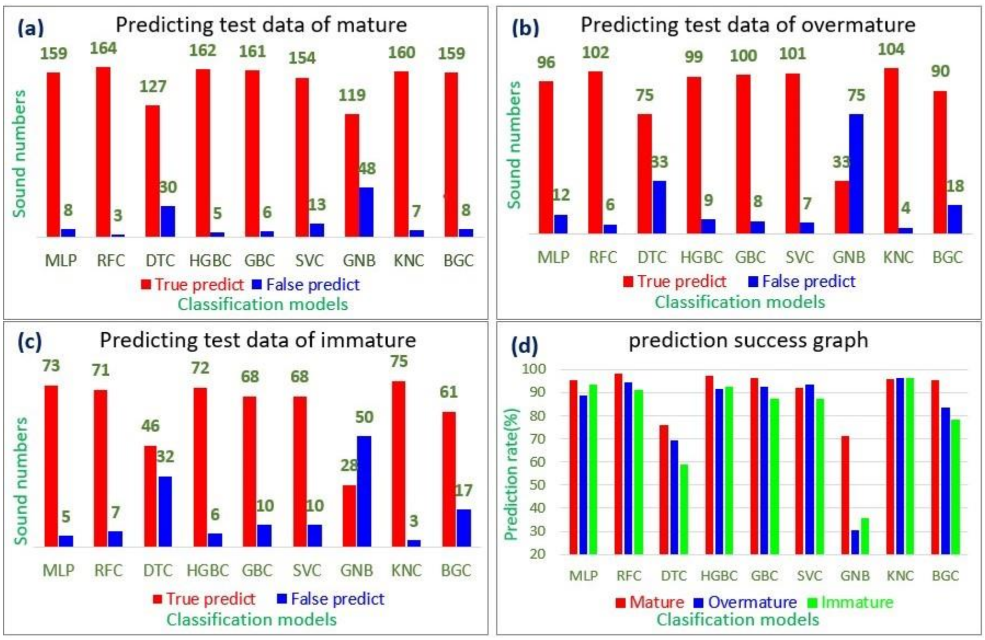

Figure 6. Predicting results for classification model

4. DISCUSSION AND CONCLUSION

This study demonstrates the potential of sound-based classification for assessing watermelon maturity, achieving a high accuracy of 96.66% using the KNC model. These results are comparable to or surpass those reported in similar studies across various domains, including animal sounds, environmental noises, and health-related audio data. The findings highlight the robustness of the proposed method, particularly under real-world conditions with ambient noise, and underscore the effectiveness of machine learning models in processing complex audio signals.

Previous studies have successfully applied sound classification, as shown in the research summarized in Table 3. For example, VGGish-MFCC features combined with KNC achieved 94.79% accuracy for honey bee sound classification [25], transformer CNN achieved 96.05% for domestic pig sounds [39], subspace KNC ensembles reached 96.23% for environmental sounds using spectrograms [40], and RFC-SVC models achieved 96.9% for fish sound classification [23]. Health-related studies, such as heartbeat and respiratory sound analysis, achieved accuracies ranging from 85% to 93% [21,41].

Table 3. Sample studies for sound classification

The studies	Sound dataset	Feature Extraction	Classification method	Accuracy (%)
[25]	Colony honey bee	VGGish-MFCC	KNC	94.79
[39]	Domestic pig sound	LM, MFCC, spectrogram	Transformer CNN	96.05
[22]	Bird sound	Mel-spectrogram, harmonic-component based spectrogram	CNN	86.31
[21]	Heartbeat sound	MFCC	ABC-ANFIS	93
[40]	Enviromental sound	Spectrogram image, CNN	Subspace KNC ensembles	96.23
[41]	Respiratory Sound	Chrome vector, MFCC, ZCR	MLP	85
[23]	Fish sound	time, frequency, and cepstral domains	RFC-SVC	96.9
[42]	Chewing sound	Bottleneck-DNN	Bi-LSTM	97.42
[17]	Watermelon Tapping sound	Image processing	GBC	92
This study	Hand hitting sound	Spectrogram, MFCC, ZCR, RMS energy	KNC	**96.66**

In comparison, the 96.66% accuracy obtained in this study using KNC is competitive with the highest performing methods in the literature. Notably, this performance was achieved despite background noise in recordings, reflecting real-world conditions in grocery stores, fields, and markets, unlike many studies that minimize ambient noise. Feature extraction employed spectrograms, mel-spectrograms, MFCC, ZCR, and RMS energy curves, which are widely recognized as effective for audio analysis. This combination allowed the extraction of 120 distinct features each sound sample, providing a robust dataset for classification and enabling reliable assessment of watermelon maturity.

Among the machine learning models tested, KNC and RFC demonstrated superior performance, achieving the highest accuracy and precision. KNC excelled in classifying immature and overmature watermelons, with success rates of 96.2% and 96.3%, respectively, while RFC achieved 98.2% for the mature class. These results align with previous studies, where KNC and RFC performed well in sound classification tasks [23, 40].

A key innovation of this study is the use of sound data recorded under natural conditions, including ambient noise. This is the first study to employ sound recordings obtained by tapping under natural conditions for watermelon maturity classification. Unlike traditional approaches that minimize background noise, this study reflects real-world scenarios where consumers and producers select watermelons based on tapping sounds. The high prediction accuracy under such conditions demonstrates the robustness of the proposed method and its potential for practical agricultural applications.

This study also builds on research linking watermelon sound characteristics to internal qualities such as sweetness and maturity. By applying machine learning, it provides an objective and scalable alternative to subjective methods like visual inspection or manual tapping, offering a reliable, non-destructive assessment of watermelon quality. However, there are limitations. The dataset was limited to a specific region and recordings were conducted under controlled but natural conditions. Future studies should include watermelons from diverse regions and growing conditions and examine the impact of varying ambient noise levels on classification accuracy. Additionally, while traditional machine learning models were employed, future work could explore deep learning architectures, such as CNN or transformer-based models, which have shown success in other sound classification tasks [33,42]. Integrating these techniques could further enhance the accuracy and robustness of watermelon maturity classification.

Overall, this study demonstrates the effectiveness of sound-based classification for assessing watermelon maturity, achieving high accuracy despite ambient noise. By combining advanced feature extraction with robust machine learning models, particularly KNC and RFC, it provides a practical and innovative solution for non-destructive quality assessment in agriculture. Future research should expand datasets and explore advanced machine learning methods to improve both accuracy and applicability.

Estimating the quality and maturity of fruits without cutting is a critical challenge for both producers and consumers. This study demonstrated the effectiveness of sound-based, non-destructive methods for assessing watermelon maturity. By recording the sounds produced when watermelons were hand-tapped and extracting features such as MFCC, ZCR, RMS energy, spectrogram, and mel spectrogram, a total of 120 features per sample were obtained. These features were used to train multiple machine learning models, with KNC achieving the highest overall accuracy (96.66%) and RFC showing the highest precision for mature watermelons (98.2%). KNC also performed best for overmature (96.3%) and immature (96.2%) classes.

The study highlights that KNC and RFC models are highly effective for predicting watermelon maturity, even in natural environments with ambient noise, reflecting real-world conditions. Compared to previous studies reporting accuracies between 85% and 99.6%, the results of this work are competitive, underscoring the robustness and practical applicability of the proposed method.
A key contribution of this study is the integration of traditional empirical practices, such as tapping and auditory evaluation, with advanced feature extraction and machine learning techniques. This approach provides a scalable, objective, and non-destructive solution for quality assessment, enabling consumers to make informed decisions and assisting producers in maintaining product standards.

Future work could expand the dataset to include watermelons from diverse regions and cultivation conditions, investigate the impact of varying ambient noise levels, and explore advanced deep learning models, such as CNN or transformer-based architectures, to further improve classification accuracy and robustness. Overall, the findings demonstrate that sound-based machine learning methods offer a practical, innovative, and reliable approach for non-destructive quality evaluation in agriculture.

REFERENCES

[1] Gondchawar N, Kawitkar RS. IoT-based smart agriculture. Int J Adv Res Comput Commun Eng. 2016; 5(6): 838-42.
[2] O'Grady MJ, Langton D, O'Hare GMP. Edge computing: A tractable model for smart agriculture? Artif Intell Agric. 2019; 3: 42-51.
[3] Femling F, Olsson A, Alonso-Fernandez F. Fruit and vegetable identification using machine learning for retail applications. In: 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). IEEE; 2018 Nov. p. 9-15.
[4] Barbole DK, Jadhav PM, Patil SB. A review on fruit detection and segmentation techniques in agricultural field. In: International Conference on Image Processing and Capsule Networks. Springer, Cham; 2021 Sep 10. p. 269-88.
[5] Kyriacou MC, Leskovar DI, Colla G, Rouphael Y. Watermelon and melon fruit quality: The genotypic and agro-environmental factors implicated. Sci Hortic. 2018; 234: 393-408.
[6] Jie D, Wei X. Review on the recent progress of non-destructive detection technology for internal quality of watermelon. Comput Electron Agric. 2018; 151: 156-64.
[7] Kyriacou MC, Soteriou GA, Rouphael Y, Siomos AS, Gerasopoulos D. Configuration of watermelon fruit quality in response to rootstock-mediated harvest maturity and postharvest storage. J Sci Food Agric. 2016; 96(7): 2400-9.
[8] Soteriou GA, Kyriacou MC, Siomos AS, Gerasopoulos D. Evolution of watermelon fruit physicochemical and phytochemical composition during ripening as affected by grafting. Food Chem. 2014; 165: 282-9.
[9] Arendse E, Fawole OA, Magwaza LS, Opara UL. Non-destructive characterization and volume estimation of pomegranate fruit external and internal morphological fractions using X-ray computed tomography. J Food Eng. 2016; 186: 42-9.
[10] Wiktor A, Gondek E, Jakubczyk E, Nowacka M, Dadan M, Fijalkowska A, Witrowa-Rajchert D. Acoustic emission as a tool to assess the changes induced by pulsed electric field in apple tissue. Innov Food Sci Emerg Technol. 2016; 37: 375-83.
[11] Zhang B, Huang W, Li J, Zhao C, Fan S, Wu J, Liu C. Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. Food Res Int. 2014; 62: 326-43.
[12] Nelson SO, Guo WC, Trabelsi S, Kays SJ. Dielectric spectroscopy of watermelons for quality sensing. Meas Sci Technol. 2007; 18(7): 1887.
[13] Wang L, Sun DW, Pu H, Zhu Z. Application of hyperspectral imaging to discriminate the variety of maize seeds. Food Anal Methods. 2016; 9: 225-34.
[14] Fan Y, Lai K, Rasco BA, Huang Y. Determination of carbaryl pesticide in Fuji apples using surface-enhanced Raman spectroscopy coupled with multivariate analysis. LWT Food Sci Technol. 2015; 60(1): 352-7.
[15] Mohd Ali M, Hashim N, Bejo SK, Shamsudin R. Quality evaluation of watermelon using laser-induced backscattering imaging during storage. Postharvest Biol Technol. 2017; 123: 51-9.
[16] Mao J, Yu Y, Rao X, Wang J. Firmness prediction and modeling by optimizing acoustic device for watermelons. J Food Eng. 2016; 168: 1-6.
[17] Chawgien K, Kiattisin S. Machine learning techniques for classifying the sweetness of watermelon using acoustic signal and image processing. Comput Electron Agric. 2021; 181: 105938.
[18] Chen M, Challita U, Saad W, Yin C, Debbah M. Artificial neural networks-based machine learning for wireless networks: A tutorial. IEEE Commun Surv Tutor. 2019; 21(4): 3039-71.
[19] Probst P, Wright MN, Boulesteix AL. Hyperparameters and tuning strategies for random forest. WIREs Data Min Knowl Discov. 2019; 9(3): e1301.
[20] Rani KV, Jawhar SJ. Lung lesion classification scheme using optimization techniques and hybrid (KNN-SVM) classifier. IETE J Res. 2022; 68(2): 1485-99.

[21] Keikhosrokiani P, Naidu AB, Iryanti Fadilah S, Manickam S, Li Z. Heartbeat sound classification using a hybrid adaptive neuro-fuzzy inference system (ANFIS) and artificial bee colony. Digit Health. 2023; 9: 20552076221150741.

[22] Xie J, Hu K, Zhu M, Yu J, Zhu Q. Investigation of different CNN-based models for improved bird sound classification. IEEE Access. 2019; 7: 175353-61.

[23] Malfante M, Mars JI, Dalla Mura M, Gervaise C. Automatic fish sounds classification. J Acoust Soc Am. 2018;143(5): 2834-46.

[24] Salau AO, Jain S. Feature extraction: A survey of the types, techniques, applications. In: 2019 International Conference on Signal Processing and Communication (ICSC). IEEE; 2019 Dec 23. p. 158-64.

[25] Di N, Sharif MZ, Hu Z, Xue R, Yu B. Applicability of VGGish embedding in bee colony monitoring: Comparison with MFCC in colony sound classification. PeerJ. 2023; 11: e14696.

[26] Sharma G, Umapathy K, Krishnan S. Trends in audio signal feature extraction methods. Appl Acoust. 2020; 158: 107020.

[27] Jie D, Zhou W, Wei X. Nondestructive detection of maturity of watermelon by spectral characteristic using NIR diffuse transmittance technique. Scientia Horticulturae, 2019; 257: 108718.

[28] Gu Q, Li T, Hu Z, Zhu Y, Shi J, Zhang L, Zhang X. Quantitative analysis of watermelon fruit skin phenotypic traits via image processing and their potential in maturity and quality detection. Computers and Electronics in Agriculture, 2025; 230:109960.

[29] Zhang YX, Han JL, Yao W. Non-destructive watermelon maturity detection by acoustic response. In 2010 2nd International Conference on Information Engineering and Computer Science, 2010; 1-4. IEEE.

[30] Garvin J, Abushakra F, Choffin Z, Shiver B, Gan Y, Kong L, Jeong N. Microwave imaging for watermelon maturity determination. Current Research in Food Science, 2023; 6: 100412.

[31] Vega-Castellote M, Sánchez MT, Torres I, De la Haba MJ, Pérez-Marín D. Assessment of watermelon maturity using portable new generation NIR spectrophotometers. Scientia Horticulturae, 2022; 304: 111328.

[32] Pamungkas WA, Bintoro N. Evaluation of watermelon ripeness using self-developed ripening detector. In IOP Conference Series: Earth and Environmental Science, 2021; 653:1, 012020. IOP Publishing.

[33] Babu PA, Nagaraju VS, Vallabhuni RR. Speech emotion recognition system with Librosa. In: 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT). IEEE; 2021 Aug 12. p. 421-4.

[34] Wu CH, Ho JM, Lee DT. Travel-time prediction with support vector regression. IEEE Trans Intell Transp Syst. 2004; 5(4): 276-81.

[35] Pekel E. Estimation of soil moisture using decision tree regression. Theor Appl Climatol. 2020; 139(3-4): 1111-9.

[36] Breiman L. Random forests. Mach Learn. 2001; 45: 5-32.

[37] Soucy P, Mineau GW. A simple KNN algorithm for text categorization. In: Proceedings of the 2001 IEEE International Conference on Data Mining. IEEE; 2001; p. 647-8.

[38] Han HG, Qiao JF. A structure optimisation algorithm for feedforward neural network construction. Neurocomputing. 2013; 99: 347-57.

[39] Liao J, Li H, Feng A, Wu X, Luo Y, Duan X, Li J. Domestic pig sound classification based on TransformerCNN. Appl Intell. 2023; 53(5): 4907-23.

[40] Kim CI, Cho Y, Jung S, Rew J, Hwang E. Animal sounds classification scheme based on multi-feature network with mixed datasets. KSII Trans Internet Inf Syst. 2020; 14(8): 3384-98.

[41] Demir F, Abdullah DA, Sengur A. A new deep CNN model for environmental sound classification. IEEE Access. 2020; 8: 66529-37.

[42] Monaco A, Amoroso N, Bellantuono L, Pantaleo E, Tangaro S, Bellotti R. Multi-time-scale features for accurate respiratory sound classification. Appl Sci. 2020; 10(23): 8606.

[43] Khan MI, Acharya B, Chaurasiya RK. iHearken: Chewing sound signal analysis-based food intake recognition system using Bi-LSTM Softmax network. Comput Methods Programs Biomed. 2022; 221: 106843.

[44] Akbal E, Tuncer T, Dogan S. Vehicle interior sound classification based on local quintet magnitude pattern and iterative neighborhood component analysis. Appl Artif Intell. 2022; 36(1): 2137653.