Dong Thap University Journal of Science, Vol. 14, No. 04S (2025): 174-188

DONG THAP UNIVERSITY JOURNAL OF SCIENCE
Tạp chí Khoa học Đại học Đồng Tháp
Natural Sciences Issue
ISSN 0866-7675 | e-ISSN 2815-567X

DOI: https://doi.org/10.52714/dthu.14.04S.2025.1577

CLASSIFICATION OF BANANA STAGES USING MICROWAVE SPECTROSCOPY BY MACHINE LEARNING

Ja-Lam Nguyen[1,2], Hai-Anh Nguyen-Thi[1,2], Dang-Khoa Trinh-Vo[1,2] , Huy-Bao Bui-Minh[1,2], Thanh-Ven Huynh[1,2], Huyen-Tran Nguyen-Thi[3], Hoang-Minh Le[1,2], and Anh-Tu Tran[1,2]*

[1]*Laboratory of Laser Technology, Faculty of Applied Science, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, 72409, Vietnam*

[2]*Vietnam National University Ho Chi Minh City, Ho Chi Minh City 71308, Vietnam*

[3]*Language and Communication, Faculty of Foreign Languages, Ho Chi Minh City University of Foreign Languages and Information Technology (HUFLIT), Ho Chi Minh City 72409, Vietnam*

**Corresponding author, Email: tranatu@hcmut.edu.vn*

Article history

Received: 30/5/2025; Received in revised form: 08/7/2025; Accepted: 09/7/2025

Abstract

In the food industry, the need to assess and predict the ripening process of fruits plays an important role in optimizing storage and transportation strategies, ensuring the product quality when reaching consumers. The study proposes a method using microwave spectrum based on vector network analyzer combined with machine learning models. It evaluates numerous machine learning models and predicts the number of days required for an unripe banana to semi ripe and then ripe banana. Data is collected through scattering parameters, including reflection parameter S11 and transmission parameter S21, in the frequency range from 1 GHz to 5 GHz. The S-parameters are processed, analyzed and extracted characteristic data and fed into the machine learning models to perform the comparison and prediction process.

Keywords: *Machine learning, ripeness prediction, vector network analyzer.*

Cite: Nguyen, J. L., Nguyen, T. H. A., Trinh, V. D. K., Bui, M. H. B., Huynh, T. V., Nguyen, T. H. T., Le, H. M., & Tran, A. T. (2025). Classification of banana stages using microwave spectroscopy by machine learning. *Dong Thap University Journal of Science, 14*(04S), 174-188. https://doi.org/10.52714/dthu.14.04S.2025.1577

PHÂN LOẠI CÁC GIAI ĐOẠN CHUỐI SỬ DỤNG PHỔ VI SÓNG BẰNG HỌC MÁY

Nguyễn Ja Lam[1,2], Nguyễn Thị Hải Anh[1,2], Trịnh Võ Đăng Khoa[1,2], Bùi Minh Huy Bảo[1,2], Huỳnh Thanh Vẹn[1,2], Nguyễn Thị Huyền Trân[3], Lê Hoàng Minh[1,2], Trần Anh Tú[1,2*]

[1]Phòng thí nghiệm Công nghệ Laser, Khoa Khoa học Ứng dụng, Trường Đại học Bách khoa Thành Phố Hồ Chí Minh (HCMUT), Thành phố Hồ Chí Minh, 72409, Việt Nam

[2]Đại học Quốc gia Thành phố Hồ Chí Minh, Thành phố Hồ Chí Minh, 71308, Việt Nam

[3]Bộ môn Ngôn ngữ và Giao tiếp, Khoa Ngoại ngữ, Trường Đại học Ngoại ngữ và Tin học Thành phố Hồ Chí Minh, Thành phố Hồ Chí Minh, Việt Nam

**Tác giả liên hệ, Email: tranatu@hcmut.edu.vn*

Lịch sử bài báo

Ngày nhận: 30/5/2025; Ngày nhận chỉnh sửa: 08/7/2025; Ngày duyệt đăng: 09/7/2025

Tóm tắt

Trong ngành thực phẩm, nhu cầu đánh giá và dự đoán quá trình chín của trái cây đóng vai trò quan trọng trong việc tối ưu hóa các chiến lược lưu trữ và vận chuyển, đảm bảo chất lượng sản phẩm tốt nhất khi đến tay người tiêu dùng. Nghiên cứu đề xuất một phương pháp sử dụng phổ vi sóng dựa trên máy phân tích mạng vectơ (VNA) kết hợp với các mô hình học máy. Phương pháp này được sử dụng để đánh giá giữa nhiều mô hình học máy và dự đoán số ngày cần thiết để một quả chuối chưa chín chuyển sang chuối chín một nửa và sau đó là chín. Dữ liệu được thu thập thông qua các tham số tán xạ, bao gồm tham số phản xạ S11 và tham số truyền S21, trong dải tần số từ 1 GHz đến 5 GHz. Các tham số S được xử lý, phân tích và trích xuất dữ liệu đặc trưng và đưa vào các mô hình học máy để thực hiện các quy trình so sánh và dự đoán.

Từ khóa: *Dự đoán độ chín, học máy, máy phân tích mạng vectơ.*

1. Introduction

In the food industry, fruit quality always plays an important role in the food supply chain. In order to ensure the sustainability of the supply-demand relationship between producers and consumers, fruits, when reaching consumers, must meet the needs of flavor, color, and nutritional value (Bahinipati, 2014). Providing unripe fruits will reduce the inherent nutrition and flavor of the fruit, overripe fruits will change the appearance and chemical composition of the fruit in a negative direction, affecting the quality of the product and the taste of consumers. This error not only affects the health of consumers but also seriously affects manufacturers' reputation and profits. Therefore, evaluating and predicting the ripeness of fruits is extremely necessary to find a method of preservation and calculate reasonable transportation time.

Traditional methods for assessing and predicting fruit ripeness rely on human senses such as observing color, touching to check firmness, etc (Bashir et al., 2020). However, these methods lack accuracy, do not ensure consistency in test results, and are not suitable for large-scale production models due to the time and effort required. Some types, such as watermelon and orange, do not have obvious changes in firmness or color to have enough information to assess fruit stages (Narendra & Hareesh, 2010). Therefore, this study proposes a scientific method that is cost-effective, non-invasive, and more accurate when using microwave signals to assess and predict fruit ripening time.

2. Theoretical overview

For foods with high water content and moisture such as fruits, dielectric properties are one of the most effective indicators of evaluation, because water has a very high dielectric constant and constitutes a large proportion of the fruit structure. Therefore, changes in quality, ripeness, or freshness will lead to significant alterations in the dielectric properties of the sample.

Research based on dielectric properties has been of interest for many years (Nelson, 2006; Venkatesh & Raghavan, 2004). Among them, microwave sensors stand out compared to other technologies due to a series of advantages such as non-invasive, non-destructive, fast measurement, and the ability to penetrate deeply into the material structure. These characteristics are particularly important among fruits, which are soft, easily damaged, and difficult to maintain shape during measurement.

Not only does it help evaluate dielectric properties effectively, but microwave technology has also been proven to be sensitive to biochemical and microbiological processes in food. A typical example is the SEQUID project (Kent et al., 2007), showing that changes such as protein degradation, alteration of water structure, and formation of polar compounds all cause significant changes in the dielectric spectrum, especially in the microwave frequency range. Due to these characteristics, the application of microwave technology in experiments evaluating ripeness, freshness, and food quality is becoming increasingly popular (Clerjon & Damez, 2007; Guo et al., 2010; Pacquit et al., 2006; Schimmer et al., 2008).

In 2023, a study examined the quality of mangosteen by using a microwave method combined with a dual-ring microstrip resonance sensor to detect the quality between normal mangosteen flesh, yellow resin-leached mangosteen, and transparent mangosteen, based on the difference in dielectric constant (Muvianto et al., 2023). This study successfully classified the quality based on certain value ranges (2.98-3.28) compatible with each type of mangosteen flesh without invasiveness and sample destruction. Another study by Van Lic used a microwave method combined with machine learning methods (ANN, KNN) to classify fruits and ripeness (Tran et al., 2023). The study had an accuracy of up to 98.75% - 99.75% in fruit classification and 98.4%-96.6% in ripeness classification. Syaiful Redzwan and his colleagues also used the microwave method combined with the Split Ring Resonator (SRR) sensor to assess the quality between fresh and aged fruits (Redzwan et al., 2018). The results showed that the use of microwaves is capable of quickly assessing the quality and ripeness of fruits, although the research is still limited to analyzing non-linear data. A study by Leekul on the classification of sweet and unsweet tangerines based on microwave signals obtained by a multi-array antenna operating in the 2.5 GHz frequency range (Leekul et al., 2016). The system showed an accuracy of up to 95% and the potential in classifying tangerines by flavor. Navid Ghavami's study showed the applicability of microwave signals to algorithms for recognizing seeds inside lemons, grapefruits, and classifying seeded and seedless fruits (Ghavami et al., 2019). Africa highlights the importance of integrating technology into the agricultural sector, especially for fruit ripeness detection and quality assessment (Africa, 2020). Traditional methods that rely on physical attributes such as shape, color, and texture are prone to human error and lack consistency. To address this, the study presents various modern approaches, including machine learning, computer vision, deep learning, image illumination, Faster-CNN and gas chromatographic systems for detecting ethylene gas. A study by Leekul presented a microwave sensor operating at 2.45 GHz using a dual patch antenna system to detect internal defects in fruits, with oranges as the test case (Leekul et al., 2016b). By analyzing the mean and standard deviation of S-parameter magnitudes at different positions, the system successfully identified granulated oranges. The sensor, designed to match sweet oranges and arranged in a perpendicular configuration, showed clear variations in reflected and coupled signals, demonstrating its potential as a low-cost solution for fruit quality classification. The study of Richard Torrealba-Melendez investigated the dielectric properties of litchi fruit using the open-ended coaxial probe method across the frequency range of 0.5-20 GHz at microwave frequencies over a 3-day storage period at room temperature (~24°C) (Torrealba-Melendez et al., 2020). Results showed that the dielectric properties increased with storage time. When measured at different temperatures (24, 30, 40 and 50°C), the dielectric constant decreased with rising temperature in the 0.5-5 GHz range but increased at higher frequencies. The loss factor exhibited a U-shaped pattern-rising at frequencies above 2 GHz but generally decreasing with temperature. A study by Garvin presents a novel microwave imaging (MWI) system designed to determine watermelon ripeness (Garvin et al., 2023). The system features a circular array of 10 Coplanar Vivaldi antennas, offering wide bandwidth, high gain, and efficient signal penetration. Automated channel switching enables rapid S-parameter collection and fast image generation. Eight watermelon samples with different ripeness levels and origins were scanned, and the resulting images were compared with physical cross-sections and sugar concentration measurements. The results showed clear differences in image characteristics based on ripeness, confirming the system's effectiveness in assessing watermelon maturity non-destructively.

This study proposed a method for predicting banana ripening stages based on microwave spectrum combined with many machine learning algorithms, aiming at a non-destructive, fast and effective solution for predicting banana ripening status.

3. Method and Experiment Design

3.1. Method

The LiteVNA-64 vector network analyzer (VNA) is a compact device used for transmitting and receiving microwave signals, which was employed to develop a practical on-body measurement system using wireless communication circuits, non-invasive sensors, and nano VNAs (Elmiladi et al., 2024). The frequency range from 1 GHz to 5 GHz was chosen in this study because it achieves an optimal balance between wave penetration depth and measurement resolution (Meng et al., 2018). In this frequency range, microwave signals can penetrate the outer skin of the fruit and interact with the water content as well as the internal structure, two key factors that reflect the ripening process. Additionally, this frequency range is suitable for the operation of VNA, which helps perform S-parameter measurements accurately. The antenna, which is made of an open-ended microstrip transmission line loaded with a complementary split-ring resonator (CSRR) (Ebrahimi et al., 2020) was directly connected to the ports of the device to perform signal transmission and reception, with the data collection model layout, shown in Fig. 1.

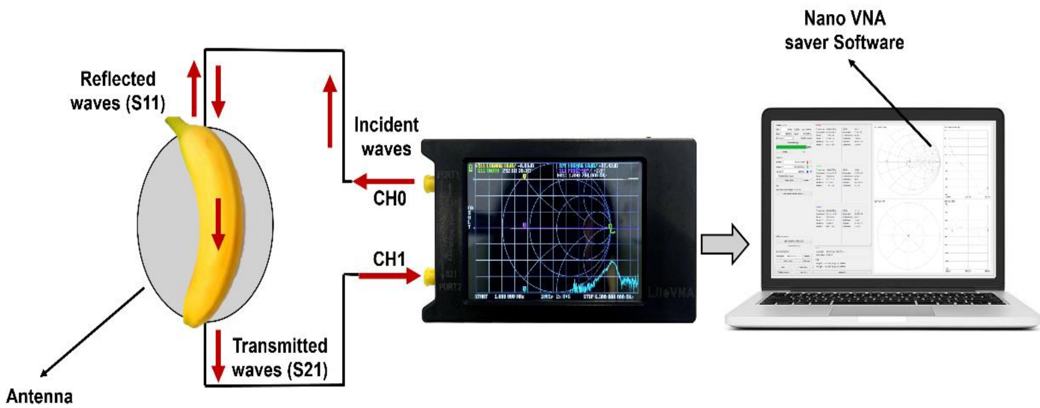

Figure 1. Data acquisition system

The banana was chosen in this study because it is a common fruit in Vietnam and has a ripening process that occurs quickly enough to suit the experiment's purpose. Furthermore, bananas contain a high amount of water, making them suitable for microwave evaluation and exhibiting clear color changes during different ripening stages.

The study was conducted through two complementary experiments, serving both objectives: (1) distinguishing among different ripening stages of banana (unripe, semi-ripe, and ripe stages), and (2) predicting the onset of ripening based on microwave signal features.

3.2. Experiment Setup

3.2.1. Daily Monitoring

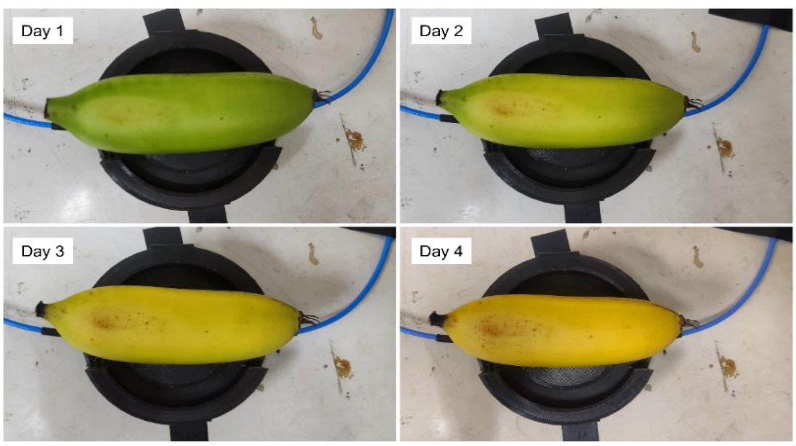

Figure 2. Visual changes in banana ripening from day 1 to day 4

In Experiment 1, a banana sample was continuously monitored from the unripe (green) state until the onset of yellowing, as shown in Fig. 2. Measurements were taken every 4 hours over 4 days. Each measurement was repeated 10 times, and the average value was taken to reduce random measurement errors. Based on the relationship between signal features (S11, S21) and the change in banana peel color over time, the "ripening onset" time point was identified.

The data collected from this experiment served as the foundation for evaluating the effect of microwaves on detecting changes in bananas during the ripening process, shown in Fig. 3.

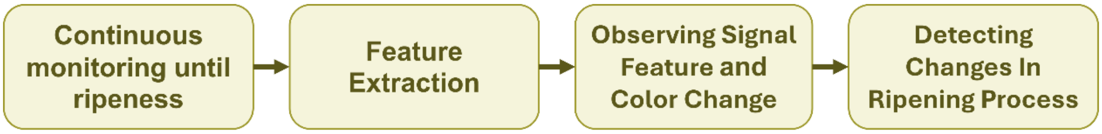

Figure 3. Block diagram procedure of Experiment 1.

To ensure measurement consistency and minimize signal loss or noise, a custom-designed circular platform was employed to secure both the antenna and the banana sample during data acquisition. This configuration enhanced signal stability and improved measurement accuracy.

3.2.2. Day Until Ripeness Classification

In Experiment 2, the S21 parameter was measured by a VNA, as illustrated in Fig. 4. The measurement process was carried out continuously once per day, starting from the point when the banana was unripe until it was fully ripe.

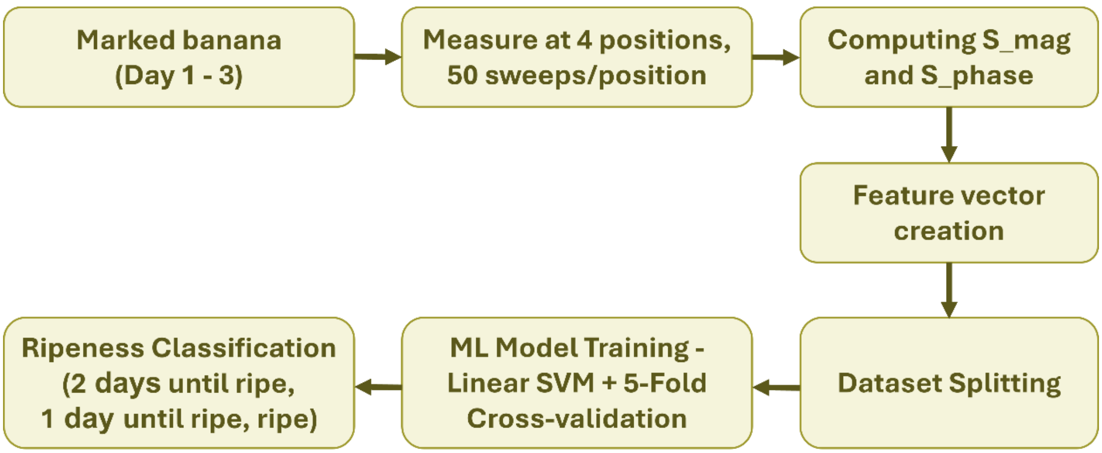

Figure 4. Block diagram procedure of Experiment 2

During a period of 3 days, a total of 30 marked bananas were randomly selected for measurement, as demonstrated in Fig. 5. For each fruit, measurements were taken at 4 different positions to ensure coverage of the entire fruit's surface.

At each measurement position, the VNA performed 50 sweeps, each collecting S (Real) and S (Image), then calculating 101 magnitude and 101 phase values of the S21 parameter using an equation (1) (Narendra & Hareesh, 2010) and equation (2) (Rauf et al., 2019), corresponding to 101 frequency points being scanned.

Figure 5. Color changes in the banana from day 1 to day 3

$$S\ (Magnitude) = 20log\left(\sqrt{S(Real) + S(Image)}\right) \tag{1}$$

$$S(Phase) = arctan\left(\frac{S(Image)}{S(Real)}\right) \tag{2}$$

A machine learning model was developed to classify bananas into three ripeness stages: 2 days until ripe, 1 day until ripe, and ripe. Each sample was characterized by two features, particularly magnitude and phase, resulting in a total of 202 features. Support Vector Machine (SVM) was chosen due to its effectiveness in high-dimensional spaces.

The training dataset comprised 16,200 samples, equally distributed among three ripeness classes, with each class containing 5,400 samples. To further ensure a rigorous and reliable evaluation of the model's performance, five-fold cross-validation was applied exclusively to the training set. Specifically, the training dataset was randomly divided into five equally sized subsets; in each iteration, one subset (20%, or 3,240 samples) served as the validation set, while the remaining four subsets (80%, or 12,960 samples) were used for model training. This process was repeated five times, ensuring each training sample was utilized exactly once for validation.

Features are normalized using the z-score normalization, transforming the data to have a mean of 0 and a standard deviation of 1 as in equation (3):

$$x'_i = \frac{x_i - \mu}{\sigma}, \tag{3}$$

where μ is the mean, σ is the standard deviation of the feature across all samples. This ensured that features with different scales contributed equally to the model. The model performance was evaluated through the following metrics (Naidu et al., 2023).

Accuracy quantifies the overall correctness of the model by calculating the proportion of correctly classified samples out of the total number of samples, as in equation (4):

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ samples}. \tag{4}$$

Weighted Precision measures the proportion of correctly predicted positive instances for each class, weighted by the class's support (number of samples in that class) to account for class imbalance as in equation (5):

$$Weighted\ Precision = \sum_{i=1}^{K} \left(\frac{Support_i}{Total\ Samples} \times \frac{TP_i}{TP_i + FP_i} \right). \tag{5}$$

Weighted Recall measures the proportion of actual positive instances correctly identified for each class, weighted by the class's support. It is computed as in equation (6):

$$Weighted\ Recall = \sum_{i=1}^{K} \left(\frac{Support_i}{Total\ Samples} \times \frac{TP_i}{TP_i + FN_i} \right). \tag{6}$$

4. Results and discussion

4.1. Results of daily measurements

Figure 6 illustrates the measured reflection characteristics of banana samples, focusing on the S11 parameter, which is a key indicator in microwave analysis. Specifically, Fig. 6(a) shows the S11 magnitude in decibels (dB) across a frequency range of 1 GHz to 5 GHz, with a zoomed-in inset highlighting the frequency band around 2.15 GHz to 2.2 GHz, where notable variations are observed. At this frequency range, the temporal change of the S11 amplitude shows a "turning point" where the initial increasing trend is reversed into a decreasing one. This phenomenon could be an important characteristic, suggesting potential use as an indicator for transitional stages in the ripening process and deserves further investigation.

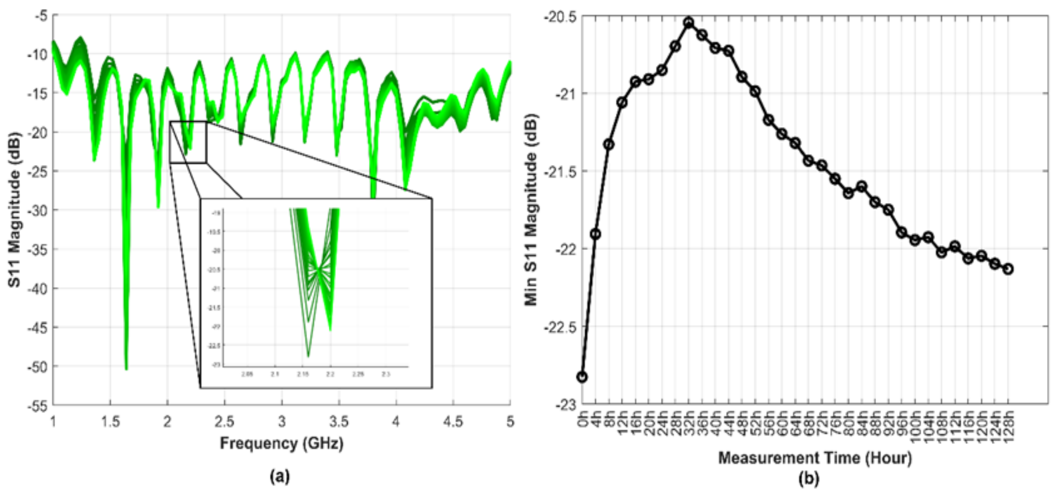

Figure 6. Measured reflection characteristics of the banana samples: (a) S11 magnitude with frequency, and (b) minimum S11 magnitude over time

Figure 6(b) tracks the minimum S11 magnitude in dB over a measurement period of 128 hours. It is evident that the minimum S11 value reaches its peak at approximately 32 hours (day 2), which corresponds to the onset of the banana's ripening process. This significant peak suggests that the microwave reflection properties, as captured by the S11 parameter, are highly sensitive to the biochemical changes occurring during ripening. This shows that peak values, especially the maximum peak of S11 can be used as indicators to identify the ripening stage, helping to mark important time points in the typical ripening process of bananas.

This demonstrates the potential of microwave technology to predict the ripeness of bananas by monitoring the S11 magnitude, offering a non-invasive and effective method for assessing fruit maturity over time.

Figure 7 shows the S21 values by frequency of banana samples measured continuously, with each line representing a corresponding measurement batch. The color of the line graph represents the measurement time: dark blue corresponds to the initial measurements and gradually shifts to light blue for the later measurements, corresponding to the increasing ripeness level. The zoomed-in section of the chart highlights the trend of the curves gradually moving from the bottom to the top (increasing amplitude) over time. This indicates the change in the dielectric properties of the sample during the ripening process. At the same time, it confirms that the S21 parameter in this experiment is highly sensitive to the natural ripening process of bananas and can be used to identify and assess ripeness.

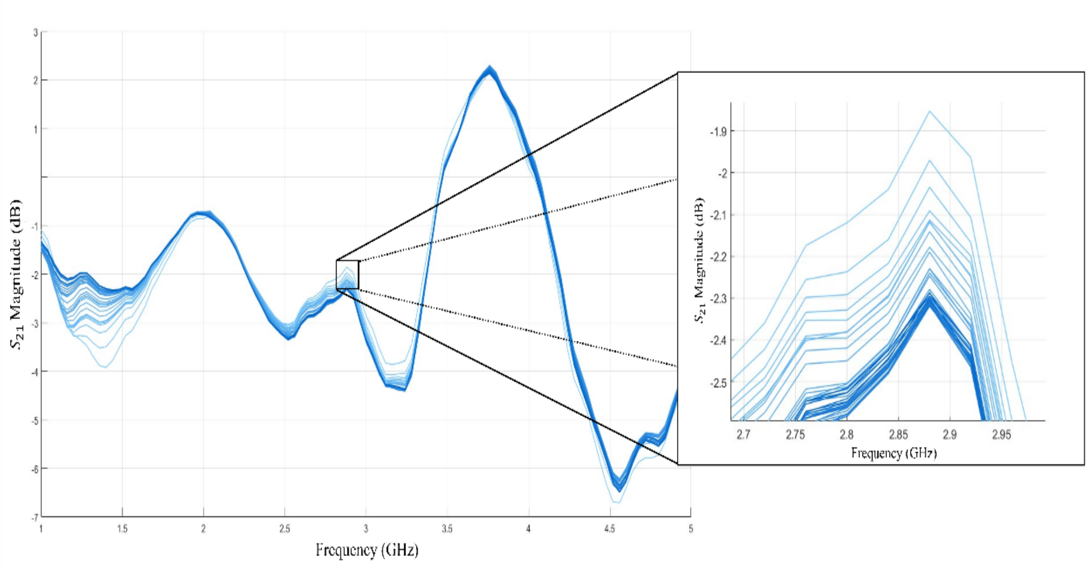

Figure 7. Measured S21 magnitude of the banana sample over frequency across ripening stages

The ripening process of bananas leads to a series of biochemical changes such as starch breakdown, water restructuring, and the formation of polar compounds. These changes affect the dielectric constant and are clearly reflected in the S11 and S21 parameters in the microwave spectrum. Particularly, the S11 signal has a peak at the transition point, indicating a change from an increasing to a decreasing trend, whereas S21 has an amplitude that tends to increase over time. These characteristics are considered potential dielectric markers for the stages of the ripening process. These changes serve as the basis for strengthening the relationship between the obtained microwave signal and the biochemical processes within the banana sample.

4.2. Results of the predictive model

Table 1 presents the performance metrics of various SVM models. The Linear SVM consistently exhibited the highest overall performance across all evaluated metrics, achieving an accuracy of 0.99844672, a precision of 0.99844912, a recall of 0.99844672, and an F1-score of 0.99844678.

Table 1. SVM model training metrics

Model	Accuracy	Precision	Recall	F1-Score
Linear SVM	0.99844672	0.99844912	0.99844672	0.99844678
Quadratic SVM	0.76812675	0.77599151	0.76812675	0.76148342
Cubic SVM	0.81317179	0.81611474	0.81317179	0.81028060
Fine Gaussian SVM	0.99167443	0.99167857	0.99167443	0.99167620
Medium Gaussian SVM	0.85641504	0.85642113	0.85641504	0.85450265
Coarse Gaussian SVM	0.64050947	0.65339561	0.64050947	0.64378005

In contrast, the Quadratic SVM demonstrated notably lower performance, with an accuracy of 0.76812675, precision of 0.77599151, recall of 0.76812675, and an F1-score of 0.76148342. The Cubic SVM model exhibited a modest improvement over the Quadratic SVM, achieving an accuracy of 0.81317179, precision of 0.81611474, recall of 0.81317179, and an F1-score of 0.81028060. The Fine Gaussian SVM model displayed competitive performance, with consistent metrics of 0.99167443 for accuracy, recall, and F1-score, and a precision of 0.99167857. Despite its robust performance, it remained slightly inferior to the Linear SVM. Meanwhile, the Medium Gaussian SVM and Coarse Gaussian SVM models demonstrated the lowest performances among the radial basis function (RBF) kernels, with accuracies of 0.85641504 and 0.64050947, respectively.

Table 2. The cross-validation metrics table.

Fold	Accuracy	Precision	Recall	F1-Score
1	0.99813607	0.99813723	0.99813607	0.99813607
2	0.99813607	0.99813732	0.99813607	0.99813611
3	0.99875738	0.99875854	0.99875738	0.99875738
4	0.99751476	0.99752112	0.99751476	0.99751494
5	0.99627213	0.99627709	0.99627213	0.99627213

Table 2 presents the cross-validation metrics obtained from the five-fold cross-validation procedure performed on the training dataset. The model consistently demonstrated exceptionally high performance across all five folds. Specifically, the accuracy ranged from a minimum of 0.99627213 (Fold 5) to a maximum of 0.99875738 (Fold 3). Similarly, precision values consistently fell within the range of 0.99627709 (Fold 5) to 0.99875854 (Fold 3), and recall values ranged from 0.99627213 (Fold 5) to 0.99875738 (Fold 3). The F1-score also remained remarkably high across all validation folds, spanning from 0.99627213 (Fold 5) to 0.99875738 (Fold 3). These consistently high metrics across all folds collectively indicate that the model possesses robust predictive performance and excellent generalization capability, thereby further confirming the reliability and stability of the chosen classification approach.

Table 3 presents the detailed classification performance of the model, evaluated on the independent test dataset, which comprises a total of 1600 samples. For the “2 days until ripe” class, the model achieved a high precision of 0.98484848 but a comparatively lower recall of 0.86666667, resulting in an F1-score of 0.92198582. This suggests that while instances predicted as “2 days until ripe” were highly accurate (few false positives), a notable portion of actual “2 days until ripe” samples were not correctly identified (more false negatives). The “1 day until ripe” class demonstrated perfect recall of 1.00000000, indicating that all samples belonging to this category were correctly identified by the model. However, this was coupled with a lower precision of 0.83857442, suggesting that some samples from other classes were incorrectly classified as “1 day until ripe.” The resulting F1-score for this class was 0.91220068. The “Ripe” class exhibited excellent performance across all metrics, with a precision of 0.99495798, recall of 0.98666667, and an F1-score of 0.99079498. These values reflect near-perfect identification accuracy and minimal misclassifications for this category.

Overall, the weighted average performance across all classes remained consistently high, with a precision of 0.95207103, recall of 0.94500000, and an F1-score of 0.94534297. This comprehensive evaluation underscores the model's robust predictive reliability and its strong capability to differentiate between the different stages of ripeness on unseen data.

Table 3. Class prediction performance

Class	Precision	Recall	F1-Score
2 days until ripe	0.98484848	0.8 6666667	0.92198582
1 day until ripe	0.83857442	1.000000000	0.91220068
Ripe	0.99495798	0.98666667	0.99079498
Weighted	0.95207103	0.94500000	0.94534297

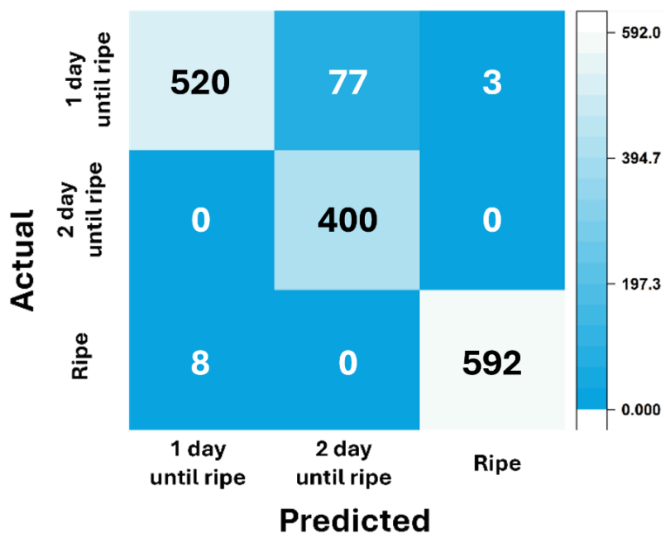

Figure 8. Confusion matrix for Linear SVM on testing data

The confusion matrix presented in Fig. 8 illustrates the performance of the Linear SVM model on the independent test dataset across three ripeness categories. The “1 day until ripe” class had 520 correctly identified samples, while 77 samples were misclassified as “2 days until ripe” and 3 as “ripe”. This misclassification is likely due to the similarity in unripe signal characteristics between the “2 days until ripe” and “1 day until ripe” stages, which can lead to confusion during classification. The “2 days until ripe” class achieved perfect identification, with all 400 samples correctly classified. For the “ripe” class, the model correctly classified 592 samples, with only 8 incorrectly identified as “1 day until ripe”.

This confusion matrix highlights the model’s overall strong performance, particularly its capability to accurately distinguish the “2 days until ripe” and “ripe” classes, while showing minor challenges in classifying the intermediate “1 day until ripe” category.

5. Conclusion

This study has demonstrated the feasibility of using S-parameters collected from a VNA to monitor and predict the ripeness of bananas. By measuring daily at various positions on the fruit’s surface, and then converting the amplitude and phase data into input feature vectors, machine learning models have been trained to effectively classify the ripening stages.

Among the evaluated models, the Linear SVM achieved the highest overall accuracy of 99.84% and exhibited strong capability in differentiating between ripeness stages. Notably, it performed particularly well in distinguishing the “1 day until ripe” and “ripe” categories. The validation results further confirm the model’s effectiveness, showing that it accurately identified nearly all samples with a high degree of reliability and consistency across the classes.

This result demonstrates the potential application of microwaves in identifying fruit ripeness, while indicating the limitations when the signal characteristics overlap between stages. Improvements in data preprocessing and model design could be the next research directions to increase classification accuracy.

This study opens a new direction in the application of microwave spectroscopy in food evaluation, particularly in real-time prediction of fruit ripeness and quality. Instead of merely distinguishing between fresh–spoiled or ripe–unripe states, the use of S-parameters combined with machine learning enables more continuous and accurate monitoring of biochemical changes.

The application of microwave technology in the food industry can be integrated into automated sorting lines, enabling rapid, non-destructive assessment of fruit ripeness. This allows businesses to optimize storage and transportation, minimize post-harvest losses, and ensure product quality reaches consumers effectively.

Acknowledgements: We would like to thank Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, for supporting this study.

References

Aaron Don M. Africa, A. R. V. T. M. A. A. T. (2020). Ripe fruit detection and classification using machine learning. *International Journal of Emerging Trends in Engineering Research*, *8*(5), 1845–1849. https://doi.org/10.30534/ijeter/2020/60852020

Bahinipati, B. K. (2014). The procurement perspectives of fruits and vegetables supply chain planning. *International Journal of Supply Chain Management*, *3*(2), 111–131.

Bashir, S., Jabeen, A., Makroo, H. A., & Mehraj, F. (2020). Application of computer vision system in fruit quality monitoring. In *Sensor-Based Quality Assessment Systems for Fruits and Vegetables* (pp. 267–290). Apple Academic Press. https://doi.org/10.1201/9781003084174-11

Clerjon, S., & Damez, J. L. (2007). Microwave sensing for meat and fish structure evaluation. *Measurement Science and Technology*, *18*(4), 1038–1045. https://doi.org/10.1088/0957-0233/18/4/011

Ebrahimi, A., Scott, J., & Ghorbani, K. (2020). Microwave reflective biosensor for glucose level detection in aqueous solutions. *Sensors and Actuators, A: Physical*, *301*, 111662. https://doi.org/10.1016/j.sna.2019.111662

Elmiladi, L. K., Hora, K. Y., Aaen, P. H., & Elsherbeni, A. Z. (2024). Wireless monitoring of S-parameters measurement using a Nano-VNA for biomedical applications. *2024 International Applied Computational Electromagnetics Society Symposium, ACES 2024*, 1–2.

Garvin, J., Abushakra, F., Choffin, Z., Shiver, B., Gan, Y., Kong, L., & Jeong, N. (2023). Microwave imaging for watermelon maturity determination: Fruit maturity determination. *Current Research in Food Science*, *6*. https://doi.org/10.1016/j.crfs.2022.100412

Ghavami, N., Sotiriou, I., & Kosmas, P. (2019). Experimental investigation of microwave imaging as means to assess fruit quality. *13th European Conference on Antennas and Propagation, EuCAP 2019*, 1–5.

Guo, W., Zhu, X., Liu, H., Yue, R., & Wang, S. (2010). Effects of milk concentration and freshness on microwave dielectric properties. *Journal of Food Engineering*, *99*(3), 344–350. https://doi.org/10.1016/j.jfoodeng.2010.03.015

Kent, M., Knöchel, R., Daschner, F., Schimmer, O., Oehlenschläger, J., Mierke-Klemeyer, S., Kroeger, M., Barr, U. K., Floberg, P., Tejada, M., Huidobro, A., Nunes, L., Martins, A., Batista, I., & Cardoso, C. (2007). Intangible but not intractable: The prediction of fish "quality" variables using dielectric spectroscopy. *Measurement Science and Technology*, *18*(4), 1029–1037. https://doi.org/10.1088/0957-0233/18/4/010

Leekul, P., Chivapreecha, S., & Krairiksh, M. (2016a). Microwave sensor for tangerine classification based on coupled-patch antennas. *International Journal of Electronics*, *103*(8), 1287–1300. https://doi.org/10.1080/00207217.2015.1092602

Leekul, P., Chivapreecha, S., & Krairiksh, M. (2016b, March 7). Microwave sensor for defected fruit classification. *Proceeding - 2015 IEEE International Conference on Antenna Measurements and Applications, IEEE CAMA 2015*. https://doi.org/10.1109/CAMA.2015.7428144

Meng, Z., Wu, Z., & Gray, J. (2018). Microwave sensor technologies for food evaluation and analysis: Methods, challenges and solutions. *Transactions of the Institute of Measurement and Control*, *40*(12), 3433–3448. https://doi.org/10.1177/0142331217721968

Muvianto, C. M. O., Yuniarto, K., Ariessaputra, S., Al Sasongko, S. M., Darmawan, B., & Syafaruddin, C. (2023). Microwave non-destructive technique using a double-ring resonator for classification of transparent flesh and yellow gum mangosteens. *E3S Web of Conferences*, *465*, 2063. https://doi.org/10.1051/e3sconf/202346502063

Naidu, G., Zuva, T., & Sibanda, E. M. (2023). A review of evaluation metrics in machine learning algorithms. *Lecture Notes in Networks and Systems*, *724 LNNS*, 15–25. https://doi.org/10.1007/978-3-031-35314-7_2

Narendra, V. G., & Hareesh, K. S. (2010). Prospects of computer vision automated grading and sorting systems in agricultural and food products for quality evaluation. *International Journal of Computer Applications*, *1*(4), 1–12. https://doi.org/10.5120/111-226

Nelson, S. O. (2006). Agricultural applications of dielectric measurements. *IEEE Transactions on Dielectrics and Electrical Insulation*, *13*(4), 688–702. https://doi.org/10.1109/TDEI.2006.1667726

Pacquit, A., Lau, K. T., McLaughlin, H., Frisby, J., Quilty, B., & Diamond, D. (2006). Development of a volatile amine sensor for the monitoring of fish spoilage. *Talanta*, *69*(2 SPEC. ISS.), 515–520. https://doi.org/10.1016/j.talanta.2005.10.046

Rauf, H. T., Saleem, B. A., Lali, M. I. U., Khan, M. A., Sharif, M., & Bukhari, S. A. C. (2019). A citrus fruits and leaves dataset for detection and classification of citrus diseases through machine learning. *Data in Brief*, *26*, 104340. https://doi.org/10.1016/j.dib.2019.104340

Redzwan, S., Perez, M. D., Velander, J., & Augustine, R. (2018). Study of maturity fruit assessment using permittivity and microwave reflectivity measurements for quality classification. *2018 IEEE Conference on Antenna Measurements and Applications, CAMA 2018*, 1–3. https://doi.org/10.1109/CAMA.2018.8530481

Schimmer, O., Daschner, F., & Knochel, R. (2008). UWB-sensors in food quality management - the way from the concept to market. *Proceeedings of The 2008 IEEE International Conference on Ultra-Wideband, ICUWB 2008*, *2*, 141–144. https://doi.org/10.1109/ICUWB.2008.4653371

Torrealba-Melendez, R., Tamariz-Flores, E. I., Sosa-Morales, M. E., Colín-Beltran, E., Miranda-Díaz, J. E., & Hernández-Ruíz, L. (2020). Dielectric properties of litchi fruit (Litchi chinensis Sonn) at microwave frequencies. *Journal of Food Science and Technology*, *57*(7), 2758–2763. https://doi.org/10.1007/s13197-020-04490-7

Tran, V. L., Doan, T. N. C., Ferrero, F., Huy, T. Le, & Le-Thanh, N. (2023). The novel combination of nano vector network analyzer and machine learning for fruit identification and ripeness grading. *Sensors*, *23*(2), 952. https://doi.org/10.3390/s23020952

Venkatesh, M. S., & Raghavan, G. S. V. (2004). An overview of microwave processing and dielectric properties of agri-food materials. *Biosystems Engineering*, *88*(1), 1–18. https://doi.org/10.1016/j.biosystemseng.2004.01.007