iJRASET

International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83414

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Design of a Spectral Signature Database for Fruit Adulteration Detection Using Machine Learning

Snehal S. Datwase[1], Dr. Charansing N. Kayte[2], Dr. R.R. Deshmukh[3]
[1, 3]*Department of computer Science and It Dr.Babasaheb Ambedkar Marathwada university,Chh.Sambhajinagar,India*
[2]*Dept. Of digital and cyber forensicDr.Babasaheb Ambedkar Marathwada UniversityChh.Samnhajinagar,India*

Abstract-The detrimental consequences of artificial fruit ripening on human health and food quality have made food adulteration a serious concern. This study suggests a hyperspectral spectral database framework that uses machine learning and spectral signature analysis to identify fruits that have been chemically and naturally ripened. The ASD Field Spec® 4 spectroradiometer was used to gather the spectral signature of banana, mango, and apple samples in the 350–2500 nm wavelength range. Under both natural and chemically ripened circumstances, 1430 spectral signatures were obtained from several fruit varietals. For effective storage, preprocessing, retrieval, and spectral library creation, the gathered hyperspectral data were processed using RS3 software and arranged into a structured Fruit Database. To increase feature extraction and spectrum quality, preprocessing methods including normalisation and second derivative analysis were used. Reflectance spectra in the Visible (VIS), Near Infrared (NIR), and Short-Wave Infrared (SWIR) regions could be visualised and compared thanks to the created spectral library. The suggested framework offers a dependable, non-destructive, and scalable method for intelligent fruit adulteration detection and food quality monitoring applications. It also facilitates the effective management of high-dimensional hyperspectral data.

Keywords: ASD FieldSpec® 4, Spectral Library, Fruit, Hyperspectral, Reflectance.

I. INTRODUCTION

Food adulteration has become a major concern in the agricultural and food industries due to its direct impact on human health and food quality [1], [2]. In order to meet market demand and speed up the ripening process, artificial fruit ripening using dangerous chemicals like calcium carbide has expanded dramatically in recent years [3]. Acetylene gas, which is released by calcium carbide and functions similarly to natural ethylene, may contain hazardous contaminants like phosphorus and arsenic, which can lead to major health problems like headaches, vertigo, neurological disorders, respiratory problems, and even cancer after prolonged exposure [3], [4]. As a result, distinguishing between chemically and naturally ripened fruits has grown to be a significant difficulty in food safety and quality assurance systems [5].

Artificial fruit ripening using hazardous chemicals like calcium carbide has grown significantly in recent years to meet market demand and expedite the ripening process [3]. Calcium carbide releases acetylene gas, which functions similarly to natural ethylene but may contain dangerous contaminants like phosphorus and arsenic. Prolonged exposure to acetylene gas can cause serious health issues like headaches, vertigo, neurological disorders, respiratory issues, and even cancer [3], [4]. Because of this, differentiating between fruits that have been chemically and naturally ripened has become a major challenge in food safety and quality assurance systems [5].

Due to their ability to record comprehensive reflectance information across a broad range of wavelengths, hyperspectral and spectral signature analysis have become viable methods for food adulteration detection [2], [8]. Fruits' distinct chemical and physical properties are contained in their spectral fingerprints, which make it possible to identify ripening patterns and adulteration effects that are invisible to the human eye [8]. For fruit quality evaluation, the combination of hyperspectral analysis and machine learning algorithms offers high accuracy, quick analysis, and non-invasive detection capabilities [1], [9].

In order to store, arrange, and retrieve vast amounts of hyperspectral data produced during spectral collecting, spectrum database creation is essential [10]. Effective administration of wavelength data, preprocessing outcomes, feature extraction data, and classification labels is made possible by a structured spectral database [10], [11]. For machine learning applications, these databases facilitate scalability, reproducibility, and quick access to spectrum data [11]. Because hyperspectral datasets contain high-dimensional information gathered across hundreds or thousands of wavelength bands, efficient data storage and retrieval procedures are particularly crucial [2], [10].

To overcome these constraints, this study suggests a structured spectral signature database system that uses machine learning and hyperspectral analysis to identify fruit adulteration. The study's main objectives are to gather spectral data from both chemically and naturally ripened fruits, preprocess the data, store it in a structured database, and utilise machine learning techniques to accurately classify the data. By combining hyperspectral data analysis, database management, and machine learning into a single framework for intelligent fruit adulteration detection, the suggested method seeks to support safer food monitoring systems.

II. LITERATURE REVIEW

This paper proposes a structured spectral signature database system that employs hyperspectral analysis and machine learning to detect fruit adulteration in order to get around these limitations. The primary goals of the study are to collect spectral data from fruits that have been chemically and organically ripened, preprocess the data, store it in a structured database, and use machine learning algorithms to appropriately identify the data.

The proposed approach aims to enable safer food monitoring systems by integrating hyperspectral data analysis, database administration, and machine learning into a unified framework for intelligent fruit adulteration detection.

Spectral databases are crucial for effectively storing, organising, and retrieving hyperspectral data, according to a number of studies [15]. Database management systems are crucial for managing preprocessing outputs, feature extraction results, and classification labels since hyperspectral datasets have hundreds of wavelength bands and substantial amounts of reflectance data [15], [16]. By keeping standardised datasets, structured spectrum databases help enhance repeatability and make training machine learning models easier [16].

Using spectral information, machine learning approaches have been widely used to categorise fruits that have been chemically and naturally matured [17]. Due to its ability to handle high-dimensional hyperspectral data and small sample sizes, Support Vector Machine (SVM) is one of the most often utilised algorithms [17], [18]. Due to their resilience and feature selection capabilities, Random Forest (RF) and Partial Least Squares Discriminant Analysis (PLS-DA) have also demonstrated encouraging results in spectral classification tasks [18], [19].

Convolutional Neural Networks (CNNs) and other deep learning techniques have recently been presented for the automatic feature extraction and categorisation of hyperspectral fruit datasets [20]. Although CNN-based models often demand big datasets and considerable processing resources, they can capture complicated spectral-spatial patterns and achieving high classification accuracy [20], [21]. Combining dimensionality reduction methods like Principal Component Analysis (PCA), t-SNE, and UMAP with machine learning algorithms has been shown to boost performance in numerous studies [18], [21].

Hyperspectral analysis has shown great promise in detecting dangerous chemical treatments like ethephon and calcium carbide in food adulteration [12], [22]. Preprocessing techniques like normalisation, smoothing, and derivative spectroscopy can be employed to analyse spectral reflectance fluctuations brought on by artificial ripening agents [22]. Second derivative analysis is a popular preprocessing method because it improves class separability, eliminates baseline drift, and amplifies tiny spectral peaks [23].

When it comes to identifying hazardous chemical treatments like ethephon and calcium carbide in food adulteration, hyperspectral analysis has demonstrated a lot of potential [12], [22]. Preprocessing techniques like normalisation, smoothing, and derivative spectroscopy can be employed to analyse spectral reflectance fluctuations brought on by artificial ripening agents [22]. Because it enhances class separability, removes baseline drift, and amplifies minute spectral peaks, second derivative analysis is a widely used preprocessing technique [23].

For precise and scalable fruit adulteration detection systems, an integrated framework including hyperspectral imaging, structured spectral databases, preprocessing methods, dimensionality reduction, and machine learning algorithms is thus desperately needed [16], [24].

TABLE 1: LITERATURE REVIEW TABLE

Author	Method	Dataset	Accuracy	Limitation
Huang et al. [12]	SVM with HSI	Mango spectral dataset	92%	Small dataset size
Mishra et al. [13]	CNN	Banana image dataset	90%	No spectral analysis
Qin et al. [14]	PLS-DA	Apple hyperspectral data	94%	Limited preprocessing
Lu et al. [15]	Random Forest	Multi-fruit spectral dataset	95%	Lack of database framework
Gowen et al. [16]	ANN	Hyperspectral fruit images	91%	High computational complexity

Author	Method	Dataset	Accuracy	Limitation
Wu et al. [17]	SVM + PCA	Chemically ripened banana spectra	96%	Limited real-world validation
ElMasry et al. [18]	PCA + HSI	Fruit quality spectral data	93%	Storage and retrieval issues
Sun et al. [19]	CNN + HSI	Mixed fruit dataset	97%	Requires large training data
He et al. [20]	Deep Learning	Hyperspectral mango dataset	95%	High hardware requirement
Geladi et al. [21]	Spectral Analysis	Agricultural spectral data	89%	Low scalability
Chen et al. [22]	RF + Spectroscopy	Fruit adulteration dataset	94%	Limited wavelength optimization

III. METHODOLOGY

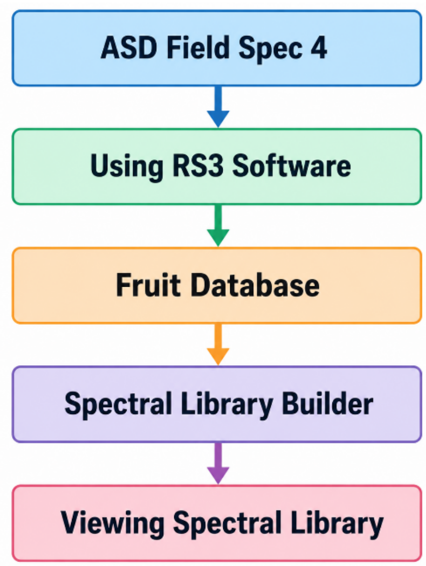

Fig 1: Proposed Framework for Building Spectral Library

The suggested framework for creating hyperspectral spectral libraries is depicted in this figure. The ASD FieldSpec® 4 spectroradiometer was used to obtain spectral signatures, which were then processed using RS3 software. The Spectral Library Builder was used to organise and store the processed data in the Fruit Database. Lastly, spectral comparison and adulteration detection were performed by visualising and analysing the spectral libraries.

IV. DATABASE AND DATASET DESIGN

A. Data Collection

The dataset utilised in this study was created to perform spectral signature analysis to assess and categorise apples that were chemically and naturally ripened. Three popular fruits—bananas, mangos, and apples—were used to gather the spectral data. These fruits were chosen because to their strong susceptibility to artificial ripening with chemicals like calcium carbide. To guarantee appropriate class representation and enhance machine learning model performance, both naturally ripened and chemically ripened samples were used.

Fruit samples were gathered in controlled environments from nearby agricultural markets and storage facilities. To replicate the adulteration circumstances frequently seen in commercial fruit marketplaces, chemically ripened samples were created using calcium carbide treatment. To obtain accurate and consistent spectral reflectance data, each fruit sample was scanned several times. A hyperspectral spectrometer that covers the visible, near-infrared (VNIR), and short-wave infrared (SWIR) regions was used to obtain the spectral signatures. Fruits' interior biochemical and physical characteristics can be thoroughly examined thanks to the spectrometer's ability to record reflectance data over a wide wavelength range.

B. Spectrometer Details

ASD Field Spec® 4 Spectroradiometer Specifications

TABLE 2: ASD FIELDSPEC® 4 SPECTRORADIOMETER SPECIFICATIONS.

Parameter	Specification
Spectral Range	350–2500 nm
Spectral Resolution	3 nm @ 700 nm 8 nm @ 1400/2100 nm
Spectral Sampling Bandwidth	1.4 nm @ 350–1000 nm 1.1 nm @ 1001–2500 nm
Scanning Time	100 milliseconds
Stray Light Specification	VNIR 0.02%, SWIR 1 & 2 0.01%
Wavelength Reproducibility	0.1 nm
Wavelength Accuracy	0.5 nm
Maximum Radiance	VNIR 2X Solar, SWIR 10X Solar
Channels	2151
Detectors	VNIR detector (350–1000 nm): 512 element silicon array SWIR 1 detector (1001–1800 nm): Graded Index InGaAs Photodiode, Two Stage TE Cooled SWIR 2 detector (1801–2500 nm): Graded Index InGaAs Photodiode, Two Stage TE Cooled
Input	1.5 m fiber optic (25° field of view). Optional narrower field of view fiber optics available.
Noise Equivalent Radiance (NeDL)	VNIR 1.0×10^{-9} W/cm²/nm/sr @ 700 nm SWIR 1 1.4×10^{-9} W/cm²/nm/sr @ 1400 nm SWIR 2 2.2×10^{-9} W/cm²/nm/sr @ 2100 nm
Weight	5.44 kg (12 lbs)
Calibrations	Wavelength, absolute reflectance, radiance, irradiance. All calibrations are NIST traceable (radiometric calibrations are optional).
Computer	Windows® 7 64-bit laptop (instrument controller)
Warranty	One-year full warranty including expert customer support

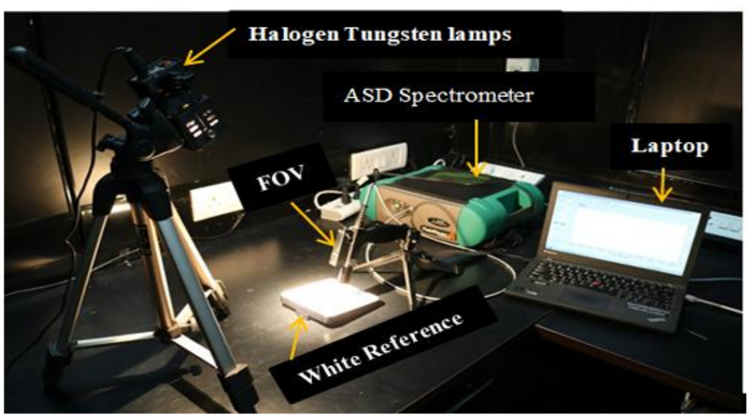

Fig 2: Laboratory Setup of the ASD FieldSpec 4 Spectroradiometer

Using machine learning and hyperspectral analysis techniques, the suggested hyperspectral spectral signature database was created for the identification and categorisation of chemically and naturally ripened fruits.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

High-dimensional spectral data gathered from various fruit samples can be stored, arranged, pre-processed, and retrieved using the database's structured platform. An effective database structure is crucial for data management, analysis, and model training because hyperspectral datasets comprise thousands of reflectance values over several wavelength bands.

Spectral signatures from three main fruit categories—bananas, mangos, and apples—are included in the database. Samples that were chemically and naturally ripened were included to facilitate comparative spectrum analysis and supervised machine learning categorisation. The dataset was created to capture spectral heterogeneity brought on by sample region, ripening condition, and fruit variety. The database has 1430 spectral signatures in total. 360 spectral signatures were obtained from the peel and flesh of banana samples. It is easier to capture the internal and exterior biochemical changes brought on by artificial ripening when peel and flesh spectra are included. To strengthen the dataset's spectral variability resilience, local banana varieties were considered.

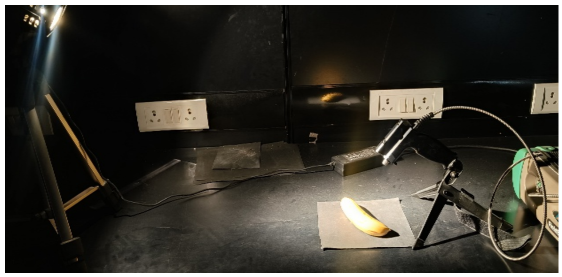

Fig 3: Data collection of banana

660 spectral signatures make up the mango dataset, which makes up the majority of the database. Several mango cultivars, including Dasheri, Totapari, Kesar, Mallika, Badam, and Raw mango, had their whole fruit spectral scans collected. Including a variety of mangoes broadens the dataset and enhances machine learning models' capacity for generalisation.

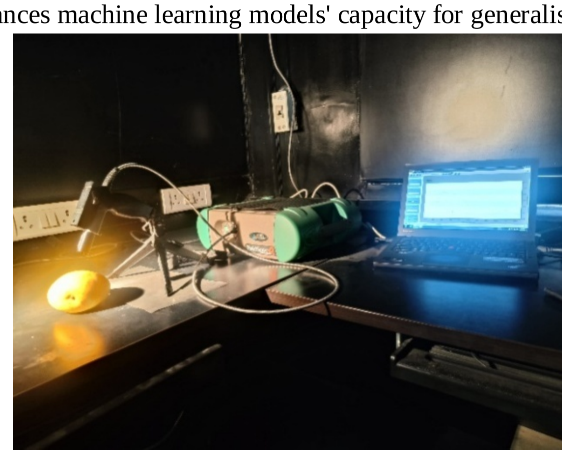

Fig 4: Data collection of Mango

410 spectral characteristics from entire fruit samples of the Royal Gala, Irani, and Kashmiri apple cultivars are included in the apple dataset. In order to preserve balanced class representation and enhance classification reliability, both naturally ripened and chemically ripened apples were included to the database.

Fig 5: Data collection of apples

TABLE 3: DATABASE TABLE

Fruit	Total spectral signature	Type of Samples	Part Used	Varieties Included
Banana	360	Natural & Chemically Ripened	Peel & Flesh	Local banana varieties
Mango	660	Natural & Chemically Ripened	Whole fruit	Dasheri, Totapari, Kesar, Mallika, Badam, Raw mango
Apple	410	Natural & Chemically Ripened	Whole fruit	Kashmiri Apple, Irani Apple, Royal Gala
Total	1430			

The ASD FieldSpec® 4 Spectroradioeter, which operates in the wavelength range of 350–2500 nm, was used to get the spectral signatures. Reflectance information from the Visible (VIS), Near Infrared (NIR), and Short-Wave Infrared (SWIR) regions is included in the collected hyperspectral data, allowing for a thorough examination of pigment concentration, moisture content, sugar composition, and internal chemical characteristics linked to fruit ripening and adulteration.
The database architecture was created to systematically store extracted features, dimensionality reduction results, preprocessing outputs, raw spectral data, and classification labels. Effective spectrum data retrieval, preprocessing, feature engineering, visualisation, and machine learning model construction are all supported by this structured database platform. Additionally, the database enhances hyperspectral datasets scalability, reproducibility, and standardisation for next studies on food adulteration detection and practical intelligent food quality monitoring systems.

V. CONCLUSION

This study suggested a hyperspectral spectral database system that uses machine learning and spectral signature analysis to identify both chemically and naturally ripened crops. The ASD Field Spec® 4 spectroradiometer was used to gather spectral data of banana, mango, and apple samples in the 350–2500 nm wavelength range. Hyperspectral data could be efficiently stored, pre-processed, visualised, and analysed thanks to the developed Fruit Database and spectral library system. For intelligent fruit adulteration detection and food quality monitoring applications, the suggested framework offers a dependable, non-destructive, and scalable method.

VI. ACKNOWLEDGMENT

The Department of Science and Technology, under the Funds for Infrastructure under Science and Technology (DST- FIST) program, has provided support to the Department of Computer Science and Information Technology at Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India. The grant with sanction no. SR/FST/ETI340/2013 has facilitated the infrastructure and resources required for conducting this research. The authors express their gratitude to the Department and University Authorities for their support in carrying out the study. Additionally, the authors extend their thanks to the Chhatrapati Shahu Maharaj Research Training & Human Development Institute (SARTHI) for their financial assistance in this endeavour.

REFERENCES

[1] Y. Huang, R. Lu, and K. Chen, "Detection of fruit ripeness using hyperspectral imaging and machine learning," *Journal of Food Engineering*, vol. 256, pp. 1–10, 2022.
[2] J. Qin, K. Chao, and M. Kim, "Hyperspectral imaging technology for food and agricultural products," *Food Engineering Reviews*, vol. 5, no. 3, pp. 157–171, 2021.
[3] S. Siddiqui and M. Dhua, "Eating artificially ripened fruits is harmful," *Current Science*, vol. 99, no. 12, pp. 1664–1668, 2010.
[4] World Health Organization, "Calcium carbide and food safety concerns," WHO Food Safety Reports, 2020.
[5] A. A. Gowen, C. P. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias, "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590–598, 2021.

[6] D. Wu, D. W. Sun, and Y. He, "Application of hyperspectral imaging for food quality analysis and adulteration detection," *Critical Reviews in Food Science and Nutrition*, vol. 59, no. 1, pp. 1–15, 2023.
[7] A. Mishra, S. Sharma, and P. Verma, "Deep learning approaches for banana quality assessment using image analysis," *Computers and Electronics in Agriculture*, vol. 185, pp. 106139, 2021.
[8] M. ElMasry, N. Wang, and C. Vigneault, "Detecting apple bruises using hyperspectral imaging," *Computers and Electronics in Agriculture*, vol. 71, pp. S27–S35, 2020.
[9] L. Sun, X. Wang, and H. Zhang, "Deep learning-based hyperspectral imaging for agricultural product classification," *IEEE Access*, vol. 11, pp. 22456–22470, 2024.
[10] R. Lu, D. Wu, and Y. He, "Spectral database systems for food quality monitoring," *Journal of Food Measurement and Characterization*, vol. 15, no. 4, pp. 2100–2112, 2022.
[11] P. Geladi and H. Grahn, *Hyperspectral Image Analysis*, Wiley Publications, 2021.
[12] Y. Huang, R. Lu, and K. Chen, "Detection of fruit ripeness using hyperspectral imaging and machine learning," *Journal of Food Engineering*, vol. 256, pp. 1–10, 2022.
[13] A. Mishra, S. Sharma, and P. Verma, "Deep learning approaches for banana quality assessment using image analysis," *Computers and Electronics in Agriculture*, vol. 185, pp. 106139, 2021.
[14] J. Qin, K. Chao, and M. Kim, "Hyperspectral imaging technology for food and agricultural products," *Food Engineering Reviews*, vol. 5, no. 3, pp. 157–171, 2021.
[15] R. Lu, D. Wu, and Y. He, "Spectral database systems for food quality monitoring," *Journal of Food Measurement and Characterization*, vol. 15, no. 4, pp. 2100–2112, 2022.
[16] A. A. Gowen, C. P. O'Donnell, P. J. Cullen, G. Downey, and J. M. Frias, "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590–598, 2021.
[17] D. Wu, D. W. Sun, and Y. He, "Machine learning and hyperspectral imaging for chemically ripened fruit detection," *IEEE Access*, vol. 11, pp. 22456–22470, 2024.
[18] M. ElMasry, N. Wang, and C. Vigneault, "Detecting apple bruises using hyperspectral imaging," *Computers and Electronics in Agriculture*, vol. 71, pp. S27–S35, 2020.
[19] L. Sun, X. Wang, and H. Zhang, "Deep learning-based hyperspectral imaging for agricultural product classification," *IEEE Access*, vol. 11, pp. 22456–22470, 2024.
[20] Y. He, F. Liu, and D. Wu, "Deep learning applications in hyperspectral food analysis," *Food Analytical Methods*, vol. 16, no. 5, pp. 1123–1138, 2023.
[21] P. Geladi and H. Grahn, *Hyperspectral Image Analysis*, Wiley Publications, 2021.
[22] T. Chen, Y. Zhang, and H. Li, "Random forest based spectroscopy analysis for food adulteration detection," *Sensors*, vol. 22, no. 14, pp. 5400–5415, 2022.
[23] J. Workman and L. Weyer, *Practical Guide to Interpretive Near-Infrared Spectroscopy*, CRC Press, 2021.
[24] Food and Agriculture Organization, "Food safety and quality monitoring using advanced sensing technologies," FAO Technical Report, 2023.