Pakistan Journal of Scientific Research
Multidisciplinary & Peer Reviewed
Volume: 4, Number: 2(Suppl.), Pages: 42-55, Year:2025

Article

Non-Destructive Quality Grading of Apples and Oranges using Hyperspectral Imaging and Machine Learning

Muhammad Aqeel[1,*], Muhammad Allah Razi[1], Syed Aun Muhammad[1], Umair Siddique[1], Ahmed Sohaib[1], Musa Ahmad[1] and Muhammad Naeem[1]

[1] Advance Image Processing Research Lab (AIPRL), Institute of Computer and Software Engineering, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Punjab, Pakistan
* Correspondence: engr.aqeel.kfueit@gmail.com

Submitted: 22-02-2025**, Revised:** 10-05-2025, **Accepted:** 01-06-2025

Abstract

Hyperspectral Imaging (HSI) offers a powerful non-destructive technique for fruit quality evaluation by simultaneously capturing both spatial and spectral information. In this research, apples and oranges were analyzed using a 224-band visible–near-infrared (400–1000 nm) imaging system with a controlled illumination setup and a standardized region-of-interest (ROI) extraction protocol. To enhance signature quality, preprocessing methods such as Savitzky–Golay smoothing and median filtering were applied, effectively reducing noise while retaining meaningful spectral characteristics. The dataset comprised 34,394 spectra (224 bands each), including 19,595 fresh and 14,799 spoiled samples. Class balance was preserved during training and testing to avoid bias. Several supervised learning algorithms—k-nearest neighbors (KNN), support vector machines (SVM), random forests (RF), decision trees (DT), linear discriminant analysis (LDA), and logistic regression (LR)—were employed to classify fruit freshness both within individual species and across a four-class scenario (apple-fresh, apple-spoiled, orange-fresh, orange-spoiled). Model performance was assessed using accuracy, precision, recall, and the F1-score. The top-performing classifiers achieved nearly 99% accuracy with balanced error rates across all categories. These findings demonstrate that the integration of HSI with supervised machine learning (ML) provides a reliable and automated solution for grading apples and oranges. The approach is highly applicable to real-time sorting lines, supporting improved quality assurance, food safety, and waste reduction. Furthermore, the methodology is adaptable to other agricultural commodities and can be optimized for industrial deployment through band selection and lightweight models that reduce computational requirements while maintaining classification accuracy.
Keywords: Artificial intelligence; classification; fruit quality; hyperspectral imaging; machine learning.

1. Introduction

Global demand for quality fruits has risen in recent years, propelled by growing consumer understanding of nutrition, food safety, and general health [1]. Fruits are rich in vitamins, fiber, and antioxidants and are a critical component of human health, such that their quality is a primary focus for producers, distributors, and consumers [2]. Ensuring consistent fruit quality, however, particularly at industrial levels, is still a significant challenge [3]. Conventional inspection techniques, which mainly rely on human vision or rudimentary imaging systems, are subjective, time-consuming, and unable to detect internal defects like bruising, over-ripeness, or early spoilage [4]. These shortcomings not only lower efficiency but also result in considerable economic loss as well as low consumer confidence when poor-quality products are released into the market [5]. Amidst all these challenges, hyperspectral imaging (HSI) has arrived as a formidable, non-invasive technology, recording spatial information as well as spectral data along hundreds of thin wavelength bands [6]. HSI contrasts with ordinary RGB cameras by granting the capacity for picking up microscopic changes in biochemical composition in the fruits, marking internal quality determinants such as sugar levels, acidity, and firmness [7]. This high-density spectral information introduces new opportunities for automated fruit sorting systems that are able to recognize defects not discernible by human vision [8]. Nevertheless, the intricate nature and elevated dimensionality of hyperspectral information necessitate the use of complex analysis tools that can derive worthwhile insights in real time [9]. The comprehensive literature review of fruit quality assurance is given below:

Hyperspectral Imaging (HSI) combined with ML models has shown significant effectiveness in fruit disease detection and quality assessment in a related study. Chun et al. [10] used 1D-CNN and ResNet-50 with HSI (380–1030 nm) to detect Botrytis cinerea in strawberries with 96.86% accuracy, while Genangeli et al. [11] used ANN-AP and BC models for moldy core detection in apples (863.38–877.69 nm) with 97% accuracy. SVM-based analysis by Feng et al. [12] identified bruising in strawberries with 85% accuracy, and Siedliska et al. [13] reached 97% accuracy for fungal infection detection using BPNN across 705–2239 nm. Gao et al. [14] estimated strawberry ripeness in the field with 98.6% accuracy using SVM on HSI data (370–1015 nm). Ye et al. [15] used various models, including CNN and ResNet, to detect pesticide residues in grapes (915–1699 nm) with 93% accuracy. Chu et al. [16] achieved 94.35% accuracy in banana maturity classification using PLSDA (400–1000 nm), while Pu et al. [17] applied spatial-measured Vis-HSI with k-NN and PLSDA to classify ripeness (400–740 nm) at 93.3% accuracy. Peach fungal disease classification reached 92.5% with SPA-PLSDA in the 400–1000 nm range [18]. Mango damage evaluation by Xu et al. [19] using RMSEP in the 900–1700 nm range showed 77.8% accuracy. Apple bruise detection with AlexNet-SVM achieved 87.5% accuracy within 400–700 nm [20], and Xu et al. [21] predicted Kyoho grape shelf-life with 98.125% accuracy using SVM and CNN on HSI data (400.68–1001.61 nm).

Machine learning (ML) methods, especially supervised methodologies such as Support Vector Machine (SVM), Random Forest, Decision Tree, K-Nearest Neighbors (KNN), and Logistic Regression, present a data-oriented solution to the issue [22]. When trained on hyperspectral datasets with labels, these models are capable of classifying fruits both by their external appearance and internal structure [23]. In this research, more than 7.7 million spectral data points from apples and oranges are employed to train and test these algorithms, showing their capability to surpass conventional methods in speed, accuracy, and consistency [24]. Advanced preprocessing operations—like denoising, normalization, and dimensionality reduction—are utilized to enhance model robustness and scalability [25]. The combination of HSI with ML not only makes fruit sorting systems more accurate but also fits with the worldwide transition toward smart agriculture and automated food processing [26]. Reduced human error, waste minimization, and providing safer, better-quality produce assist in larger agendas of food security and sustainability [27][28].

This research contributes to the growing body of evidence that intelligent, technology-driven solutions are essential for meeting the evolving demands of modern agriculture and delivering safe, nutritious fruits to markets worldwide. The contribution of this research work is as follows. Created a high-quality spectral dataset for fresh and spoiled fruits with the Specim FX-10 HSI system, recording 224 spectral bands per sample to capture detailed biochemical and physical alterations due to spoilage. Used sophisticated data preprocessing methods such as the Median filter, Savitzky-Golay smoothing, Butterworth filter, Gaussian filter, and Moving Average filter to eliminate noise and improve the quality of the spectral signal. Applied predictive modeling based on ML algorithms, including LDA, SVM, RF, DT, KNN, and LR, with good classification accuracy on LDA and SVM and excellent generalization on RF and KNN.

2. Materials and Methods

A fruit sorting and classification system needs to be efficient for the purpose of ensuring the quality and uniformity of agricultural products. In this study, HSI is combined with supervised ML algorithms to classify two fruit varieties: Apple and Orange. The research methodology follows (see Figure 1) structured approach to guarantee accuracy and consistency of classification. A collection of fruit samples from local markets was done to capture the full spectrum of fruit variations in terms of quality, ripeness, and size. A hyperspectral camera (Specim FX10) was used for the imaging process in a push broom/line scanner system. The spectral data acquired in this system covers 224 spectral bands in a range of wavelengths 400–1000 nm at a spatial resolution of 445 × 512 pixels. The spectral data were controlled with lighting conditions to minimize external variability. Highly resolute spatial and spectral information was obtained through scanning each fruit individually. Assessment of the acquired data was done by means of specialized imaging software (Lumo Scanner), which stored and processed the data for further analysis.

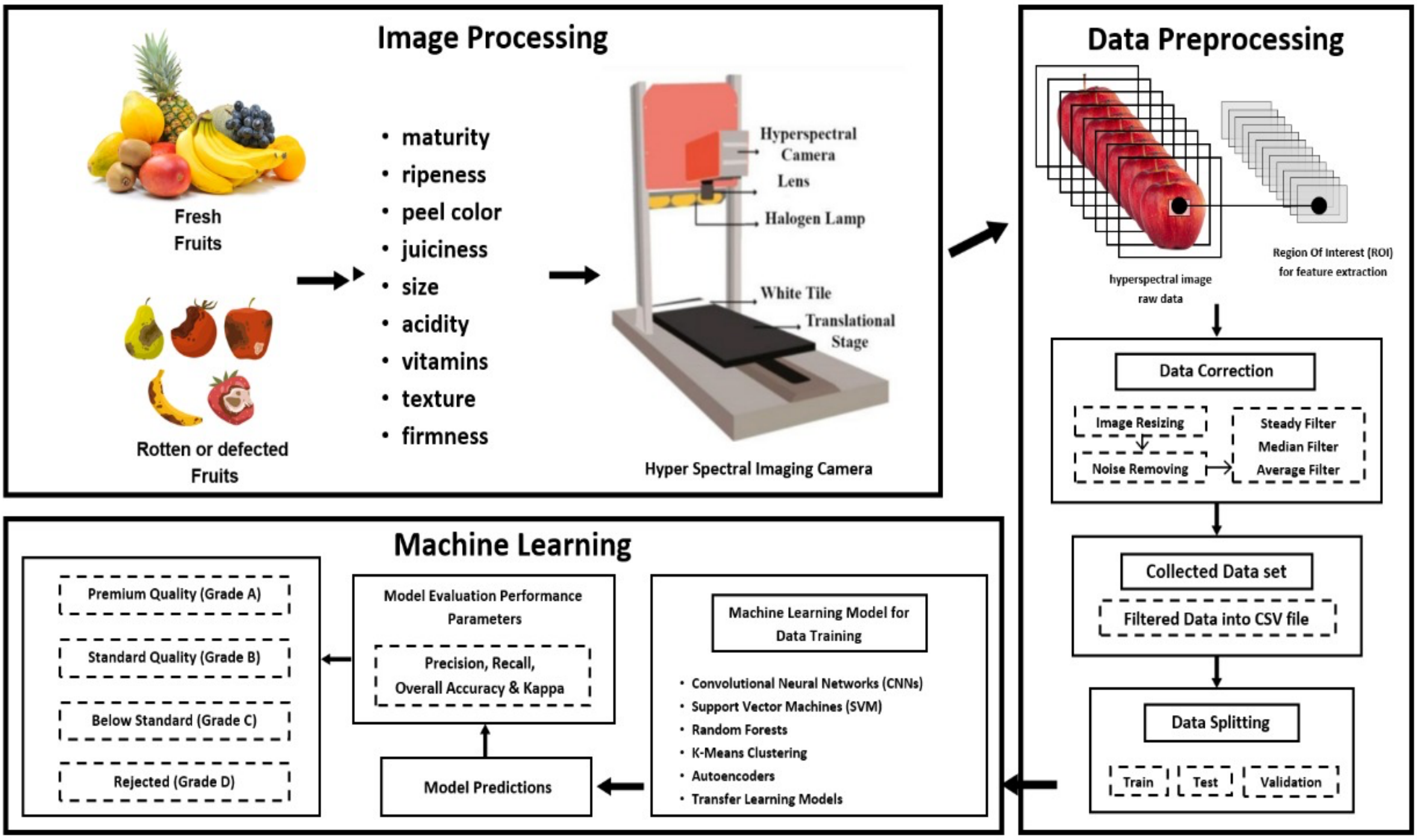


Figure 1. Working flow diagram of HSI for fruit quality assurance.

HSI provided both spatial and spectral information; reflectance spectra were computed for each fruit type. High-quality representative data was extracted by performing a careful selection of Regions of Interest (ROI). Since an HSI image has a unique spectral signature for each pixel, the mean reflectance spectrum per ROI was determined. Each ROI provided 224 spectral features to represent the spectral signature for each fruit type. Preprocessing techniques, such as noise reduction and smoothing, were applied for the purpose of enhancing data quality. Median filtering, Savitzky-Golay filtering, and steady filtering were addressed in it. Of these filters, the Savitzky-Golay filter was most effective at reducing noise without loss of spectral integrity. After preprocessing was done on the dataset, it was split into training (80%) and testing (20%) subsets, in order to ensure training of the model and its validation were on a balanced format. After obtaining the spectral data, it was converted to a text or CSV format to be seamlessly integrated with the machine learning framework. To train classification models of classification, various supervised ML algorithms like SVM, LR, LDA, RF, DT, and KNN were implemented. To recognize the different fruit types, these models were trained to recognize them depending on their spectral signature. Evaluation was done to test the performance of models using accuracy, recall, precision, and F1 score. LDA was found to be the most successful classifier among all tested, achieving better accuracy and better robustness across all evaluation metrics. Data has shown that HSI, along with ML techniques, are promising means to build an automatic yet scalable fruit sorting and classification system, which has great significance to modern agriculture and the food industry.

2.1 Dataset Preparation

Apple and orange samples were purchased from local markets in Rahim Yar Khan (Punjab, Pakistan). To keep the dataset consistent, we defined "fresh" and "spoiled" using straightforward visual and textural criteria (e.g., firmness, discoloration). Fresh fruit was imaged first using hyperspectral imaging (HSI). The same batches were then stored under natural conditions to allow spoilage, and we re-imaged them at intervals so the labels reflected the actual stage of deterioration. As a focused proof of concept, the study was limited to apples and oranges, with comparable numbers of fresh and spoiled samples for each. All images were acquired with an HSI camera capturing 224 spectral bands across 400–1000 nm under fixed illumination to minimize variability. We selected Regions of Interest (ROIs) conservatively in each hyperspectral cube to extract clean, representative spectra. Each pixel carries its own reflectance signature, allowing fresh and spoiled tissue to be separated when the spectra are aggregated over the ROI. Figure 2 summarizes the processed dataset: 7,704,256 spectral values arranged as 224 bands by 34,394 samples. Of these samples, 19,595 are labeled fresh and 14,799 are labeled spoiled, providing substantive coverage for both fruit types. The full spectral range (400–1000 nm) is retained after preprocessing so that downstream models see the informative variation rather than noise. This large corpus was then split into training and test partitions and used to fit and evaluate standard supervised classifiers. Together, hyperspectral imaging and machine learning enable accurate, automated grading with clear relevance to agriculture, food safety, and routine quality assurance.

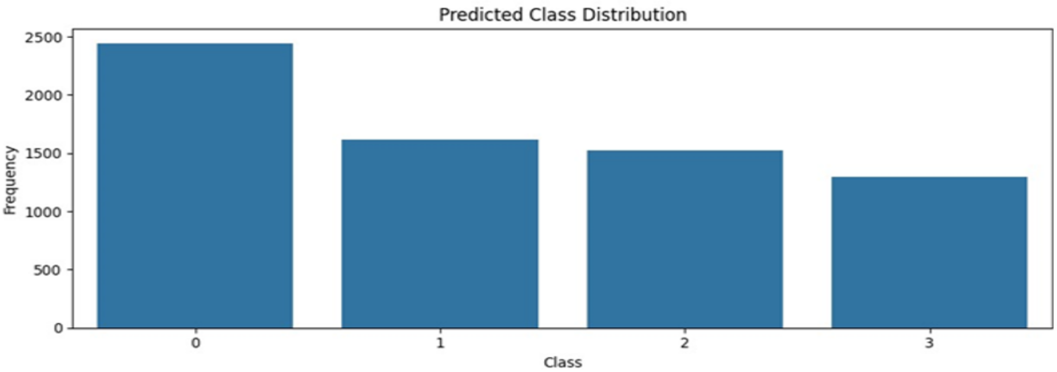


Figure 2. Predicted class distribution graph.

2.2 HSI Experimental Setup

For this research, the hyperspectral imaging (HSI) system used is Specim FX-10 camera (Specim, Spectral Imaging Ltd., Oulu, Finland; software version 1.2.0, 2017) fitted with a Schneider Kreuznach 12 mm Cinegon 1.4/12-0906 C-Mount lens and an integrated line scanner (Lumo Scanner 2017, Specim, Spectral Imaging Ltd.). The experimental setup includes three halogen lamps, a scanning platform that is mobile to obtain correct positioning and an image capture, and an adjustable height camera. The scanner is accessed through a serial communication port with a laptop to carry out trouble-free data pickup and transfer, and the Lund University Machine Operations software (Lumo Scanner) carries out the data processing stream. A computer can be linked with the camera through the connection provided by the GigE-Vision interface (Pleora Technologies) for fast communication. The outcome is three types of raw data files from the system: sample data, dark reference, and white reference. The dark reference 445 frames are obtained by closing the camera shutter, and the white reference 445 frames are taken using a calibrated white tile. The HSI system captures hyperspectral data in the form of the 445 × 512 × 224 hypercube and 397–1003 nm spectral range (cutoff wavelengths but in this study, the effective usable range of 400–1000 nm was applied for analysis, with an average subsampling interval of 2.7 nm. This configuration yields high-resolution spatial and spectral data for fresh and damaged fruit.

2.3. Spectral Reflectance Calculation

Because the capability to obtain accurate spectral reflectance data with such high spatial and spectral resolution was a prerequisite for successful classification of fresh and spoiled fruits, hyperspectral imaging (HSI) was utilized for this purpose. Calibration was done using a 99.9% white reflectance white reference tile so that the spectral measurements would be correct. For a fixed position of the camera (15 cm above the samples), the system ran at a 50 Hz frame rate and 16 ms exposure. To ensure standardized data collection for all samples, conveyor movement maintained a conveyor speed of 11.72 mm/s.

Radiance was captured by the hyperspectral sensor and affected by conditions such as lighting, viewing, and sensor. For correcting radiance values, two targets (white and dark references) were utilized, and radiance values were corrected through the Empirical Line Method (ELM) to reduce the effects of shading and background noise to yield spectral reflectance values. The formula to calculate the spectral reflectance (RR) for every wavelength band is as follows:

$$Reflectance = \frac{R_{specimen} - D_{ref}}{w_{ref} - D_{ref}} \quad (1)$$

The recorded radiance of the fruit sample $R_{specimen}$, D_{ref} , is the radiance of the dark reference, W_{ref} , is the radiance of the white reference can be calibrated into reflectance data that forms a basis to discriminate between fresh and spoiled produce through analysis of spectral signatures (See Figure 3). This approach is accurate and allows for making food quality assessments and reducing post-harvest losses.

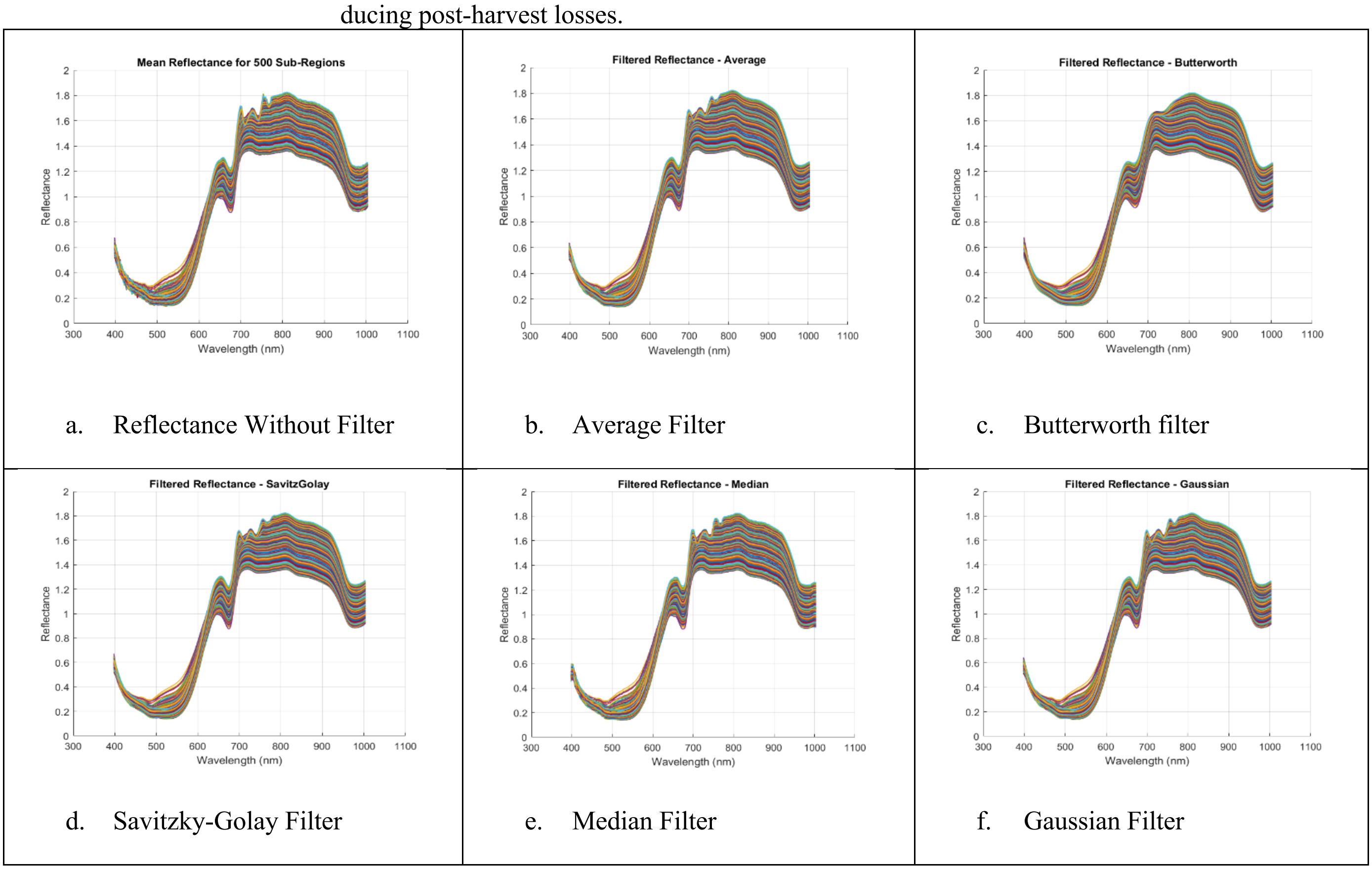
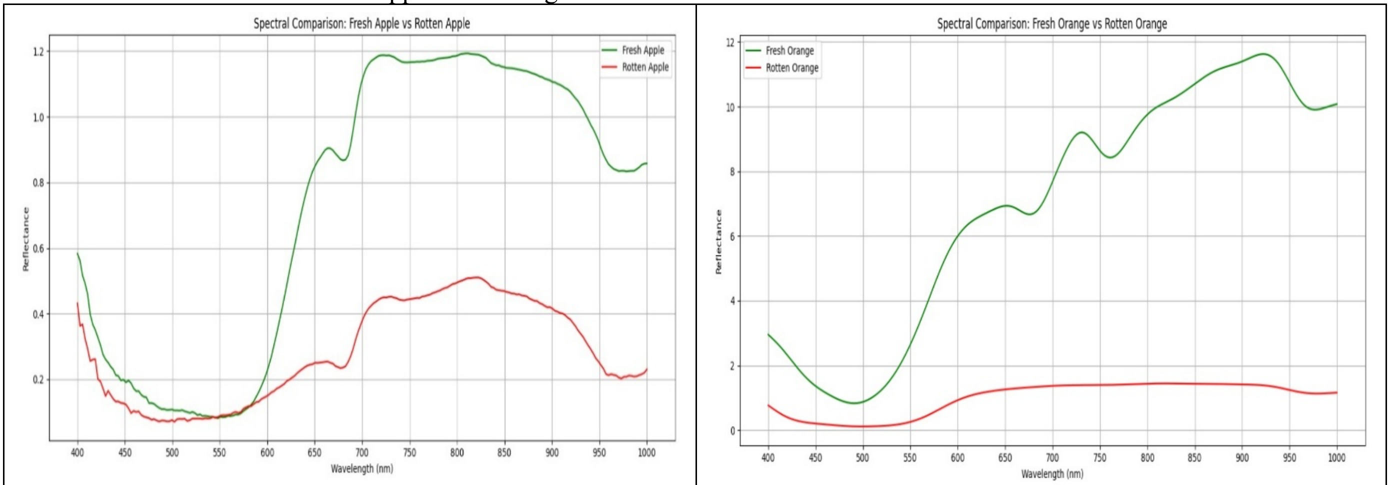


Figure 3. Processed and non-processed spectral reflectance signature.

2.4 Spectral Analysis for Fresh and Rotten Fruits

Spectral analysis allows for the detection of fruit quality changes through the analysis of distinctive reflectance patterns associated with characteristics such as ripeness, firmness, and moisture. Raw hyperspectral data, while informative, has sensor noise, lighting noise, and surface variation noise. For clarity enhancement, filters including Savitzky-Golay, Median, Gaussian, Average, and Butterworth were used. Butterworth filtering produced the smoothest, most classifiable spectra, as evident in Figure 3 illustrates all filters, demonstrating the system's flexibility to dataset-specific preprocessing requirements. Through noise reduction and the improvement of spectral clarity, these filtering methods helped greatly enhance the discrimination between quality grades so that machine learning models could detect better and more significant patterns in the spectral data. The application of filtered reflectance also helps reinforce real-time grading systems in industrial settings by providing data consistency and interpretability. Figure.4(a)and(b) show distinct spectral patterns for apples and oranges.



4(a) 4(b)

Figure 4. Apple and orange fruits' spectral analysis comparison.

2.5. Machine Learning Algorithms

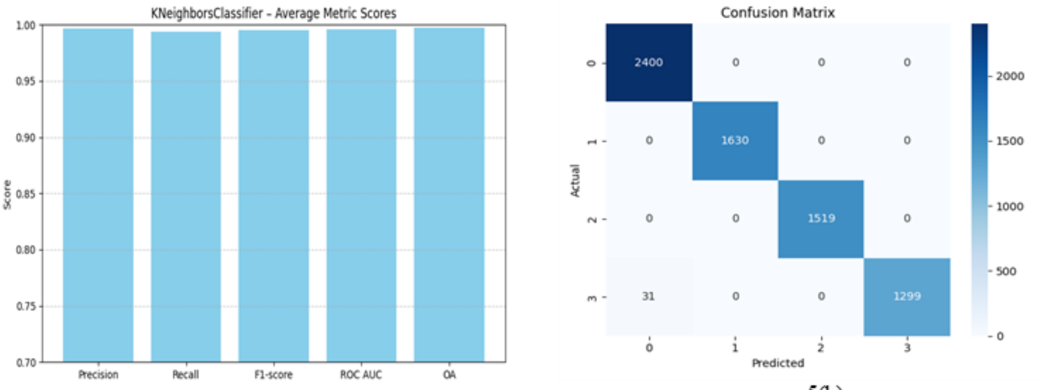
To categorize fresh and spoiled fruits based on HSI data, several machine learning algorithms were utilized to examine slight spectral differences associated with spoilage. These were LR, LDA, SVM, KNN, DT, and RF. All models were trained on the complete spectral dataset to identify minor biochemical and surface changes that indicate fruit spoilage. Of them, LDA was notable for dimensionality reduction and class separation efficiency, and SVM provided stable boundary definition among new and spoiled samples. Random Forest and Decision Tree models successfully projected reflectance changes to spoilage patterns, and KNN showed simplicity with good accuracy. Logistic Regression provides consistent performance in classification. In total, these models offered accurate and effective early spoilage detection tools, which helped promote smarter quality control and less food waste in the fruit supply chain.

3. Results and Discussion

Fruits are essential for health, but quality control is still challenging on a large scale because of the limitations of manual inspection. Rotten fruits lead to economic loss and health hazards. Hyperspectral imaging (400–1000 nm) was used in this study to non-destructively examine fruit quality from spatial and spectral information. In contrast to conventional methods, it records both external and internal characteristics. We used machine learning algorithms (KNN, LDA, LR, RF, SVM, and DT) to grade fruits into four quality grades. Performance evaluation with metrics such as accuracy, F1 score, and ROC AUC indicated good performance, validating the method's worth for intelligent, autonomous fruit grading.

3.1. Classification

The KNN classifier exhibited excellent performance, which is also observed in Figure 5(a) for fruit quality grading using hyperspectral data, with almost perfect scores for all the performance metrics. Its precision, recall, F1-score, ROC AUC, and overall accuracy (OA) were all around 100%. Figure 5(b) confusion matrix verifies KNN's accuracy, with most predictions placed correctly on the diagonal. Class 0, for example, had 2400 correct out of 2400, showing few misclassifications in all classes. Figure 6 demonstrates uniformly strong performance in precision, recall, F1-score, ROC AUC, and OA for all four classes, which reflects the model's strength and capability to generalize well across fruit varieties and quality status. The KNN model showed nearly perfect accuracy, uniformity across all classes, and no overfitting, and thus is a good option for real-time fruit quality inspection using hyperspectral imaging.



5(a) 5(b)

Figure 5. KNN model performance evaluation and confusion metric scores.

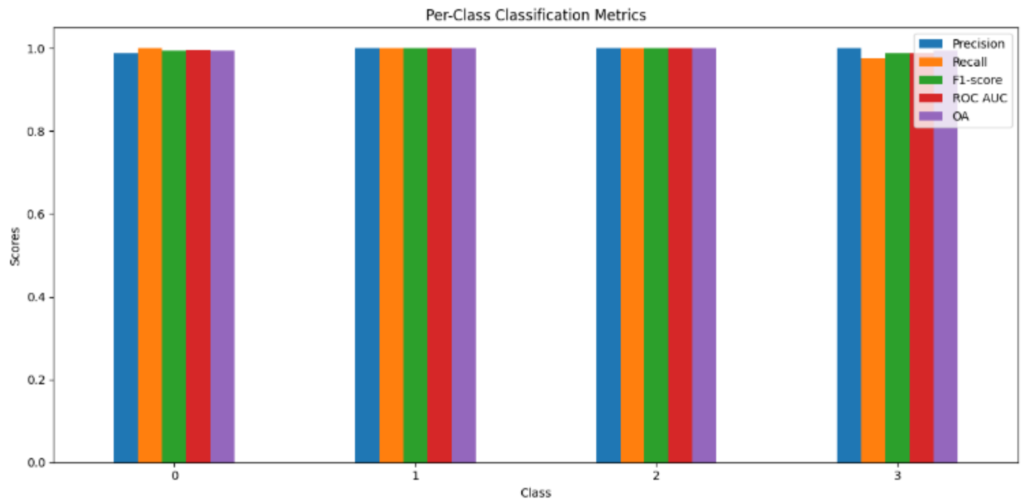
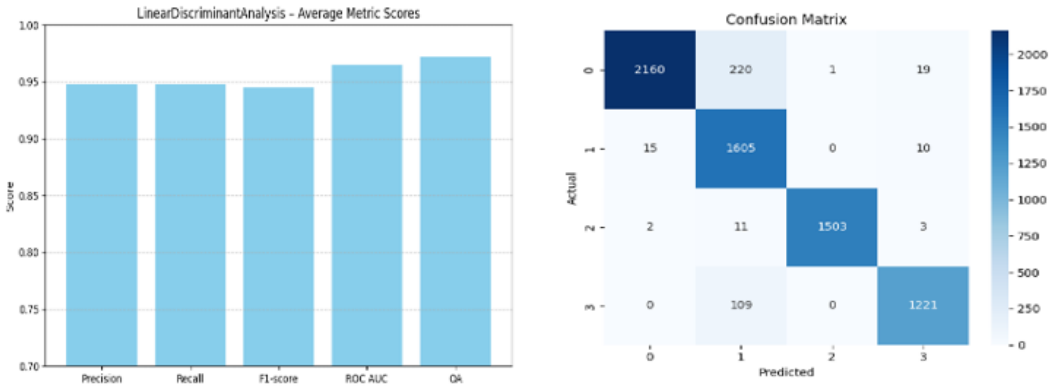


Figure 6. KNN Per-Class Classification Metrics.

The LDA model performed well in classifying fruit quality grades with precision, recall, and F1-scores less than 0.95, ROC AUC greater than 0.97, and an OA close to 0.99, showing high reliability for all 4 fruit classes. The anticipated class distribution, as depicted in Figure 7(a), closely resembles the real distribution, with class 0 being the most common, then classes 1 and 2. Small differences indicate occasional overlap, which is also observed in Figure 7(b), particularly among mid-sized classes such as class 0 and class. The confusion matrix in Figure 7(b) shows that most of the predictions were accurate, but in some instances few classes are misclassified. These indicate possible improvements in differentiating spectrally close quality stages. Figure 8 provides high, consistent performance for all classes with minor dips in precision and F1-score for a few Classes corresponding to the results of the confusion matrix. Still, most scores are above 0.90, validating the strong generalization and predictive potential of the model. LDA was found to be a very good classifier for hyperspectral signature-based fruit quality grading. Its accuracy and balanced performance are very good, and thus recommend it for deployment in real-world applications in real-time fruit sorting, particularly where simplicity of computation and speed are desired.



7(a) 7(b)

Figure 7. Linear discriminant analysis average metric scores.

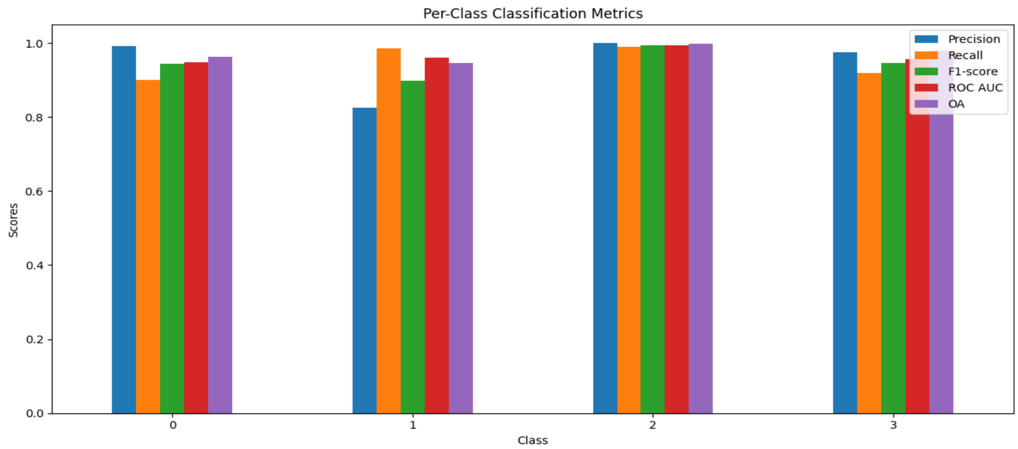
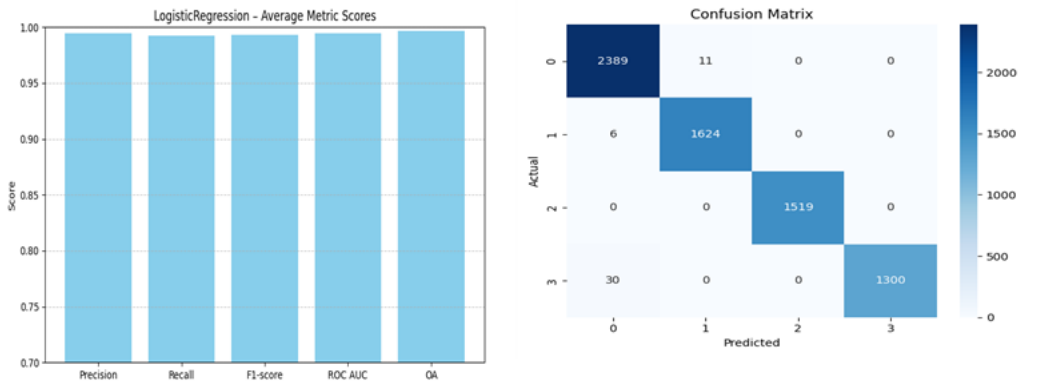


Figure 8. LDA per-class classification metrics score.

In Figure 9(a), the logistic regression model worked very well, with all the metrics—Precision, Recall, F1-score, ROC AUC, and OA—approaching or higher than 0.99 reflects its good generalization and efficacy in multi-class fruit classification, as can also be seen in Figure 9(b) of the Confusion matrix.



9(a) 9(b)

Figure 9. LR performance evaluation score and confusion matrix.

Figure 10 validates the high accuracy of the model, with predictions along the diagonal for most and very few scattered misclassifications. Class 0, for instance, has 2389 correct predictions, showing the precision of the model and low inter-class confusion. LR was a strong, high-performance classifier with nearly perfect performance for all categories of fruits. It is a good choice due to its simplicity and accuracy for real-time monitoring of fruit quality based on HSI.

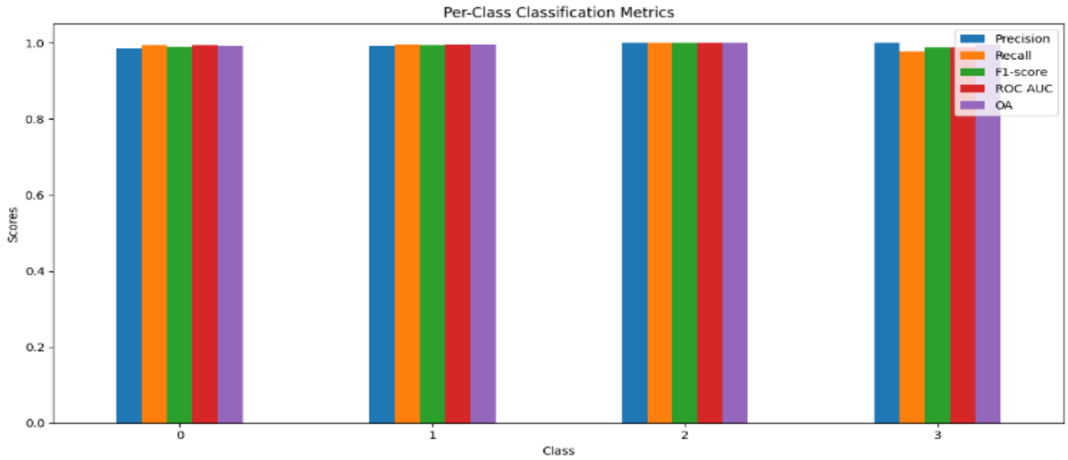


Figure 10. LR performance per-class classification metrics.

Figure 11 (a) The Decision tree model performed moderately, with a total accuracy of 99%. Lower precision (0.99) and F1-score (0.99), however, reflect difficulties in making confident and balanced classification on all fruits. Figure 11(b) confusion matrix reveals large misclassifications, specifically Class 0 being mixed up with Classes 2 and 3 more than 200 times. Comparable spillovers in Classes 0 and 1 decrease the model's reliability for precise fruit quality grading.

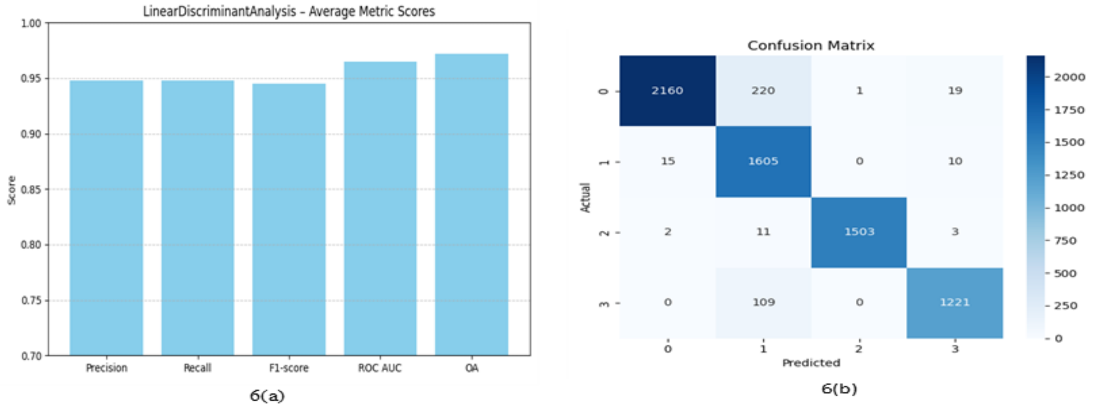


Figure 11. DT performance evaluation confusion metric scores.

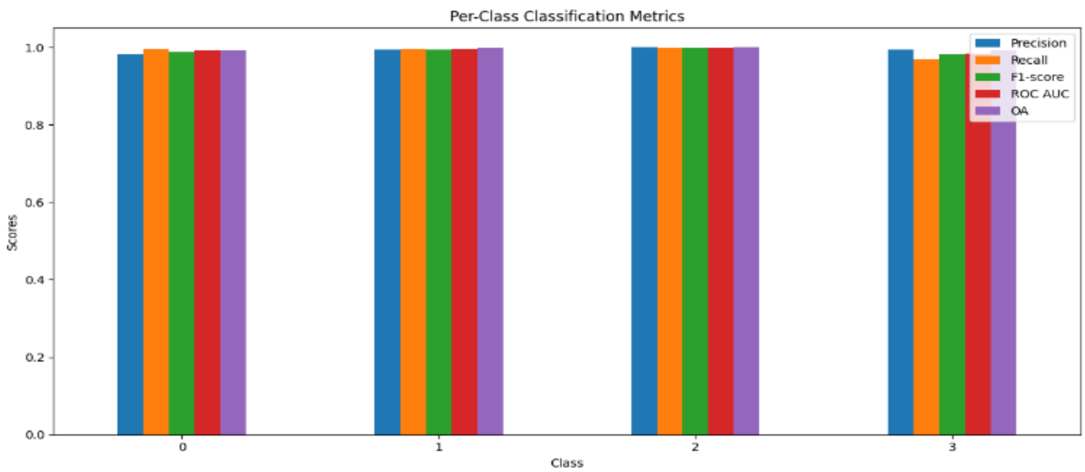
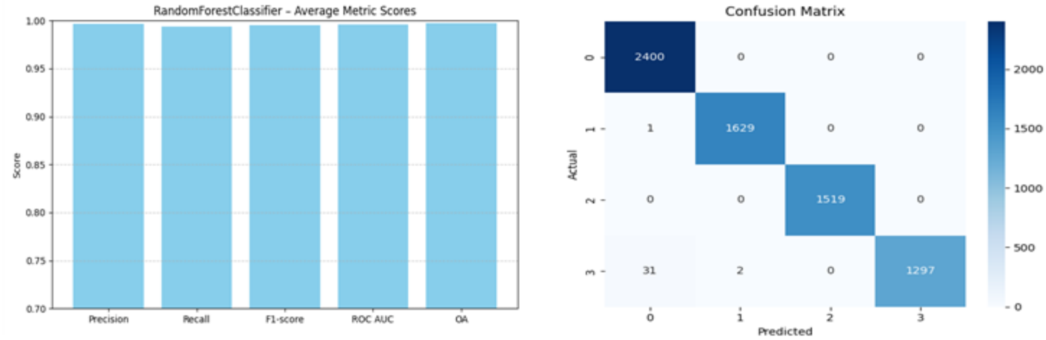


Figure 12. Decision Tree (DT) per-class classification metrics.

Random Forest classifier, as seen in Figure 13(a), performed incredibly well, and all the principal metrics—Precision, Recall, F1-Score, ROC AUC, and OA—all converged to near 1.0, clearly indicating its extreme reliability and practically perfect consistency in separating all classes. In Figure 13(b) Confusion matrix refers to strong diagonal dominance and minimum misclassification. To clarify, the predictions of 2400 True Class 0 were in a set of 2400 samples, which demonstrates the model's stability. In Figure 14, the per-class classification metrics bar chart also illustrates that all classes had high recall, precision, and F1-score with little deviation, indicating that the model's performance across all levels of adulteration is consistent.



13(a) 13(b)

Figure 13. RF performance evaluation and confusion matrix results.

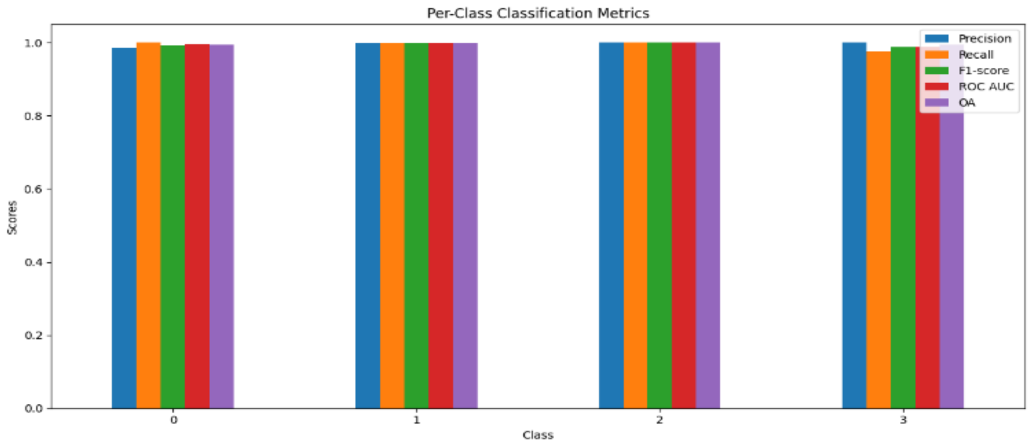
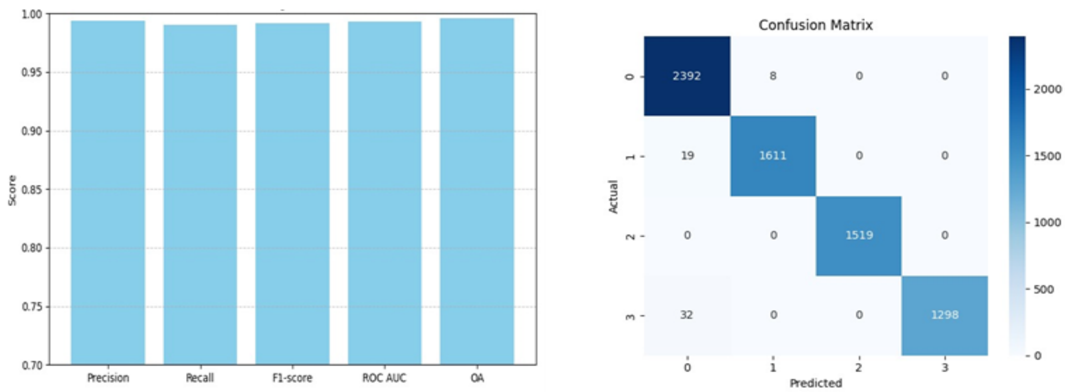


Figure 14. RF classifier per-class classification metrics.

Support Vector Machine (SVM) model performed outstandingly (Figure 15(a)), with Precision, Recall, F1-score, ROC, AUC, and Overall Accuracy all being above 0.99, evidence of its accuracy and stability for challenging tasks. İn Figure 15(b) confusion matrix reveals high diagonal values and little misclassifications, with Class 0 possessing 2392 correct classifications and just 8 misclassified samples, largely into Classes 2 and 3, consistent with other highly performing models.



15(a) 15(b)

Figure 13. SVM model overall performance and confusion matrix.

İn Figure 15(b) confusion matrix reveals high diagonal values and little misclassifications, with Class 0 possessing 2392 correct classifications and just 8 misclassified samples, largely into Classes 2 and 3, consistent with other highly performing models. In Figure 16, examining the per-class classification scores, each class attained very close and high scores on all five measures, showing that the SVM did not disproportionately favor or ignore any one class. Such balanced performance is especially useful in multi-class scenarios. Per-class scores were uniformly high the model did not disproportionately favor any class.

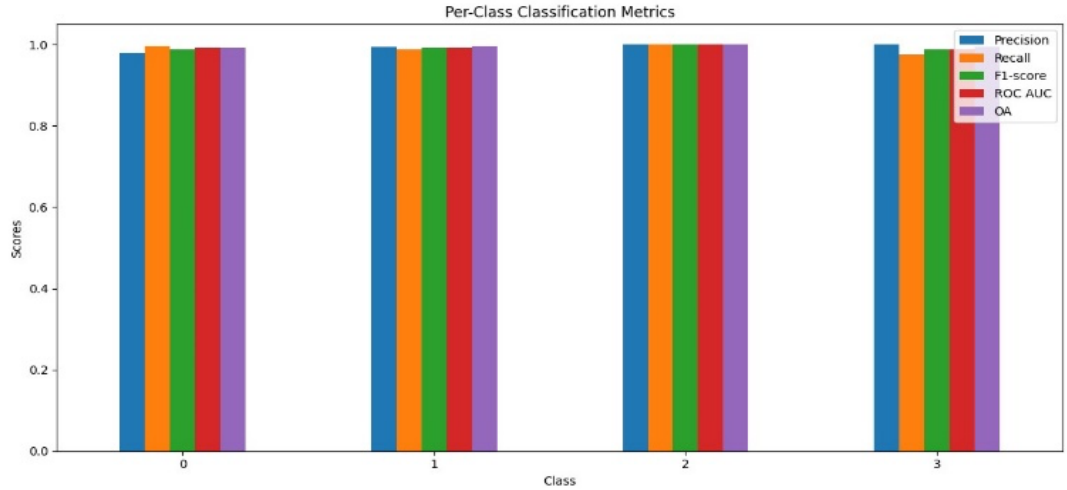


Figure 14. SVM per-class classification matrix.

4. Discussion

Table 1 presents the contrast of hyperspectral imaging (HSI)-based fruit classification methods among different studies, showing significant differences in accuracy, spectral range, and machine learning approaches. Some of the earlier studies have used models like SVM, CNN, ResNet, PLSDA, and hybrid neural networks on fruits such as strawberries, grapes, bananas, mangoes, apples, and peaches, with accuracies ranging generally between 77% and 98.6%. The used spectral ranges vary from the narrow bands, such as 400–740 nm, to wider ones extending to 2239 nm. Notwithstanding these developments, most of the models exhibited shortcomings in generalizability or consistency across quality grades and different fruit types. Conversely, this study illustrates a considerable improvement by realizing 99.1% classification accuracy for oranges and apples utilizing HSI across the range of 400–1000 nm. This research applied and contrasted a full set of supervised ML algorithms— KNN, LDA, LR, RF, SVM, and DT. These algorithms were trained and tested exhaustively on pre-processed hyperspectral data, well reflecting nuanced spectral changes associated with fruit quality. The outstanding performance of our system not only beats current benchmarks but also demonstrates the strong synergy between hyperspectral imaging and strong ML pipelines, setting a new benchmark for accurate, scalable, and non-destructive fruit quality grading.

Table 1. State-of-the-art comparison with proposed work.

Fruits	Technology	Techniques	Wavelength (nm)	Accuracy	Ref
strawberry	HSI	1D-CNN, ResNet-50	380–1030 nm	96.86%	[10]
apple	HSI	ANN-AP and BC models.	863.38–877.69 nm	97%	[11]
strawberry	HSI	SVM	380–1030 nm	85%	[12]
strawberry	HSI	backpropagation neural network (BNN) model, (TPC), (SSC)	705–2239 nm		[13]
strawberry	HSI	SVM	370 –1015 nm	98.6%	[14]
grapes	HSI	LR, SVM, RF, CNN, residual neural network (ResNet)	915-1699 nm	93%.	[15]
banana	HSI	PLSDA model	400–1000 nm	94.35%	[16]
banana	HSI	k-NN, SIMCA, PLSDA	400–740 nm	93.3%	[17]

peach	HSI	SPA-PLSDA	400–1000 nm	92.5%	[18]
mango	HSI	RMSEP	900–1700 nm	77.8%	[19]
apple	HSI	AlexNet-SVM	400–700 nm	87.5%	[20]
grapes	HSI	SVM, CNN	400.68–1001.61 nm	98.125 %	[21]
Apple, orange	HSI	SVM, KNN, LDA, DT, RF, LR	400–1000 nm	99.1%	This work

5. Conclusions

This research demonstrates that hyperspectral imaging (HSI) combined with supervised machine learning (ML) provides a robust and non-destructive solution for grading apples and oranges. Using a 224-band visible–near-infrared system (400–1000 nm) under fixed illumination and standardized region-of-interest selection, we compiled a dataset of 34,394 spectra and evaluated six widely used classifiers: k-nearest neighbors (KNN), support vector machines (SVM), random forests (RF), decision trees (DT), linear discriminant analysis (LDA), and logistic regression (LR). Across four-class classification tasks, the top-performing models achieved close to 99% accuracy, with balanced performance across categories. These findings highlight the potential of HSI to enable reliable, automated fruit grading systems that can improve quality control, reduce postharvest losses, and enhance food safety in industrial processing lines.

Future research will also focus on optimizing feature selection and developing lightweight model architectures to reduce computational cost, enabling real-time, embedded deployment. Additionally, extending the framework to other fruit types and naturally occurring defects will further strengthen its applicability to large-scale agricultural and food industry operations.

Author Contributions: Author Contributions: Conceptualization, M.A and A.S; methodology, M.A; software, U.S. and A.S.; validation, H.S, M.A, and M.N; formal analysis, H.S., M.A. and A.R; investigation, S.A; resources, M.A; data curation, A.R; writing—original draft preparation, M.N and A.S.; writing—review and editing, M.A, A.S., S.A. and H.S; visualization, M.A; supervision, A.S and S.A.; project administration, A.R; funding acquisition, M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Shahbaz, H.M.; Akram, K.; Ahn, J.J.; Kwon, J.H. Worldwide Status of Fresh Fruits Irradiation and Concerns about Quality, Safety, and Consumer Acceptance. *Crit. Rev. Food Sci. Nutr.* **2016**, *56*, 1790–1807, doi:10.1080/10408398.2013.787384.
2. Yahia, E.M.; Celis, M.E.M.; Svendsen, M. The Contribution of Fruit and Vegetable Consumption to Human Health. *Fruit Veg. Phytochem. Chem. Hum. Heal. Second Ed.* **2017**, *1*, 3–52, doi:10.1002/9781119158042.ch1.
3. Trienekens, J.; Zuurbier, P. Quality and Safety Standards in the Food Industry, Developments and Challenges. *Int. J. Prod. Econ.* **2008**, *113*, 107–122, doi:10.1016/j.ijpe.2007.02.050.
4. Wang, D.; Zhang, M.; Jiang, Q.; Mujumdar, A.S. Intelligent System/Equipment for Quality Deterioration Detection of Fresh Food: Recent Advances and Application. *Foods* **2024**, *13*, 1662, doi:10.3390/foods13111662.
5. Den Ouden, E.; Yuan, L.; Sonnemans, P.J.M.; Brombacher, A.C. Quality and Reliability Problems from a Consumer's Perspective: An Increasing Problem Overlooked by Businesses? *Qual. Reliab. Eng. Int.* **2006**, *22*, 821–838, doi:10.1002/qre.766.
6. Bhargava, A.; Sachdeva, A.; Sharma, K.; Alsharif, M.H.; Uthansakul, P.; Uthansakul, M. Hyperspectral Imaging and Its Applications: A

Review. *Heliyon* **2024**, *10*, doi:10.1016/j.heliyon.2024.e33208.

7. Martínez-Domingo, M.Á.; Valero-Benito, E.M.; Hernández-Andrés, J. Multispectral and Hyperspectral Imaging. In *Non-invasive and Non-destructive Methods for Food Integrity*; Springer, 2024; pp. 175–201 ISBN 9783031764653.
8. Maier, G.; Gruna, R.; Langle, T.; Beyerer, J. A Survey of the State of the Art in Sensor-Based Sorting Technology and Research. *IEEE Access* **2024**, *12*, 6473–6493, doi:10.1109/ACCESS.2024.3350987.
9. Signoroni, A.; Savardi, M.; Baronio, A.; Benini, S. Deep Learning Meets Hyperspectral Image Analysis: A Multidisciplinary Review. *J. Imaging* **2019**, *5*, 52, doi:10.3390/jimaging5050052.
10. Chun, S.W.; Song, D.J.; Lee, K.H.; Kim, M.J.; Kim, M.S.; Kim, K.S.; Mo, C. Deep Learning Algorithm Development for Early Detection of Botrytis Cinerea Infected Strawberry Fruit Using Hyperspectral Fluorescence Imaging. *Postharvest Biol. Technol.* **2024**, *214*, 112918, doi:10.1016/j.postharvbio.2024.112918.
11. Genangeli, A.; Allasia, G.; Bindi, M.; Cantini, C.; Cavaliere, A.; Genesio, L.; Giannotta, G.; Miglietta, F.; Gioli, B. A Novel Hyperspectral Method to Detect Moldy Core in Apple Fruits. *Sensors* **2022**, *22*, 4479, doi:10.3390/s22124479.
12. Feng, R.; Han, X.; Lan, Y.; Gou, X.; Zhang, J.; Wang, H.; Zhao, S.; Kong, F. Detection of Early Subtle Bruising in Strawberries Using VNIR Hyperspectral Imaging and Deep Learning. *Vib. Spectrosc.* **2025**, *138*, doi:10.1016/j.vibspec.2025.103786.
13. Siedliska, A.; Baranowski, P.; Zubik, M.; Mazurek, W.; Sosnowska, B. Detection of Fungal Infections in Strawberry Fruit by VNIR/SWIR Hyperspectral Imaging. *Postharvest Biol. Technol.* **2018**, *139*, 115–126, doi:10.1016/j.postharvbio.2018.01.018.
14. Gao, Z.; Shao, Y.; Xuan, G.; Wang, Y.; Liu, Y.; Han, X. Real-Time Hyperspectral Imaging for the in-Field Estimation of Strawberry Ripeness with Deep Learning. *Artif. Intell. Agric.* **2020**, *4*, 31–38, doi:10.1016/j.aiia.2020.04.003.
15. Ye, W.; Yan, T.; Zhang, C.; Duan, L.; Chen, W.; Song, H.; Zhang, Y.; Xu, W.; Gao, P. Detection of Pesticide Residue Level in Grape Using Hyperspectral Imaging with Machine Learning. *Foods* **2022**, *11*, 1609, doi:10.3390/foods11111609.
16. Chu, X.; Miao, P.; Zhang, K.; Wei, H.; Fu, H.; Liu, H.; Jiang, H.; Ma, Z. Green Banana Maturity Classification and Quality Evaluation Using Hyperspectral Imaging. *Agric.* **2022**, *12*, 530, doi:10.3390/agriculture12040530.
17. Pu, Y.Y.; Sun, D.W.; Buccheri, M.; Grassi, M.; Cattaneo, T.M.P.; Gowen, A. Ripeness Classification of Bananito Fruit (Musa Acuminata, AA): A Comparison Study of Visible Spectroscopy and Hyperspectral Imaging. *Food Anal. Methods* **2019**, *12*, 1693–1704, doi:10.1007/s12161-019-01506-7.
18. Sun, Y.; Wei, K.; Liu, Q.; Pan, L.; Tu, K. Classification and Discrimination of Different Fungal Diseases of Three Infection Levels on Peaches Using Hyperspectral Reflectance Imaging Analysis. *Sensors (Switzerland)* **2018**, *18*, 1295, doi:10.3390/s18041295.
19. Xu, D.; Wang, H.; Ji, H.; Zhang, X.; Wang, Y.; Zhang, Z.; Zheng, H. Hyperspectral Imaging for Evaluating Impact Damage to Mango According to Changes in Quality Attributes. *Sensors (Switzerland)* **2018**, *18*, 3920, doi:10.3390/s18113920.
20. Chen, J.F.; Ding, Z.; Song, J.Y.; Wang, Y.; Bian, L.F.; Yang, C. Low-Cost Monochromatic Uniform Illumination HSI System for Detection and Classification of Apple Bruise. *J. Food Meas. Charact.* **2024**, *18*, 4853–4864, doi:10.1007/s11694-024-02540-8.
21. Xu, M.; Sun, J.; Cheng, J.; Yao, K.; Shi, L.; Zhou, X. Non-Destructive Estimation for Kyoho Grape Shelf-Life Using Vis/NIR Hyperspectral Imaging and Deep Learning Algorithm. *Infrared Phys. Technol.* **2024**, *142*, 105532, doi:10.1016/j.infrared.2024.105532.
22. Pargent, F.; Pfisterer, F.; Thomas, J.; Bischl, B. Regularized Target Encoding Outperforms Traditional Methods in Supervised Machine Learning with High Cardinality Features. *Comput. Stat.* **2022**, *37*, 2671–2692, doi:10.1007/s00180-022-01207-6.
23. Zhao, Y.; Kang, Z.; Chen, L.; Guo, Y.; Mu, Q.; Wang, S.; Zhao, B.; Feng, C. Quality Classification of Kiwifruit under Different Storage Conditions Based on Deep Learning and Hyperspectral Imaging Technology. *J. Food Meas. Charact.* **2023**, *17*, 289–305, doi:10.1007/s11694-022-01554-4.
24. Safari, Y.; Nakatumba-Nabende, J.; Nakasi, R.; Nakibuule, R. A Review on Automated Detection and Assessment of Fruit Damage Using Machine Learning. *IEEE Access* **2024**, *12*, 21358–21381, doi:10.1109/ACCESS.2024.3362230.
25. Sivathapandi, P.; Krishnaswamy, P.; Muthusubramanian, M. Advanced AI Algorithms for Automating Data Preprocessing in Healthcare: Optimizing Data Quality and Reducing Processing Time. *J. Sci. Technol. ISSN 2582* **2022**, *6921*.
26. Khan, A.; Vibhute, A.D.; Mali, S.; Patil, C.H. A Systematic Review on Hyperspectral Imaging Technology with a Machine and Deep Learning Methodology for Agricultural Applications. *Ecol. Inform.* **2022**, *69*, 101678, doi:10.1016/j.ecoinf.2022.101678.

27. Farooq, M.; Rehman, A.; Pisante, M. Sustainable Agriculture and Food Security. In *Innovations in Sustainable Agriculture*; Elsevier, 2019; pp. 3–24 ISBN 9783030231699.
28. Reddy, K.S.; Ahmad, S.S.; Tyagi, A.K. Artificial Intelligence and the Internet of Things-Enabled Smart Agriculture for the Modern Era. In *AI Applications for Business, Medical, and Agricultural Sustainability*; IGI Global, 2024; pp. 68–99 ISBN 9798369352687.