Journal of Spectroscopy

WILEY

RESEARCH ARTICLE OPEN ACCESS

Framework to Detect Adulterated Honey Using Hyperspectral Imaging and Machine Learning

Kiran Yaqub[1] | Mutiullah Jamil[1] | Muhammad Nauman Saeed[1] | Saqib Ubaid[1] | Arif Mehmood[2] | Monica Gracia Villar[3,4,5] | Eduardo Silva Alvarado[3,6,7] | Imran Ashraf[8]

[1]Institute of Computer Science, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan | [2]Department of Computer Science & Information Technology, The Islamia University of Bahawalpur, Bahawalpur, Pakistan | [3]Department of Social Science and Humanities, Universidad Europea Del Atlantico, Santander 39011, Spain | [4]Department of Project Management, Universidad Internacional Iberoamericana, Campeche 24560, Mexico | [5]Department of Projects, Universidade Internacional Do Cuanza, Cuito, Bie, Angola | [6]Department of Projects, Fundacion Universitaria Internacional de Colombia, Bogota, Colombia | [7]Department of Projects, Universidad de La Romana, La Romana, Dominican Republic | [8]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan-si, Gyeongbuk, South Korea

Correspondence: Imran Ashraf (imranashraf@ynu.ac.kr)

Received: 20 January 2026 | **Revised:** 8 August 2026 | **Accepted:** 11 August 2026

Academic Editor: Pramita Mishra

Keywords: adulterated honey | hyperspectral imaging | image processing | machine learning

ABSTRACT

Honey is a natural product renowned for its nutritional and medicinal properties. However, the increasing market demand in Pakistan has led to frequent instances of adulterated and fraudulent honey. Conventional detection methods are widely used but are often ineffective in accurately identifying adulteration. This study proposes a novel approach to detecting adulteration in honey, specifically the mixing of sugar, using machine learning (ML) techniques in conjunction with hyperspectral imaging (HSI). By leveraging HSI, we capture distinct spectral features of both pure and adulterated honey. These spectral features are processed and analyzed using ML algorithms trained on a dataset comprising pure and mixed honey samples. Experiments involve binary, as well as, multiclass (5 classes) honey samples with different levels of adulteration. In addition, data balancing is also carried out using synthetic minority oversampling technique (SMOTE). The results prove the proposed approach to be efficient in detecting adulteration across various honey varieties, with linear regression achieved 99.9% while support vector machine and multilayer perceptron obtaining a 98.7% accuracy for binary class. For the multiple classes of honey, the multilayer perceptron shows a 96.67% accuracy outperforming the remaining ML and deep learning models. Among the deep learning models, the convolutional neural network (CNN) achieved the best performance with an accuracy of 94.67% after applying the SMOTE $\times$ 3 (three times of over samples of original samples of each class) data balancing strategy, whereas recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU) models showed comparatively lower performance.

1 | Introduction

Honey, a natural product celebrated for its nutritional and medicinal properties, is increasingly subjected to adulteration to meet rising market demands. Conventional methods for detecting adulterated honey, such as visual inspection and basic tests, are often ineffective and unreliable. Adulteration, commonly involving the addition of commercial sugar syrups or blending with inferior-quality honey, can have health risks, including elevated blood sugar levels, obesity, diabetes, and hypertension [1]. These unethical practices also undermine consumer trust and the authenticity of honey.

Hyperspectral imaging (HSI) and machine learning (ML) offer innovative solutions to address this challenge. HSI captures images at multiple wavelengths across the visible and near-infrared spectra (VNIR), providing detailed spectral information for every pixel [2]. These spectral data enable precise differentiation between pure and adulterated honey. ML algorithms, trained on datasets containing spectral features of pure and adulterated honey samples, can accurately classify and quantify adulteration levels [3].

The integration of ML and advanced imaging techniques has significantly advanced classification and detection tasks across various domains. These methods are increasingly being applied to land use classification, crop analysis, medical imaging, and anomaly detection.

Honey adulteration is a global issue, with implications for consumer trust and public health. Advanced technologies such as HSI and ML have proven to be effective tools for addressing this challenge. Numerous studies have explored their application to detect and quantify adulteration in honey. ML techniques have been extensively applied to classify honey adulteration. For example, studies have shown the effectiveness of algorithms such as support vector machine (SVM), random forest (RF), and neural networks (NNs) in classifying honey adulterated with fructose and other syrups. These methods utilize features such as pH, electrical conductivity, and spectral reflectance for classification. Furthermore, [4] emphasized the potential of nondestructive technologies like Fourier transform infrared spectroscopy combined with ML to detect adulterants in honey, achieving accuracies above 90%.

HSI provides significant advantages for food quality assurance by capturing detailed spectral and spatial data. Studies have demonstrated its utility for detecting adulteration in honey by analyzing spectral differences at various syrup concentrations. For instance, researchers employed HSI to classify honey adulteration with an accuracy exceeding 95%, using supervised learning like linear discriminant analysis (LDA) and NN [5]. These approaches underline the versatility of HSI to discriminate the pure honey from adulterated honey across multiple concentrations.

The integration of HSI with ML has extended beyond honey adulteration to related applications in food quality and safety. Hyperspectral microscopy has been used to analyze adulterated honey by classifying it according to syrup concentration. These methods rely on advanced algorithms like SVM and NNs to achieve high precision and recall scores, facilitating reliable and efficient detection of adulteration [6].

The reviewed literature highlights the growing adoption of HSI and ML techniques in detecting honey adulteration. These methods offer significant advantages in terms of nondestructiveness, accuracy, and adaptability to diverse adulteration scenarios. This study builds on these findings by employing HSI and ML for honey adulteration detection, focusing on spectral feature extraction, noise reduction, and model optimization.

The reviewed literature underscores the value of advanced imaging and ML techniques in enhancing classification accuracy across diverse applications. This research builds on these findings by focusing on the integration of HSI and ML for honey adulteration detection, contributing to the growing field of food quality assurance [7].

Lastra-Mejias et al. [8] proposed a novel method for adulteration detection in honey by employing algorithms that rely on chaotic parameters. This technique was applied in conjunction with laser-induced breakdown spectroscopy (LIBS) to classify honey based on its botanical origin and detect possible fraud. The study reported a high success rate, achieving over 90% accuracy when analyzing honey of various botanical origins and exceeding 95% when individual honey types were examined separately. Despite the promising results, the method faced limitations in distinguishing small quantities of adulterants, such as rice syrup, and encountered challenges in categorizing honey according to its botanical source due to the complexity and similarity of spectral patterns.

Li et al. [9] investigated the application of HSI in combination with convolutional neural network (CNN) to accurately detect adulteration in Atlantic salmon, showcasing the potential of this approach for food authenticity verification. Although the study focused on fish, its findings are relevant to honey adulteration research, as they highlight the effectiveness of HSI when integrated with deep learning (DL) models for detecting subtle compositional changes. The HSI technique captures spatial and spectral information simultaneously, allowing for comprehensive analysis of the sample. While the study reported promising results, a notable limitation of the HSI approach is its requirement for complex setup and specialized equipment, which can be expensive and not readily accessible to all users. Additionally, the effectiveness of HSI may vary depending on specific imaging parameters and sample conditions.

Mendes and Duarte [10] presented a comprehensive review highlighting the utility of mid-infrared (MIR) spectroscopy as a powerful tool for food analysis across various products, including honey. The study emphasized the effectiveness of combining MIR spectroscopy with chemometric approaches to reliably assess food authenticity and detect adulteration. This integrated method enables the extraction of meaningful patterns from complex spectral data, making it a quick and robust technique for identifying tampered food samples. However, the study also noted key challenges, such as overlapping spectral bands and poor resolution in the near-infrared (NIR) region, which can hinder the accurate interpretation of results and limit the extraction of relevant information from large spectral datasets.

Siano et al. [11] investigated HSI for improving the classification performance of food adulteration detection models. While the study focused on enhancing model accuracy, it did not explicitly report the specific outcomes of using HSI for detecting honey adulteration. Nevertheless, it acknowledged the effectiveness of HSI in capturing both spatial and spectral information, which can improve adulteration detection depending on the selected parameters and experimental conditions. One of the primary limitations identified was the complexity and cost associated with hyperspectral systems, which often require specialized equipment and setup. This can restrict accessibility and practical deployment, especially for routine or large-scale testing.

Calle et al. [12] developed an automated approach to detect adulteration in high-quality honey. The authors combine visible and near-infrared spectroscopy (Vis–NIRs) with ML for this purpose. The study evaluated the performance of several models, including SVM and RF. SVR delivered strong predictive capabilities with a better (R^2) of 0.991 than other studies. In addition, a root-mean-squared error (RMSE) of 1.894 proves its efficiency. These results underline the potential of Vis–NIR spectroscopy coupled with ML for accurate, nondestructive honey adulteration detection. However, the study did not consider factors like the generalizability and robustness of the proposed models in real-world applications.

Vis–NIR spectroscopy has also been utilized to detect honey adulteration. For example, the study [13] proposed optimal subspace wavelength reduction (OSWR) approach for efficient feature extraction to integrate it with Vis–NIR. With feature dimensions reduced to 39 wavelengths, various ML models show improved performance, reporting the highest accuracy of 96.67%. In addition, honey adulteration levels are also predicted.

The use of Fourier transform has also been investigated for detecting honey adulteration. The study [14] applied Fourier transform infrared spectroscopy, along with multivariate analysis for adulteration detection in honey. Honey adulteration was carried out at various ratios up to 50%. Experimental findings indicate that spectral patterns change with various levels of adulteration, and hierarchical clustering shows groupings based on adulteration levels. Similarly, [15] investigates ultraviolet–visible spectral band to analyze changes in physicochemical parameters of honey for honey adulteration detection. Adulteration was carried out at various levels ranging from 5% to 50%. The study used spectral features to train an NN to discriminate between pure and adulterated honey. Coupled with principal component analysis (PCA), the artificial NN shows promising results for honey adulteration and reduced latency.

The focus of the study [16] is to investigate honey crystallization through spectral analysis. The authors gathered honey from different geographical sites and analyzed the crystals to find the source of honey. Experimental findings indicate that using wavelength and light intensity percentiles, honey samples from various cities of Turkey can be identified with high accuracy. Precise findings of the literature review are given in Table 1.

Recent advancements in food adulteration detection underscore the effectiveness of artificial intelligence (AI) and sensor technologies. Techniques such as Vis–NIRs coupled with SVMs and RFs have achieved higher accuracy in identifying adulterated honey. Furthermore, DL models, such as CNNs, have demonstrated superior performance in processing hyperspectral data by leveraging high-dimensional features [17]. The introduction of 3D CNNs has further enhanced hyperspectral analysis by incorporating spatial and spectral correlations, improving the detection of adulteration. The following key aspects are explored in this study:

- The study addresses practical challenges, especially focused on issues such as small-scale adulteration detection, variability in honey properties, and the need for noise reduction during spectral feature analysis.
- The study advanced the application of HSI and ML in honey adulteration detection, offering a scalable solution to enhance food quality assurance practices. It proposes an approach combining HSI and ML for binary and multiclass honey classification.
- Data are collected using a dedicated HSI setup. Various ratios of adulterated dwarf honey are mixed with giant honey to formulate five classes. A systematic data preparation pipeline is designed including spectral feature extraction, noise reduction, normalization, and feature scaling to optimize the performance of ML models.
- Extensive experiments involve ML and DL models like multilayer perceptron (MLP), extra tree (ET) classifier, logistic regression (LR), RF, SVM, DT, CNN, gated recurrent unit (GRU), long short-term memory (LSTM), and recurrent neural network (RNN).

Further, Section 2 explains the discussion of the proposed approach, dataset, ML models, and evaluation metrics. Results are discussed in Section 3, while Section 4 concludes this study.

2 | Materials and Methods

Pure honey is a natural, unadulterated product produced by bees and contains sugar in the form of fructose, glucose, and sucrose, along with water, enzymes, vitamins, and minerals with high viscosity, consistent natural aroma, and distinct spectral signatures when analyzed through HSI. In contrast, impure honey is adulterated with external substances such as sugar syrups, sucrose solutions, or water volume or alters its properties. The impure honey is produced by blending the determined ratio of dwarf honey with calculated amounts of sugar syrup.

In this study, two labels are used: pure "0" and impure "1". It is observed during the literature review that sugar syrup is mixed less than 30%, honey is considered pure, and above is impure. Low-level adulteration (5%–15%) shows minor spectral deviations; moderate adulteration (16%–30%) leads to noticeable changes in viscosity and spectral properties, while high-level adulteration (> 30%) significantly alters honey's natural composition, making it easier to detect. By leveraging HSI and ML models, this study captures the spectral variations at different impurity levels, identifies the substances added, and defines thresholds for classifying pure and impure honey with high accuracy. In addition, experiments with multiple classes are also conducted. For that, the impurity levels are systematically varied as 10%, 20%, 30%, and 50%, enabling the detection and classification of adulterated honey across a range of impurity levels.

This research focuses on identifying pure and impure honey by leveraging HSI and ML techniques. Honey samples were collected from the commercial market, including two distinct types, giant honey and dwarf honey, chosen for their diversity and the opportunity they provide to examine how granule size influences impurity absorption when mixed with sugar syrups. Adulterated honey samples were prepared by blending calculated amounts of sugar syrup, composed of sugar and water, maintained at a consistent temperature of 40°C. Reflectance data were collected using a Specim FX-10 hyperspectral camera operating in the visible and NIR spectra (395–1000 nm) with 224 spectral bands. Adulterated honey samples were prepared by mixing pure honey with food-grade sugar syrup at four predefined adulteration levels of 10%, 20%, 30%, and 50% (w/w). The mixtures were thoroughly homogenized before hyperspectral image acquisition to ensure uniform distribution of the adulterant. Glucose was selected as the adulterant because it is widely used as a commercial adulterant in honey authenticity studies. Its composition is similar to natural honey, making it difficult to detect honey adulteration using traditional methods. The original giant honey

TABLE 1 | Main finding of the literature review.

Ref.	Methodology	Results	Limitations
9	Makes use of models relying on chaotic parameters like shifting (lag-*k*) autocorrelation coefficients	Accuracy higher than 90% in examining samples from various botanical origins and higher than 95% when examining honey types	Difficult to detect small amounts of adulteration.
10	IRS, Raman spectroscopy (RS), and HS to detect honey adulteration	Detect adulteration based on the spectral responses of different components in the sample	Being singe-point measurement methods, IRS and RS may not reflect the real adulteration due to the uneven distribution of adulterated elements
11	Makes use of chemometric approaches MIR for quick and better detection of honey adulteration.	MIR spectroscopy joined with appropriate chemometric can provide accurate results	Interpreting large spectral data is difficult due to overlapping bands of NIR region
12	Demonstrate the capability of HSI in detecting honey adulteration	Accuracy varies with respect to parameters and conditions	Requirement of complex setup and specialized equipment leading to high cost
13	Combining ML with Vis–NIRs.	SVM and RF achieved the best accuracy, while SVR has the best RoF 0.991 and RMSE of 1.894	The impact of honey processing and smaller adulteration levels are not considered.
14	Feature selection integrated with visible near-infrared spectroscopy	Improved 96.67% accuracy for honey adulteration	Accuracy depends on preprocessing, no cross-spectrometer validation, and smaller dataset.
15	Fourier transform infrared (FTIR) spectroscopy	Hierarchical cluster analysis indicates clusters at various adulteration levels	FTIR alone is not enough to report accurate purity; chemometrics are needed.
16	Ultraviolet–visible spectroscopy with PCA	Rapid processing and promising results for various adulteration levels	Poor performance for honey containing 10% or less adulteration.

had a 14.2 ± 0.57 moisture level (%), while the dwarf honey had a 13.8 ± 0.28 moisture (%). On the other hand, the total sugar for the giant and dwarf honey was approximately 64 and 72, respectively.

Noise removal and preprocessing were performed using various filters, including median, steady, moving average, and Savitzky–Golay filters, with the latter providing the most effective noise reduction and spectral enhancement. The preprocessed data were formatted into CSV files and split into training (75%) and testing (25%) subsets. ML algorithms, including MLP, RF, ET, KNN, DT, LR, and SVM, were trained on the spectral data to classify honey samples as pure or adulterated. Among these models, the SVM demonstrated the highest accuracy of 99.99%, establishing it as the most effective method for honey adulteration detection. The methodology and performance evaluation are illustrated in Figure 1.

2.1 | Dataset Collection and Acquisition

A hyperspectral camera, Specim FX-10 (Specim, Spectral Imaging Ltd, Oulu, Finland), was used to acquire the samples, as shown in Figure 2. This camera captures the visible and near-infrared (VNIR) electromagnetic spectrum, covering 224 spectral bands with a spectral resolution of 2.7 nm, across a range of 395–1000 nm. It is equipped with a Schneider Cinegon 1.4/8 mm lens for enhanced image clarity. The laboratory setup included a translational platform measuring 21 cm by 40 cm and three halogen lights to ensure uniform illumination. The camera system was controlled via a computer using GigE-Vision and a serial connection adapter. To reduce external noise and ensure accurate data capture, the HSI system was enclosed in a dark-colored box, with the camera positioned 30 cm above the moving tray.

2.2 | Output of Hyperspectral Processing System and Sample Layout

Hyperspectral cameras capture a large amount of detailed spectral information, as shown in Figure 3, which makes the dataset quite large and requires careful planning for storage and access. The dataset includes white references, dark references, and files in HDR and RAW formats, all increasing its quality and flexibility. White reference files are taken when the camera captures light from a uniformly lit white surface. These are used to correct differences in lighting and sensor sensitivity across different wavelengths. Dark references, on the other hand, are recorded without any light to measure the camera's built-in noise. HDR files help capture subtle differences in light intensity, while RAW files store unprocessed, original data directly from the camera. These RAW files contain pure sensor readings, without any changes or corrections applied.

The reflectance values from these files act like unique fingerprints, helping to thoroughly identify the composition of honey samples. By creating graphs for both pure and impure honey, it is possible to visually compare their reflectance patterns and spot clear differences between them, as shown in Figure 4.

Studying hyperspectral data of honey is a powerful way to understand its complex makeup, which can benefit both the food industry and scientific research.

The sample output of hyperspectral header values stored in the comma-separated value (CSV) file is given in Figure 5. The CSV file is read using the Panda library available in Python language. A total of 225 different values are acquired from our samples using the hyperspectral machine, with two labels pure represented by "0" and impure represented by "1" shown in the last column of the CSV file. A screenshot of the sample output is given in Figure 6.

For multiclass honey adulteration, the original data acquired by the hyperspectral machine have 500 samples with five labels 0–4; after preprocessing, these are transformed into only two labels 0 and 1. A short description is given in Table 2. The original data have 100 samples for each of the five classes. For ML models, the data are divided with a 70:30 ratio for training and test datasets. In the test dataset, 150 samples from five classes are selected to test models.

2.3 | ML Models

Through the use of ML, automated systems can be developed for automated honey adulteration detection. Making predictions using ML models requires training the models on training data to learn relationships among various attributes and the target class.

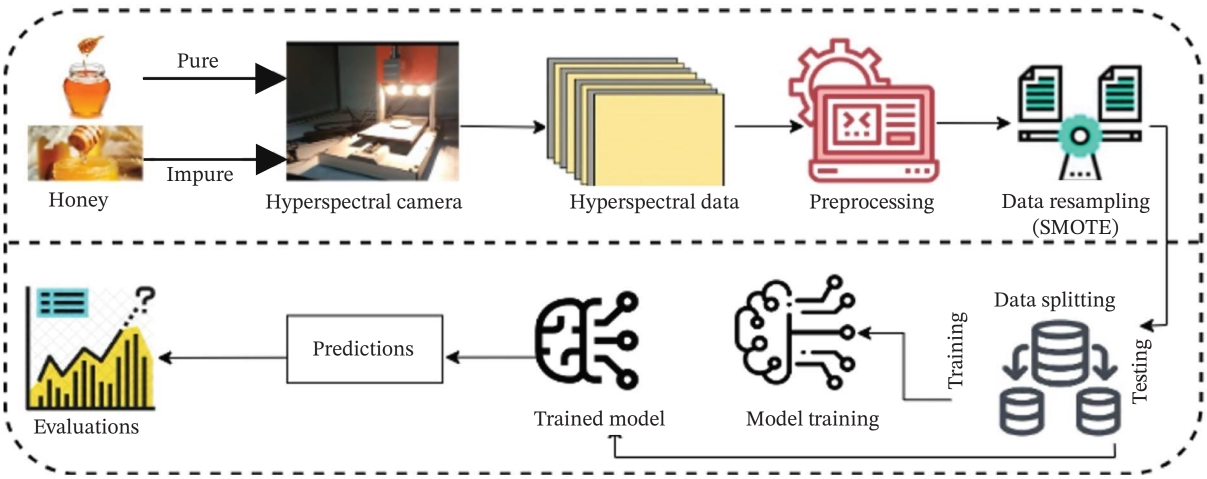

FIGURE 1 | Methodology diagram for classification of pure and impure honey.

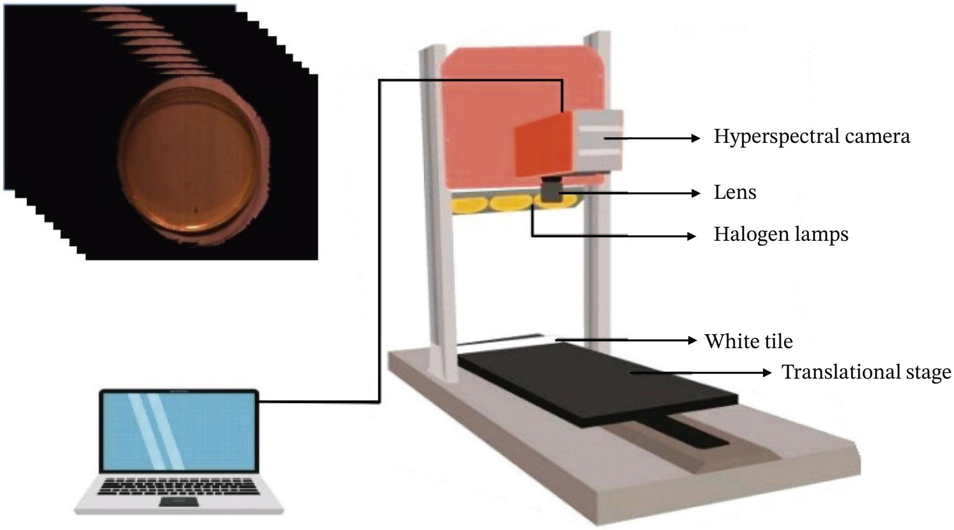

FIGURE 2 | Components of hyperspectral processing system.

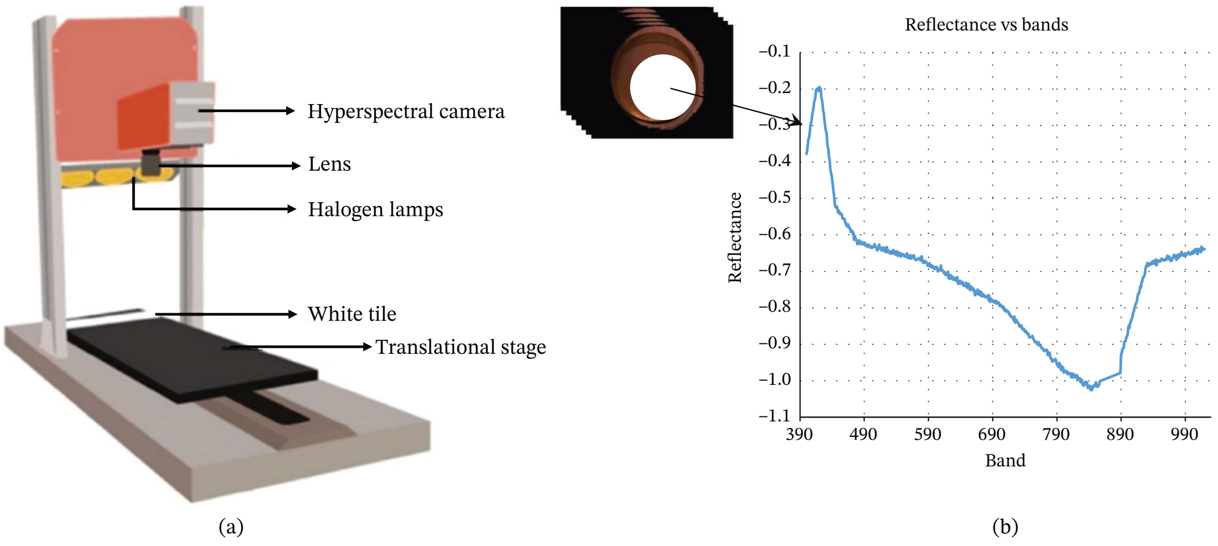

FIGURE 3 | Output from hyperspectral system. (a) Experimental setup and (b) output in the form of reflectance.

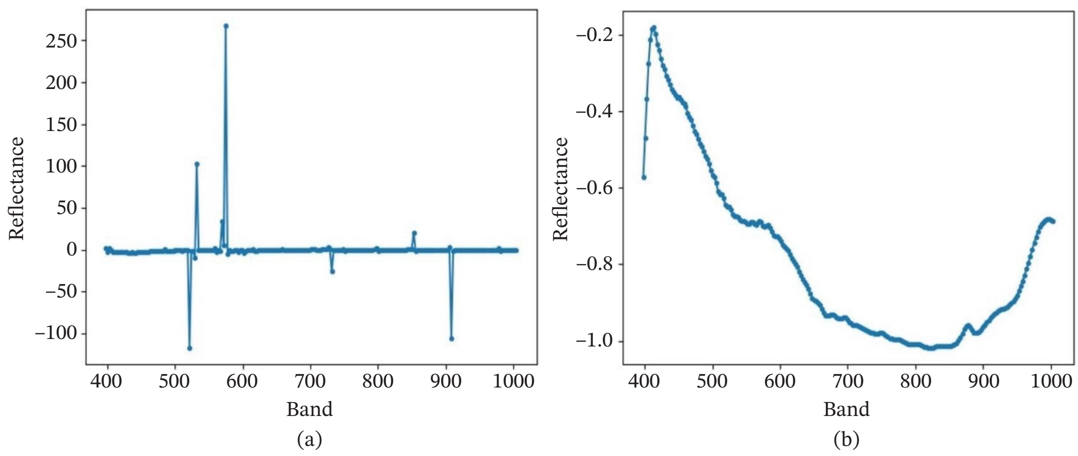

FIGURE 4 | Reflectance graph for (a) pure honey and (b) impure honey.

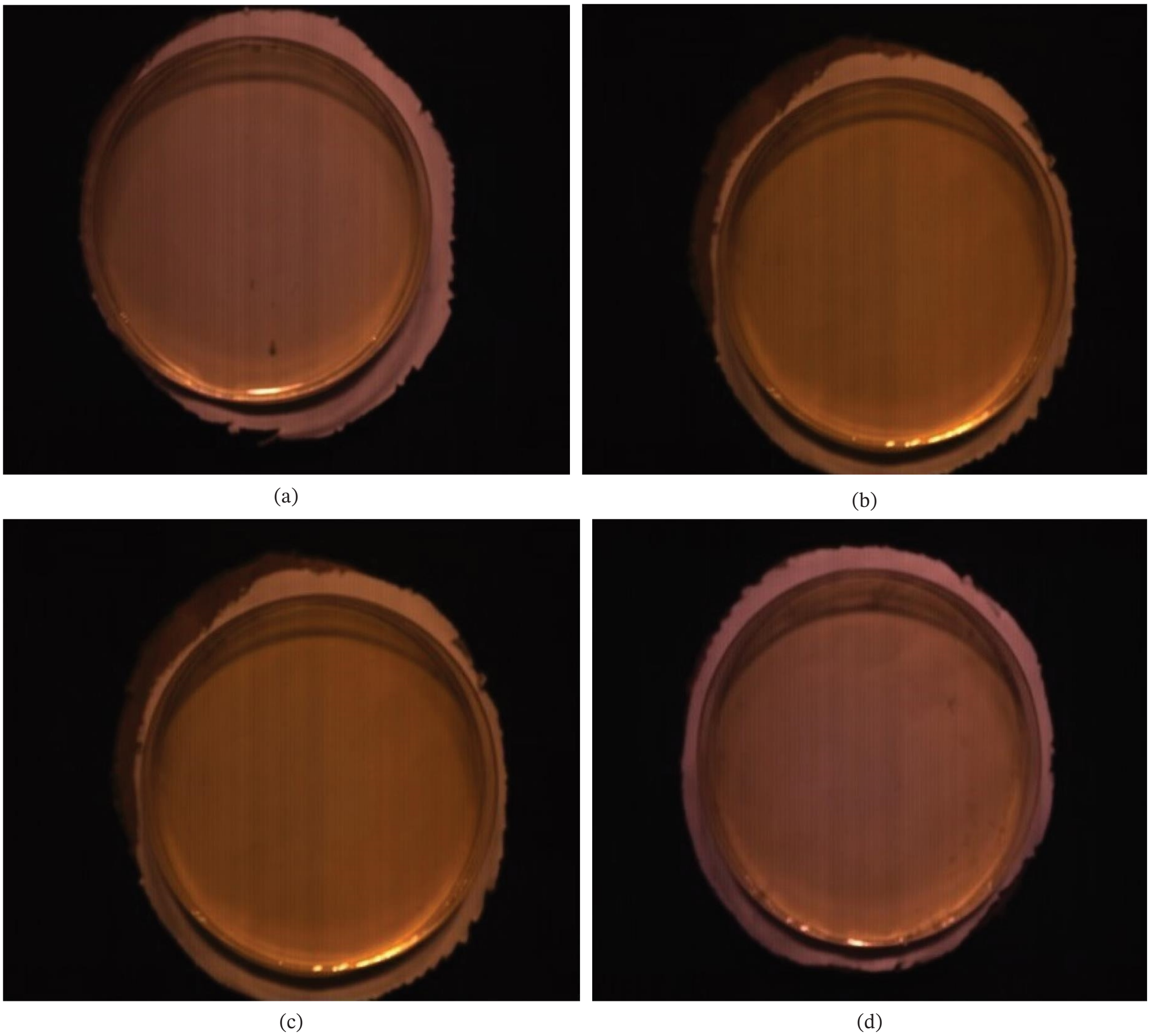

FIGURE 5 | Visual difference of the pure honey and impure honey classes: (a, b) pure honey samples and (c, d) impure honey samples.

	397.66	400.28	402.9	405.52	408.13	410.75	413.37	416	418.62	421.24	...	981.27	984.09	986.9	989.72	992.54	995.35	998.17	1000.99	1003.81	label
0	−0.581351	−0.495557	−0.413323	−0.343727	−0.298280	−0.287901	−0.295343	−0.333342	−0.378906	−0.415864	...	−0.799948	−0.791234	−0.785103	−0.780980	−0.776456	−0.772721	−0.772541	−0.773914	−0.772568	0
1	−0.663985	−0.572449	−0.484872	−0.414582	−0.365079	−0.352675	−0.362365	−0.397921	−0.447127	−0.486312	...	−0.897837	−0.889733	−0.883007	−0.878377	−0.874862	−0.870700	−0.869888	−0.871897	−0.870954	0
2	−0.571108	−0.470835	−0.370555	−0.279631	−0.211473	−0.176340	−0.154849	−0.160764	−0.172006	−0.182383	...	−0.738265	−0.727234	−0.718625	−0.713062	−0.708621	−0.705445	−0.705624	−0.707049	−0.708551	0
3	−0.571108	−0.470835	−0.370555	−0.279631	−0.211473	−0.176340	−0.154849	−0.160764	−0.172006	−0.182383	...	−0.738265	−0.727234	−0.718625	−0.713062	−0.708621	−0.705445	−0.705624	−0.707049	−0.708551	0
4	−0.518284	−0.430554	−0.331294	−0.241095	−0.175986	−0.142117	−0.122450	−0.127154	−0.136935	−0.142445	...	−0.672318	−0.660888	−0.652287	−0.646634	−0.641780	−0.637974	−0.637246	−0.639655	−0.641794	0

5 rows x 225 columns

FIGURE 6 | Sample output data in numeric form from the hyperspectral imaging setup used for training and testing.

TABLE 2 | Details of the dataset for honey.

Label	Source	No. of samples
0	dwarf & giant	100
1	10% adulteration of dwarf & giant	100
2	20% adulteration of dwarf & giant	100
3	30% adulteration of dwarf & giant	100
4	50% adulteration of dwarf & giant	100

At its core, ML relies on data; with higher and better data, it can provide better results.

HSI offers extensive information on the chemical composition of honey by recording a wide range of spectral bands, making it possible to detect even minute changes. Some of the advanced classification algorithms (MLP, RF, ETC, kNN) are used to understand the spectrum data and differentiate between pure honey and adulterants, allowing for quick and precise identification.

This technology promises to protect customer confidence and health by guaranteeing the quality and authenticity of honey used in a variety of sectors, such as food, pharmaceuticals, and cosmetics.

2.3.1 | MLP

The MLP, a type of artificial NN, was chosen for its ability to capture complex patterns and nonlinear relationships within the data. This model consists of multiple layers of interconnected neurons, where each layer transforms the input data through weighted connections and activation functions. By leveraging the MLP's DL capabilities, we aimed to accurately classify honey samples based on their hyperspectral signatures. The model was trained using a diverse dataset of hyperspectral images, enabling it to learn the subtle spectral differences between pure and adulterated honey. Performance metrics such as accuracy, precision, and recall were used to evaluate the MLP's effectiveness, demonstrating its potential as a robust tool for the reliable detection of honey adulteration.

2.3.2 | Random Forest

RF model, an ensemble learning technique, aggregates the predictions of multiple decision trees to improve classification accuracy and robustness. Each tree in the forest is trained using a subset of the data and features selected randomly. This randomness is used in understanding diverse patterns and reducing overfitting. By applying this approach to hyperspectral images, the RF model effectively distinguishes between pure and adulterated honey by analyzing the complex spectral information present in the data. The model's strength lies in its ability to handle high-dimensional datasets and its resilience to noisy data, making it well suited for the task of honey adulteration detection.

2.3.3 | ET Classifier

The ET classifier is an ensemble learning technique that builds a multitude of decision trees during training and aggregates their predictions to produce a final classification. Unlike traditional decision trees, the ET Classifier introduces greater randomness in the training process by selecting random subsets of features and splitting thresholds at each node. This increased randomness helps in reducing variance and improving the generalization of the model. For our application, the ET Classifier effectively leveraged the high-dimensional spectral data from hyperspectral images to differentiate between pure and adulterated honey.

2.3.4 | *K*-Nearest Neighbors

The KNN model is a nonparametric, instance-based learning technique that classifies data based on the similarity of neighboring samples. By calculating the distance between a query sample and its nearest neighbors in the feature space, KNN assigns a label based on the majority vote of these nearest neighbors. This method is particularly advantageous for hyperspectral data, as it effectively handles high-dimensional spaces and does not require an explicit model training phase. For THE study, KNN was employed to distinguish between pure and adulterated honey by analyzing the spectral characteristics captured in hyperspectral images. The choice of the number of neighbors (k) and the distance metric were optimized to improve classification accuracy. Evaluation results, including accuracy and classification metrics, demonstrated that KNN provided a straightforward yet effective approach for identifying adulteration, showcasing its practicality in real-world applications for honey quality assurance.

3 | Results and Discussion

This section presents the performance evaluation of the ML models applied to the honey adulteration dataset. The results are analyzed across two key state-of-the-art ML models and four DL models CNN, LSTM, RNN, and GRU models.

3.1 | Experimental Setup and Hyperparameters

The experiments are carried out using the Python programming language. Data were acquired using a dedicated setup involving hyperspectral camera and various compositions of dwarf and giant honey. For implementing traditional ML models, the scikit-learn library was used. On the other hand, DL models were implemented using TensorFlow/Keras. Experiments were performed using an Intel Core i7 system, operating on Windows 10 with 16 GB RAM. Descriptive statistics of the dataset used for binary class classification with and without applying the synthetic minority oversampling technique (SMOTE) for oversampling in the experimental setup are given in Table 3.

For optimal performance, hyperparameter selection is very important. This study used the GridSearchCV method for hyperparameter selection, which performs an exhaustive search over a predefined set of parameters to identify the optimal configuration. The range of values was defined using the existing literature on adulterated honey detection. GridSearchCV then trained the model using all possible combinations of these parameters, employing k-fold cross-validation to ensure robustness and reduce overfitting. The best set of parameters was decided based on the best average cross-validation performance, which was selected as the final configuration and used for model training and testing. Basic parameters are given in Table 4 for each ML model. The models are optimized to get the best results for honey classification.

3.2 | Binary Classification Without SMOTE

In the first step, ML models are implemented after the scaling using the StandardScaler () function available in Python language. Table 5 shows the results of models for binary

TABLE 3 | Training and test datasets of each class with balanced and imbalanced data.

Class label	Training dataset	Test dataset	Total	Original/SMOTE
0	70	30	100	Original data. Before applying SMOTE, the imbalanced data
1	280	120	400	
Total	350	150	500	
0	279	121	400	After applying SMOTE balanced data
1	281	119	400	
Total	560	240	800	

classification. After the careful analysis of confusion matrices of seven ML models, the performance of classification for pure and impure honey was done. It is observed that SVM and MLP demonstrate the highest accuracy, with minimal misclassifications, making them the best performers. LR and DTs also perform well, with only a few errors in their predictions. RF and ET show moderate performance, struggling slightly with Class 0 predictions. KNN has the weakest performance, misclassifying a notable number of samples, particularly for Class 0. Overall, SVM stands out as the best model due to its near-perfect classification accuracy.

Results for F1 score, accuracy, etc., are given in Table 6, indicating that SVM obtains a 0.99 accuracy score. It is followed by the MLP and LR with accuracy scores of 0.987 and 0.98, respectively. Other evaluation metrics for these models are also better, indicating the good fit of these models on the data.

Figure 7 visualizes the performance of the models on binary classes. It shows that despite a good performance by the SVM, MLP, and LR models, certain models show poor performance such as KNN and DT.

3.3 | Binary Classification Using SMOTE

In the case of binary classification, the pure class has less number of samples, which might lead to model overfitting on the majority classes. Often, with lower positive samples, the number of false positives is high. To avoid this problem, SMOTE is applied to the pure honey class to balance the dataset. For binary classification, a total of 240 samples are used for testing. It is important to point out that SMOTE is only applied to training samples to avoid data leakage. SMOTE is applied on the training samples, and after oversampling, the number of training samples are 560 (each class has 280 samples). All experiments for binary classification involve the same number of samples after applying SMOTE. The confusion matrix after applying the SMOTE is given in Table 7.

It is clearly observed that the confusion matrices reveal the impact of applying the SMOTE on the classification models. Before SMOTE, the models showed variability in handling imbalanced data, with SVM and MLP performing the best and KNN struggling the most. After applying SMOTE, the overall performance improved significantly, especially for minority class predictions.

Results given in Table 8 show that SVM and MLP maintained near-perfect accuracy, while RF, ET, and DT showed notable improvements, particularly reducing false negatives. KNN, despite some residual errors, demonstrated better balance in class predictions. LR also benefited from SMOTE, achieving higher accuracy with minimal misclassifications.

Figure 8 visualizes the models' performance on data balanced using SMOTE. Overall, SMOTE enhanced the predictive power of all models, making them more robust for imbalanced datasets.

3.4 | Binary Classification Using DL Models With SMOTE

Besides using ML models, four DL models are also used for experiments in this study. CNN, RNN, GRU, and LSTM models are used to classify samples into pure and impure honey. Table 9 shows the architecture and parameter details for DL models.

Since using ML models with SMOTE yielded the best results, DL models are used on the SMOTE-based data. Table 10 provides the results of DL models. CNN achieves the highest accuracy, showing consistent results for other evaluation metrics. In contrast, the RNN model demonstrates significant fluctuations concerning evaluation metrics like precision and F1 score.

The model struggles to converge, with both accuracies remaining relatively low compared to other models. The GRU model displays a more stable performance than RNN but achieves lower performance compared to CNN and LSTM. The LSTM model

TABLE 4 | Parameter settings used for the models.

Model name	Parameter settings
MLP	hidden_layer_sizes= (64), max_iter = 500, random_state = 42
Extra trees	n_estimators = 100, random_state = 42
Logistic regression	C = 1.0, penalty = 'l2', solver = 'lbfgs'
Random forest	n_estimators = 100, max_depth = None, random_state = 42
Support vector machine	C = 1.0, kernel = 'rbf', gamma = 'scale'
Decision tree	criterion = 'gini', max_depth = None, random_state = 42

TABLE 5 | Confusion matrix for models without using SMOTE.

Model	True positive	True negative	False positive	False negative
MLP	28	120	2	0
RF	23	120	7	0
ET	23	120	7	0
KNN	17	119	13	1
DT	22	119	8	1
LR	27	120	3	0
SVM	29	120	1	0

TABLE 6 | Performance of models without SMOTE.

Model	Accuracy	Precision	Recall	F1
MLP	0.987	0.984	0.98	0.982
RF	0.95	0.95	0.95	0.95
ETC	0.95	0.95	0.95	0.95
KNN	0.90	0.91	0.90	0.90
DT	0.94	0.94	0.94	0.93
SVM	0.99	0.99	0.99	0.99
LR	0.98	0.98	0.98	0.98

consistently achieves high performance for accuracy, recall, and precision, second only to CNN. In summary, CNN outperforms all other models, achieving the best accuracy score of 0.906, while LSTM is the second-best performer. The dataset used for training is smaller, and DL models do not get a good fit on the data as they need larger datasets for better training. As a result, their performance is lower compared to ML models.

3.5 | Multiclass Classification Using ML Models With SMOTE

In addition to binary classification, the multiclass problem is also investigated where the honey is divided into five classes. Class 0 shows pure honey, while classes 1–4 show different levels of impurity. The data are split into training and test splits, and the SMOTE is only applied to training data to avoid data leakage. Models are tested on the original test data without any oversampling. Furthermore, the SMOTE was applied using two oversampling levels. In SMOTE $\times 2$, the original 70 training samples per class were synthetically expanded to 140 samples per class, whereas in SMOTE $\times 3$, they were expanded to 210 samples per class. Figure 9 shows the confusion matrices of the best ML and DL models among all original and various settings of the SMOTE dataset used for experiments. Confusion matrix clearly shows that MLP achieved a 96.67% accuracy, while among DL models CNN achieved a 94.67% highest accuracy.

Table 11 presents the class-wise performance of seven ML and four DL models under three different training strategies: the original training dataset (70 samples per class), SMOTE $\times 2$ (140 samples per class), and SMOTE $\times 3$ (210 samples per class). Each model was evaluated on a fixed test set of 150 samples (30 samples per class). For each class, the numbers of correctly predicted (CP) and wrongly predicted (WP) samples are reported along with the overall classification performance.

Using the original training dataset, the MLP classifier achieved the highest performance among all ML models by correctly classifying 145 out of 150 samples (96.67%), followed by CNN, which correctly classified 138 samples (92.00%) and outperformed the remaining DL models. Ensemble models such as ET (136 correct predictions) and RF (130 correct predictions) also demonstrated competitive performance, whereas SVM (111 correct predictions) and LR (118 correct predictions) showed comparatively lower performance for the five-class classification task. Among the DL models, LSTM, RNN, and GRU produced significantly lower classification performance, indicating difficulty in learning discriminative spectral features from the available training samples.

After applying SMOTE $\times 2$, which increased the training samples from 70 to 140 per class, several models exhibited improved

TABLE 7 | Confusion matrix for models using SMOTE.

Model	True positive	True negative	False positive	False negative
MLP	121	118	0	1
RF	121	110	0	9
ETC	121	113	0	6
KNN	120	106	1	13
DT	120	110	1	9
LR	121	119	0	0
SVM	121	117	0	2

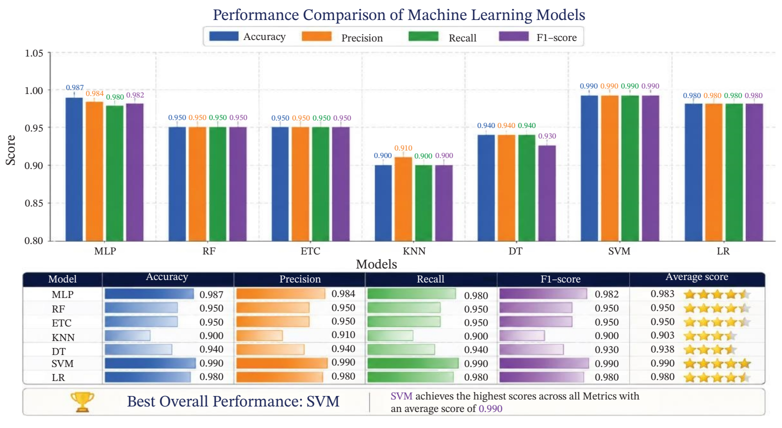

FIGURE 7 | Performance of models for binary classes on imbalanced datasets.

TABLE 8 | Performance of models using the SMOTE approach.

Model	Accuracy	Precision	Recall	F1
MLP	0.99	0.99	0.99	0.99
RF	0.96	0.96	0.96	0.96
ETC	0.975	0.976	0.975	0.975
KNN	0.94	0.946	0.941	0.941
DT	0.958	0.958	0.958	0.958
LR	1.00	1.00	1.00	1.00
SVM	0.99	0.99	0.99	0.99

classification performance. CNN improved from 138 to 140 correct predictions (93.33%), while ET improved from 136 to 138 correct predictions. Similarly, RF increased from 130 to 133, and KNN increased from 122 to 127 correctly classified samples. In contrast, the MLP performance remained unchanged at 145 correct predictions (96.67%), indicating that the classifier had already achieved near-optimal performance using the original balanced dataset. Although LSTM, RNN, and GRU also benefited from oversampling, their performance remained substantially lower than that of the ML classifiers.

With SMOTE $\times 3$, where the training samples were further increased from 70 to 210 per class, the CNN achieved its highest performance by correctly classifying 142 out of 150 samples (94.67%), reducing the number of misclassified samples from 12 to 8. This demonstrates that CNN benefited from the larger synthetically balanced training dataset. Conversely, the MLP exhibited a slight reduction from 145 to 144 correct predictions (96.00%), although it remained the overall best-performing model. The performances of RF, DT, and KNN also improved slightly compared with the original dataset, whereas LSTM, RNN, and GRU showed only marginal improvements despite the larger training set. A graphical representation of this study is given in Figure 10.

3.6 | Cross-Validation

The proposed approach is tested for 10-fold cross-validation to evaluate its robustness and ability to generalize. The results of cross-validation are given in Table 12. Results suggest that the MLP and ET models perform well using 10-folds similar to their better performance for multiclass problems. Besides accuracy, other metrics for these models also show better results compared to other models. On the other hand, SVM and LR show poor results.

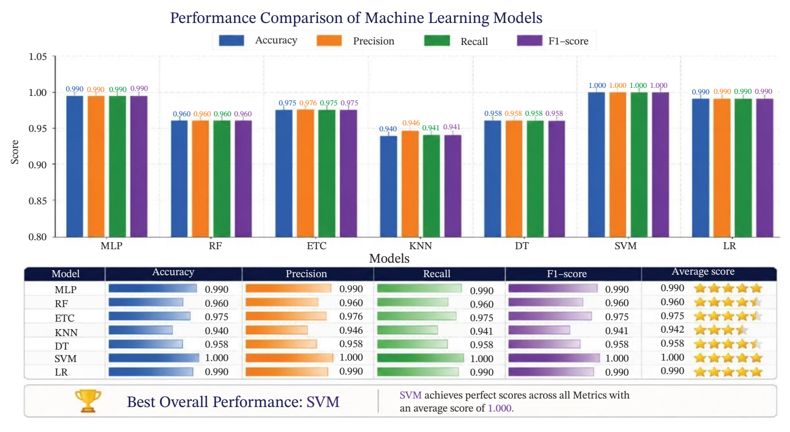

FIGURE 8 | Performance of models for binary classes using SMOTE.

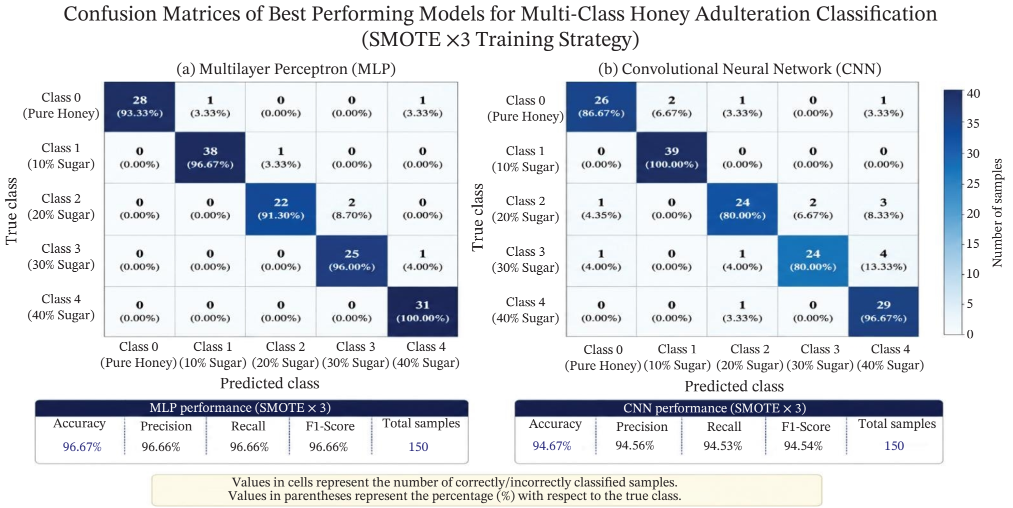

FIGURE 9 | Confusion matrices of the best-performing models for multiclass honey adulteration classification with SMOTE $\times$ 3 training strategy. (a) Multilayer perceptron (MLP). (b) Convolutional neural network (CNN).

TABLE 9 | Parameter settings of DL models with batch size 32 and epoch 20 in each case.

Model	Layers	Activation	Dropout	Optimizer	Loss function
CNN	Conv1D (64 filters, kernel size = 3) MaxPooling1D (pool size = 2) Dropout (0.2) Flat-ten, Dense (64 units), Dropout (0.5), Dense (softmax)	ReLU (Conv1D, dense) Softmax (output layer)	0.2, 0.5	Adam	Sparse categorical cross-entropy
RNN	RNN (64 units) Dropout (0.2) Dense (64 units) Dropout (0.5) Dense (Softmax)	ReLU (dense) Softmax (output layer)	0.2, 0.5	Adam	Sparse categorical cross-entropy
GRU	GRU (64 units) Dropout (0.2) Dense (64 units) Dropout (0.5) Dense (Softmax)	ReLU (dense) Softmax (output layer)	0.2, 0.5	Adam	Sparse categorical cross-entropy
LSTM	LSTM (64 units) Dropout (0.2) Dense (64 units) Dropout (0.5) Dense (Softmax)	ReLU (dense) Softmax (output layer)	0.2, 0.5	Adam	Sparse categorical cross-entropy

TABLE 10 | Performance evaluation of DL models using SMOTE.

Model	Accuracy	Precision	Recall	F1
GRU	0.84	0.83	0.84	0.81
LSTM	0.90	0.90	0.90	0.89
CNN	0.906	0.92	0.91	0.90
RNN	0.80	0.64	0.80	0.71

3.7 | Performance Comparison

Comparative performance analysis of the current study with existing approaches on adulterated honey detection is provided in Table 13. The study [4] used SVM, LDA, and NNs for honey adulterated detection and reported a 93% accuracy using the SVM model. The study [6], on the other hand, adopted several models like ANN, SVM, KNN, RF regressor (RFR), and DT, with the ANN obtaining the best accuracy of 97.9% for the binary class problem. The authors report a 95.0% accuracy using stepwise discriminant analysis (SDA) with a leave-one-out-cross-validation approach. In comparison to these approaches, the current study showed a better accuracy of 96.67% using five classes of honey.

3.8 | Discussion

The results underscore the effectiveness of HSI combined with ML and DL for robust and accurate detection of honey adulteration. The MLP model achieves perfect metrics across all performance measures, emerging as the most reliable solution for classifying pure and adulterated honey.

The MLP model outperforms RF, KNN, DT, and LR because it can effectively learn highly nonlinear, hierarchical feature representations from data. Compared to LR, MLP model is not limited to linear decision boundaries. While DT and RF rely on hard splits that may fragment the data, MLP model captures smooth and complex interactions without the need for such splits. The performance of the KNN model is reported to degrade for high-dimensional datasets, which is the case in this study as well. The MLP, on the other hand, generalizes well through learned weight sharing and regularization. Provided with a sufficient data and a proper tuning, MLP can adapt well and show

TABLE 11 | Classification results (correct predictions [CPs] and wrong predictions [WPs]) across experiments and ML models.

Experiment	ML model	Class 0		Class 1		Class 2		Class 3		Class 4		Total	
		CP	WP	CP	WP	CP	WP	CP	WP	CP	WP	CP	WP
Original data vs: test 70:30	MLP	28	2	38	2	22	1	25	0	32	0	145	5
	RF	20	10	39	1	22	1	20	5	29	3	130	20
	ET	22	8	38	2	23	0	22	3	31	1	136	14
	KNN	18	12	37	3	20	3	21	4	26	6	122	28
	SVM	15	15	38	2	18	5	20	5	20	12	111	39
	LR	28	2	37	3	17	6	23	2	13	19	118	32
	DT	22	8	38	2	22	1	22	3	27	5	131	19
	CNN	26	4	38	2	21	2	21	4	32	0	138	12
	LSTM	0	30	39	1	0	23	10	15	0	32	49	101
	RNN	23	7	0	40	10	13	0	25	8	24	41	109
	GRU	23	7	0	40	12	11	0	25	0	32	35	115
SMOTE × 2 (140/class) 140:30	MLP	28	2	38	2	22	1	25	0	32	0	145	5
	RF	24	6	38	2	21	2	19	6	31	1	133	17
	ET	25	5	38	2	23	0	21	4	31	1	138	12
	KNN	19	11	37	3	20	3	20	5	31	1	127	23
	SVM	16	14	38	2	20	3	19	6	24	8	117	33
	LR	27	3	38	2	19	4	21	4	17	15	122	28
	DT	22	8	38	2	20	3	21	4	28	4	129	21
	CNN	26	4	39	1	21	2	24	1	30	2	140	10
	LSTM	21	9	39	1	1	22	0	25	0	32	61	89
	RNN	20	10	38	2	4	19	0	25	16	16	78	72
	GRU	23	7	33	7	3	20	0	25	0	32	59	91
SMOTE × 3 (210/class) 210:1	MLP	28	2	38	2	22	1	25	0	31	1	144	6
	RF	23	7	38	2	21	2	20	5	29	3	131	19
	ET	25	5	38	2	23	0	21	4	30	2	137	13
	KNN	21	9	37	3	22	1	20	5	31	1	131	19
	SVM	16	14	38	2	20	3	19	6	28	4	121	29
	LR	28	2	38	2	18	5	23	2	18	14	125	25
	DT	23	7	37	3	21	2	22	3	28	4	131	19
	CNN	26	4	39	1	21	2	24	1	32	0	142	8
	LSTM	23	7	39	1	0	23	0	25	0	32	62	88
	RNN	26	4	39	1	0	23	0	25	0	32	65	85
	GRU	23	7	39	1	0	23	0	25	0	32	62	88

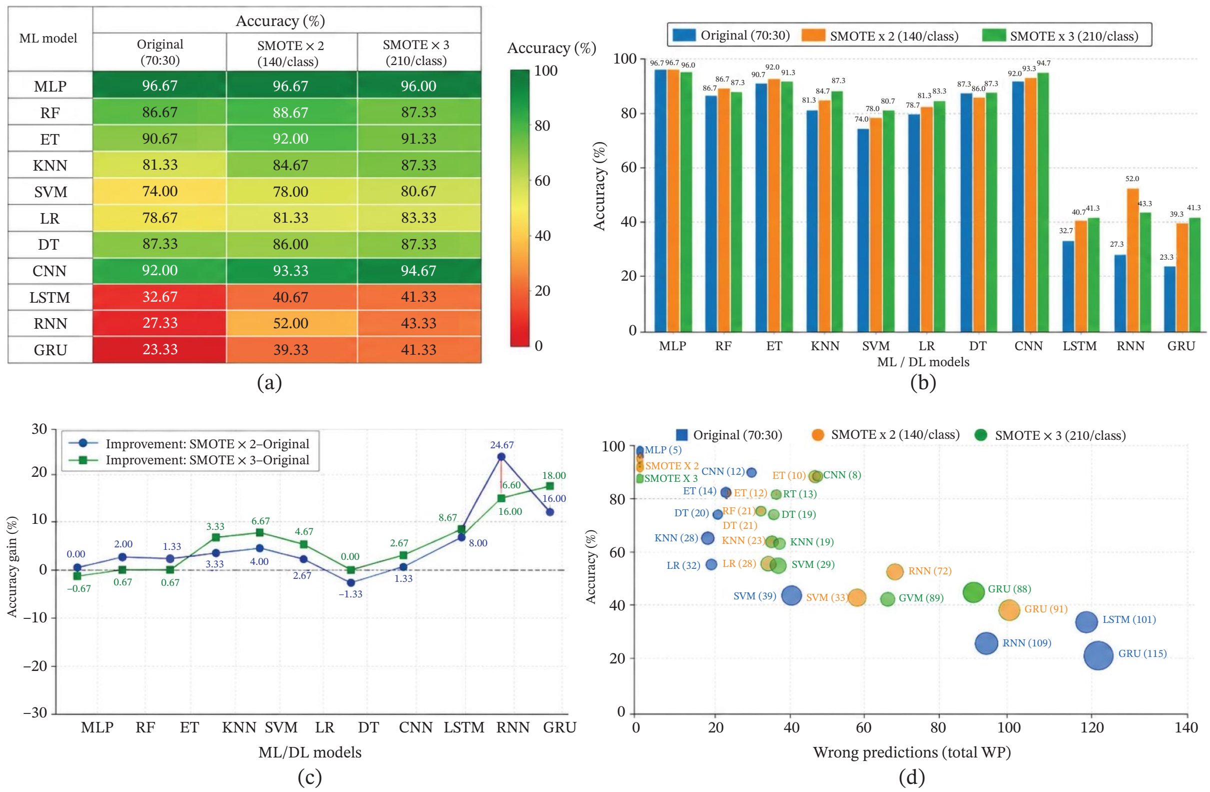

FIGURE 10 | Performance analysis under different training strategies: (a) accuracy (%) heatmap across models and sampling strategies, (b) overall accuracy (%) comparison across models, (c) accuracy improvement over original data, and (d) accuracy vs. wrong predictions (bubble size shows wrong predictions).

better results for complex, noisy, and high-dimensional datasets common in real-world applications. Its performance surpasses traditional methods, which often struggle to detect low levels of adulteration. Additionally, the ET and SVM models deliver near-perfect results, further validating the potential of ML techniques in this domain. SVM performs better because it handles high-dimensional hyperspectral data effectively, finding optimal margins that separate subtle spectral differences. Unlike trees and KNN, SVM resists overfitting, and unlike logistic regression, it captures nonlinear class boundaries, making it well suited for detecting fine adulteration patterns in honey spectra.

These findings align with recent studies, such as [12], which leveraged Vis–NIRs and ML techniques to achieve similar classification accuracies. Furthermore, the existing literature highlighted the advantages of DL models, such as CNN and LSTM, in extracting high-dimensional features from hyperspectral data, which aligns with the performance of CNN in this study. However, LSTM, RNN, and GRU exhibited comparatively lower classification performance, particularly for the intermediate adulteration levels (Classes 2 and 3), indicating difficulty in learning discriminative spectral patterns from the relatively limited training dataset. Nevertheless, CNN demonstrated superior robustness in capturing complex hyperspectral features, achieving the highest DL accuracy of 94.67%. These results suggest that the performance of recurrent DL models may further improve with the availability of larger and more diverse training datasets. Although the binary classification models achieved very high classification accuracies, these results should be interpreted carefully. The high performance is primarily attributed to the rich spectral information provided by HSI, which enables clear discrimination between pure and adulterated honey samples. To minimize data leakage, the dataset was first divided into independent training and testing sets, and the SMOTE was applied only to the training data. Consequently, the testing samples remained completely unseen during model

TABLE 12 | Results of cross-validation using 10-folds.

Model	Accuracy	Precision	Recall	F1 score
MLP	93.20	93.72	93.20	93.22
RF	91.20	91.82	91.20	91.12
ET	93.00	93.50	93.00	92.99
KNN	89.60	90.53	89.60	89.56
DT	89.20	89.98	89.20	89.17
SVM	69.40	69.73	69.40	68.31
LR	64.40	63.15	64.40	62.73

TABLE 13 | Performance analysis in comparison to existing approaches.

Ref	Model	Classes	Accuracy
5	SVM, LDA, NN	3	93% using SVM
7	ANN, SVM, KNN, RFR, DT	2	0.979 with ANN
12	SDA	3	95%
Proposed	MLP	5	96.67%

training. Nevertheless, because the dataset is relatively small, there remains a possibility that the reported performance may be optimistic. Furthermore, hyperspectral data often exhibit strong spectral and spatial correlations that can increase classification performance if samples are not completely independent. Although every effort was made to maintain an independent train–test split, future work will validate the proposed framework using larger datasets collected from different acquisition sessions, honey batches, and external sources to further assess the generalization capability and robustness of the developed models.

Despite achieving high classification accuracy, some limitations persist. The variability in honey's botanical origin and environmental factors may impact the spectral signatures, emphasizing the need for a more diverse dataset to ensure generalizability. Additionally, while both ML and DL models show promise, scaling these methodologies for real-time industrial applications would require addressing challenges such as computational efficiency and the variability in large-scale datasets.

Nevertheless, the integration of HSI with ML and DL techniques represents a significant advancement in food quality assurance. This approach enhances consumer trust and product authenticity, particularly in the honey industry, where adulteration remains a critical concern. The results affirm the potential of these methods in ensuring the detection of adulteration with exceptional precision and reliability.

4 | Conclusion

This study demonstrates a robust HSI-based framework for identifying adulterated honey by exploiting reflectance signatures that capture subtle impurity variations. The core technical contribution lies in systematically evaluating ML and DL models for both binary and fine-grained multiclass adulteration detection. SVM achieves a near-perfect accuracy (99.99%) in binary classification, while MLP delivers superior performance (96.67%) for five-level adulteration recognition, validated through k-fold cross-validation. Comparative analysis shows clear improvements over existing methods. Although CNN performance is promising, results indicate that larger hyperspectral datasets are essential to fully realize DL's potential in this domain.

In the future, we intend to expand the dataset to include honey samples from diverse botanical origins and geographic regions. In addition, optimizing computational models for real-time detection in industrial settings is another potential consideration. We also want to investigate advanced DL architectures, such as 3D-CNNs and Siamese neural networks, to enhance classification accuracy.

Acknowledgments

Declaration of Generative AI and AI-Assisted Technologies in the writing Process. This study used Grammarly only for typos and English grammar correction.

Funding

This work was supported by the European University of Atlantic.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. Shafiee, G. Polder, S. Minaei, N. Moghadam-Charkari, S. Van Ruth, and P. M. Kus, "Detection of Honey Adulteration Using Hyperspectral Imaging," *IFAC-PapersOnLine* 49, no. 16 (2016): 311–314, https://doi.org/10.1016/j.ifacol.2016.10.057.

2. T. Phillips and W. Abdulla, "A New Honey Adulteration Detection Approach Using Hyperspectral Imaging and Machine Learning," *European Food Research and Technology* 249, no. 2 (2023): 259–272, https://doi.org/10.1007/s00217-022-04113-9.

3. Y. Shao, Y. Shi, G. Xuan, et al., "Hyperspectral Imaging for Non-Destructive Detection of Honey Adulteration," *Vibrational Spectroscopy* 118 (2022): 103340, https://doi.org/10.1016/j.vibspec.2022.103340.

4. G. Machuca, J. Staforelli, M. Rondanelli-Reyes, et al., "Hyperspectral Microscopy Technology to Detect Syrups Adulteration of Endemic Guindo Santo and Quillay Honey Using Machine-Learning Tools," *Foods* 11, no. 23 (2022): 3868, https://doi.org/10.3390/foods11233868.

5. I. I. Robert, Z. Birech, and M. Kaniu, "Rapid Assessment of Molasses Adulterated Honey Using Laser Raman Spectroscopy and Principal Component Analysis," *Food Analytical Methods* 16, no. 11 (2023): 1702–1710, https://doi.org/10.1007/s12161-023-02538-w.

6. E. Ahmed, "Detection of Honey Adulteration Using Machine Learning," *PLOS Digital Health* 3, no. 6 (2024): e0000536, https://doi.org/10.1371/journal.pdig.0000536.

7. T. Phillips, "Botanical Origins Classification and Adulteration Detection in NZ Honey Using Machine Learning and Hyperspectral Imaging," *PhD thesis* (2022).

8. M. Lastra-Mejías, M. Izquierdo, E. González-Flores, J. C. Cancilla, J. G. Izquierdo, and J. S. Torrecilla, "Honey Exposed to Laser-Induced Breakdown Spectroscopy for Chaos-Based Botanical Classification and Fraud Assessment," *Chemometrics and Intelligent Laboratory Systems* 199 (2020): 103939, https://doi.org/10.1016/j.chemolab.2020.103939.

9. P. Li, S. Tang, S. Chen, X. Tian, and N. Zhong, "Hyperspectral Imaging Combined With Convolutional Neural Network for Accurately Detecting Adulteration in Atlantic Salmon," *Food Control* 147 (2023): 109573, https://doi.org/10.1016/j.foodcont.2022.109573.

10. E. Mendes and N. Duarte, "Mid-Infrared Spectroscopy as a Valuable Tool to Tackle Food Analysis: A Literature Review on Coffee, Dairies, Honey, Olive Oil and Wine," *Foods* 10, no. 2 (2021): 477, https://doi.org/10.3390/foods10020477.

11. D. B. Siano, W. Abdullakasim, A. Terdwongworakul, and K. Phuangsombut, "Improving the Performance of the Model Developed From the Classification of Adulterated Honey With Different Botanical Origins Based on Near-Infrared Hyperspectral Imaging System and Supervised Classification Algorithms," *Infrared Physics & Technology* 131 (2023): 104692, https://doi.org/10.1016/j.infrared.2023.104692.

12. J. L. P. Calle, I. Punta-Sánchez, G. Peredo, et al., "Rapid and Automated Method for Detecting and Quantifying Adulterations in High-Quality Honey Using Vis-NIRs in Combination With Machine Learning," *Foods* 12, no. 13 (2023): 2491.

13. M. Al-Awadhi and R. Deshmukh, "Enhancing Honey Adulteration Detection With Optimal Subspace Wavelength Reduction in Vis-NIR Reflection Spectroscopy," *IEEE Access* 11 (2023): 144226–144243, https://doi.org/10.1109/access.2023.3343731.

14. T. Damto, A. Zewdu, and T. Birhanu, "Application of Fourier Transform Infrared (FT-IR) Spectroscopy and Multivariate Analysis for

Detection of Adulteration in Honey Markets in Ethiopia," *Current Research in Food Science* 7 (2023): 100565, https://doi.org/10.1016/j.crfs.2023.100565.

15. E. I. Geana, R. Isopescu, C. T. Ciucure, C. L. Gîjiu, and A. M. Josceanu, "Honey Adulteration Detection via Ultraviolet-Visible Spectral Investigation Coupled With Chemometric Analysis," *Foods* 13, no. 22 (2024): 3630, https://doi.org/10.3390/foods13223630.

16. A. Beyaz and M. Akkul, "Honey Crystallization Analysis Based on Spectral Signature," *Journal of Food Process Engineering* 49, no. 2 (2026): e70341, https://doi.org/10.1111/jfpe.70341.

17. X. Wu, J. Zhang, S. Wang, S. Zhu, H. Deng, and P. Yuan, in *Identification of Adulterated Honey with Different Varieties and Con-Centrations Based on Hyperspectral Imaging Technology Combined With Deep Learning* (IOP Publishing, 2023), 2552.