

**DR. BABASAHEB AMBEDKAR MARATHWADA UNIVERSITY,
CHHATRAPATI SAMBHAJINAGAR-431 004 (MS) INDIA
DEPARTMENT OF COMPUTER SCIENCE & IT**



NAAC REACCREDITED A+ GRADE

A THESIS ON

**DESIGN AND CREATION OF ADULTERANT DETECTION
SYSTEM FOR FOOD PRODUCTS USING HYPERSPECTRAL
NON-IMAGING TECHNIQUES**

In Partial Fulfilment Of
The Requirements for The Award of The Degree of
Doctor of Philosophy
In
Computer Science & IT

Submitted by
MR. VIJAY DHAKU DHANGAR

Under the guidance of
DR. BHARTI WAMANRAO GAWALI
Senior Professor

Department of Computer Science and Information Technology,
Dr. Babasaheb Ambedkar Marathwada University,
Chhatrapati Sambhajinagar - 431 004 (Ms) India

2025

ABSTRACT

Food adulteration has significant health hazards and economic repercussions, especially in spices and tea, where contaminants such as synthetic dyes, starch, brick powder, and used tea leaves are often employed. Although conventional detection techniques such as chemical analysis and microscopic examination are effective, they often prove to be time-consuming, labor-intensive, damaging, and costly. This research investigates the efficacy of hyperspectral spectroscopy as a non-destructive and efficient method for identifying food adulteration.

Hyperspectral non-imaging data were acquired using the ASD FieldSpec4 Spectroradiometer, which encompasses a wide spectral range to examine both pure and contaminated food samples. The research concentrated on chili, turmeric, and tea powders, using both artisanal and commercially produced samples to provide a varied and representative dataset. Commercial samples were obtained from several retail establishments in Chhatrapati Sambhajnagar City, Maharashtra, India, namely in Waluj, Padegaon, Jubly Park, and Begumpura. To preserve sample integrity, all obtained materials were carefully kept in sealed containers prior to analysis. To guarantee precision, the ASD FieldSpec4 Spectroradiometer was preheated for 30 minutes prior to spectral data acquisition. The measurement configuration included a 60 cm interval between the light source and the spectral gun, a 10 cm interval between the spectral gun and the sample, and an 8-degree field of view (FOV) for uniform illumination. White reference calibration was conducted to normalize the measurements. Preprocessing methods, such as Savitzky-Golay filtering with second derivative transformation and Standard Normal Variate (SNV), were used to eliminate noise and augment spectral characteristics. Noisy spectral bands (350-400 nm and 2400-2500 nm) were eliminated, hence narrowing the spectral range to 400-2400 nm. The Savitzky-Golay filter was used to smooth spectral curves while maintaining critical characteristics, using a 15-point window, a third-order polynomial, and first- and second-order derivatives. Feature extraction was performed using Principal Component Analysis (PCA) to diminish data dimensionality while preserving essential information for classification. The incorporation of extracted characteristics, including statistical traits like mean, standard deviation, skewness, kurtosis, energy, and area under the curve (AUC), enhanced classification accuracy. These attributes were used to construct

spectrum libraries, facilitating accurate sample classification using machine learning algorithms. The performance metrics across chili, turmeric, and tea classifications revealed distinctive patterns among the evaluated machine learning models. In chili classification, Random Forest demonstrated exceptional performance with a perfect 100% accuracy, precision, recall, and F1-score. XGBoost followed closely with 97% accuracy and high precision (0.98) and recall (0.97). KNN and SVM showed comparatively lower performance, with KNN achieving 85% accuracy and SVM reaching 82% accuracy. For turmeric classification, the performance trends mirrored the chili results. Random Forest again exhibited perfect 100% accuracy across all metrics, while XGBoost maintained strong performance with 93.94% accuracy. KNN demonstrated 90.91% accuracy, and SVM recorded the lowest performance at 84.85%. Notably, both precision and recall for XGBoost were 0.96 and 0.94 respectively, indicating robust classification capabilities. The tea classification results were remarkably consistent with the turmeric analysis. Random Forest once more achieved perfect 100% accuracy, and XGBoost replicated its previous performance with 93.94% accuracy. KNN maintained 90.91% accuracy, and SVM remained at 84.85%. These results validate that hyperspectral spectroscopy, in conjunction with machine learning, provides a quick, accurate, and non-invasive method for identifying food adulteration. A Graphical User Interface (GUI) was created to improve the usability of the Food Adulteration Detection System. This tool enables users to enter client and sample information, upload ASD spectral data for tea, turmeric, or chili samples, and get real-time adulteration findings driven by trained machine learning models.

This study demonstrates the considerable potential of combining hyperspectral spectroscopy with sophisticated machine learning methods to enhance food quality management. This strategy not only improves detection precision but also reduces dependence on conventional, labor-intensive techniques. Subsequent investigations and improvements in machine learning models may augment detection capabilities, assisting regulatory agencies and the food sector in guaranteeing food safety and quality.

INTRODUCTION



“Safe food is not a privilege, it’s a right.”

-Dr. Margaret Chan, Former Director General, WHO

The food industry is expanding rapidly worldwide, driven by increasing consumer demand, urbanization, and international trade. India, being one of the largest producers and consumers of agricultural and processed food products, faces a significant challenge in maintaining the quality and safety of its food. One of the most alarming issues is food adulteration, which involves the intentional addition or substitution of inferior, harmful, or non-permitted substances in food items to increase quantity and profit margins. According to the Food Safety and Standards Authority of India (FSSAI), a substantial number of food samples collected and tested annually are found to be adulterated, unsafe, or misbranded [1].

As the complexity of the food supply chain increases, traditional detection methods are becoming less effective and more time-consuming. There is an urgent need to adopt rapid, non-destructive, and intelligent technologies such as spectroscopy, hyperspectral imaging, and AI-based classification models to accurately detect adulteration and ensure consumer safety. These techniques offer high sensitivity, speed, and the potential for automation, making them ideal for modern food quality control systems [2].

Food adulteration includes both physical and chemical tampering such as mixing brick powder in chili powder, lead chromate in turmeric, or water in milk which can severely compromise human health. These adulterants can cause food poisoning, organ failure, gastrointestinal disorders, and even long-term chronic diseases or cancer. In many cases, consumers are unaware that the food they consume is contaminated, resulting in both financial loss and health hazards [3].

1.1. Overview.

Food adulteration is a significant global concern, affecting both public health and food quality. The intentional or unintentional addition of harmful substances to food products compromises their safety and can lead to severe health risks such as allergies, toxicity, and long-term illnesses. Spices and tea, including chili powder, turmeric powder, and tea powder, are often adulterated with materials such as artificial colours, starch, brick powder, and exhausted tea leaves. The consumption of such adulterated products not only presents serious health hazards but also leads to economic losses and a decline in consumer trust. Ensuring food safety in this context has become increasingly difficult, calling for the development of rapid, reliable, and non-destructive detection methods [4].

Traditional techniques used to detect food adulteration include High-Performance Liquid Chromatography (HPLC), Gas Chromatography-Mass Spectrometry (GC-MS), titrimetric methods, and chemical tests. Although these methods are known for their accuracy, they involve several limitations such as high costs, complex sample preparation, destructive analysis, and lengthy procedures. These challenges highlight the need for alternative approaches that are more efficient and practical [5].

In recent years, hyperspectral spectroscopy has emerged as a powerful, non-destructive technique for assessing food quality. This technology captures rich spectral signatures of food samples, making it possible to detect adulterants based on their specific chemical compositions. Especially in non-imaging mode, hyperspectral methods enable fast and accurate detection with minimal sample preparation, offering substantial benefits for food safety monitoring and regulatory compliance [6].

1.2 Hyperspectral Remote Sensing

1.2.1 Hyperspectral Remote Sensing: Imaging and Non-Imaging Spectroscopy

Hyperspectral remote sensing encompasses both imaging and non-imaging spectroscopy techniques, offering significant capabilities in material identification and analysis. Recent advancements in remote sensing and Geographic Information Systems (GIS) have contributed to the rapid evolution of hyperspectral sensor technology. This relatively novel approach is now widely adopted by researchers for diverse applications

such as crop species discrimination, soil characterization, identification of man-made materials, vegetation monitoring, and analysis of dairy products [7].

Although the concept of spectral analysis has existed for over a century, its application through hyperspectral remote sensing for the detailed evaluation of objects based on their biological and chemical properties has become increasingly sophisticated in recent decades. Hyperspectral sensors enable the acquisition of detailed physical and chemical information from a target in a non-destructive manner, making them valuable in both laboratory and field settings [8]. Hyperspectral sensors typically cover a broad range of the electromagnetic spectrum, including the visible (VIS), near-infrared (NIR), and short-wave infrared (SWIR) regions. Devices such as spectroradiometers, which operate predominantly in the NIR range, are especially effective in analyzing material composition due to their sensitivity to overtones and combination vibrations of molecular bonds [9, 10]. Spectroscopy operates on the principle of detecting specific absorption features associated with molecular bonds present in solid, liquid, or gaseous substances. These absorption characteristics are captured as spectral signatures, which are unique to each material and allow for accurate identification [9].

These sensors generate a spectral signature or spectrogram, which represents the reflectance or absorbance of the material at specific wavelengths, thereby aiding in the identification and quantification of components within the sample.

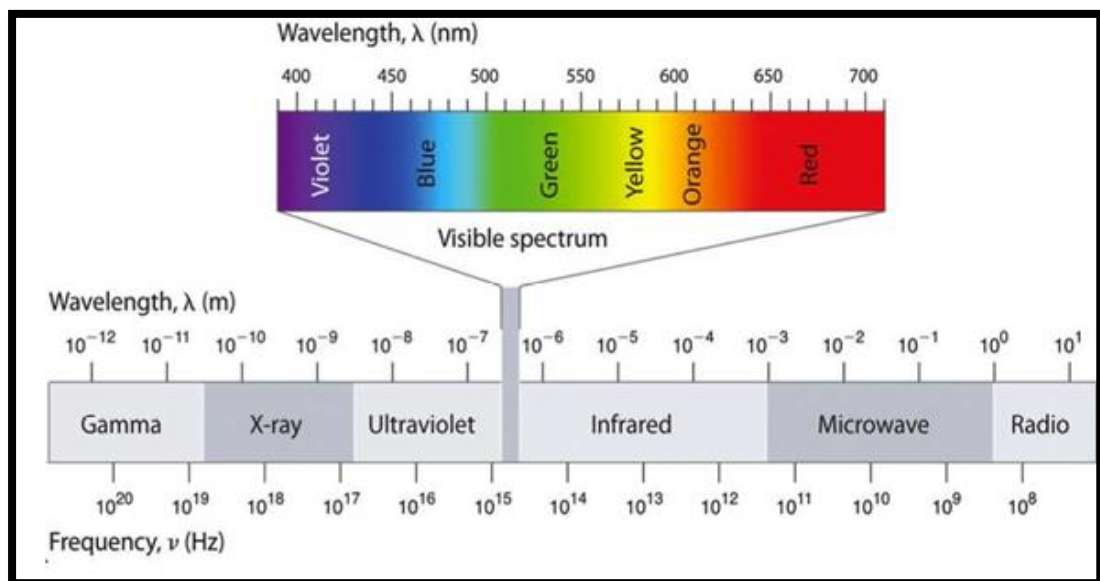


Figure 1.1: Electromagnetic Spectrum for Remote Sensing

1.3 Application of NIR Spectroscopy in the Food Adulteration Domain.

Near-Infrared (NIR) spectroscopy has emerged as a vital analytical tool in the detection of food adulteration due to its rapid, non-destructive, and reagent-free nature. Operating typically in the wavelength range of 780–2500 nm, NIR spectroscopy is capable of probing molecular overtones and combinations primarily associated with O–H, C–H, and N–H bonds [10]. These characteristics allow it to identify and quantify both organic and inorganic constituents in food products, making it highly effective for monitoring adulterants that alter the chemical composition of authentic food items.

In the food adulteration domain, NIR spectroscopy has been widely applied to detect unauthorized additives and contaminants in products such as spices, dairy, grains, edible oils, and beverages. For example, in powdered spices like chili powder and turmeric, NIR can distinguish between pure and adulterated samples based on spectral variations caused by the presence of brick powder, metanil yellow, or starch [11]. Similarly, in dairy products, it is employed to detect the addition of water, starch, or synthetic milk components that compromise quality and safety.

The strength of NIR spectroscopy lies in its ability to generate a unique spectral fingerprint of each sample, which can then be analyzed using multivariate statistical techniques such as Principal Component Analysis (PCA), Partial Least Squares (PLS), and Support Vector Machines (SVM) to classify pure and adulterated samples [12]. Furthermore, advancements in portable and handheld NIR devices have expanded its applicability for on-site testing in markets and supply chains, contributing to more efficient food safety surveillance.

Overall, NIR spectroscopy represents a powerful solution in combating food adulteration by enabling fast, accurate, and non-invasive detection of adulterants, thereby supporting regulatory enforcement and consumer protection.

1.3.1 Techniques for Analyzing Food Adulteration

Ensuring the authenticity and safety of food products has become a growing concern due to the increasing prevalence of adulteration. A wide range of analytical techniques, both conventional and advanced, are employed to detect adulterants in various food matrices. These techniques vary in terms of accuracy, sensitivity, cost, complexity, and whether they are destructive or non-destructive in nature.

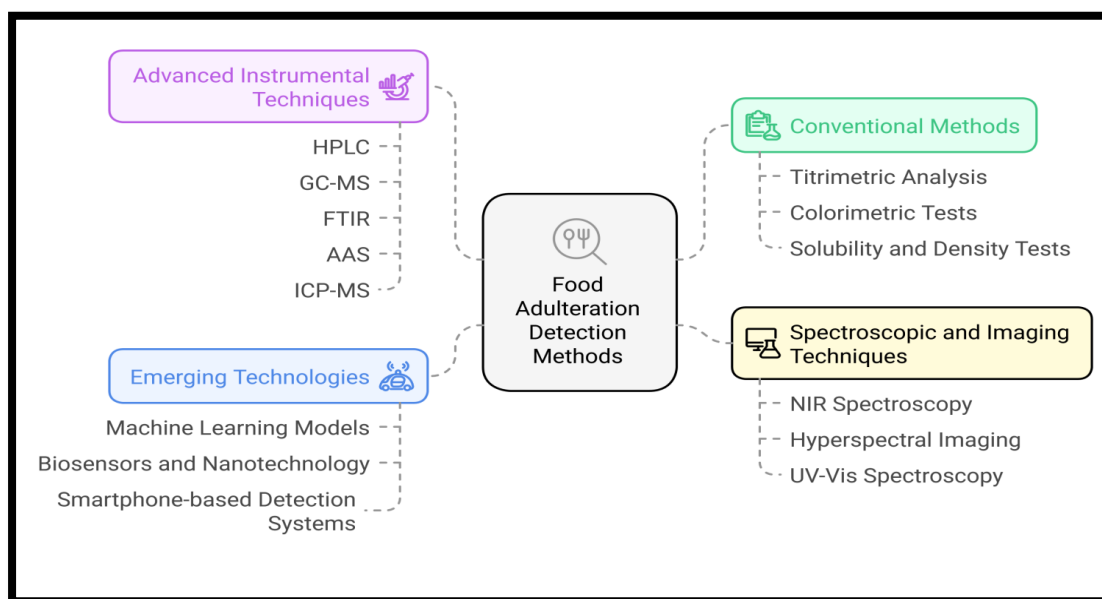


Figure 1.2: Techniques for Analyzing Food Adulteration

A. Conventional Methods

Traditional detection methods are based on basic chemical and physical tests that are widely used in households and local laboratories for preliminary screening. These include titrimetric analysis (e.g., detection of acidity or added water in milk), colorimetric tests (e.g., turmeric adulterated with metanil yellow), and physical property tests (e.g., solubility, density, float tests). While these methods are cost-effective and easy to implement, they lack sensitivity, are often qualitative, and cannot quantify the adulterant content precisely [13].

- **Merits:** These methods are based on simple chemical and physical tests that are inexpensive and easy to perform. They are useful for preliminary screening in households and local laboratories. Examples include titrimetric analysis for detecting water in milk and colorimetric tests for identifying metanil yellow in turmeric.
- **Demerits:** The main drawbacks are their lack of sensitivity, qualitative nature, and inability to precisely quantify the amount of adulterant present.

B. Advanced Instrumental Techniques

Modern food testing laboratories employ sophisticated instrumentation that offers higher sensitivity and specificity. These techniques include:

- **High-Performance Liquid Chromatography (HPLC):** Used to identify non-permitted dyes, synthetic sweeteners, and preservatives.
- **Gas Chromatography-Mass Spectrometry (GC-MS):** Effective for identifying volatile adulterants and contaminants such as pesticide residues.
- **Fourier Transform Infrared Spectroscopy (FTIR):** Detects adulterants based on molecular vibration patterns.
- **Atomic Absorption Spectroscopy (AAS):** Used for quantifying toxic heavy metals like arsenic, lead, and mercury.
- **Inductively Coupled Plasma Mass Spectrometry (ICP-MS):** Provides highly accurate detection of trace elements.

These methods offer high precision but often require extensive sample preparation, are time-consuming, expensive, and generally destructive to the sample [14].

- **Merits:** These techniques, such as High-Performance Liquid Chromatography (HPLC) and Gas Chromatography-Mass Spectrometry (GC-MS), are highly sensitive and specific. They can identify a wide range of adulterants, including non-permitted dyes, synthetic sweeteners, pesticide residues, and heavy metals.
- **Demerits:** They are often expensive, require complex and time-consuming sample preparation, and are destructive to the sample. These limitations make them impractical for large-scale, routine inspections.

C. Spectroscopic and Imaging Techniques

Recent advancements in analytical science have led to the adoption of non-destructive spectroscopic methods. These techniques provide rapid analysis and minimal sample preparation:

- **Near-Infrared (NIR) Spectroscopy:** Detects overtones and combinations of molecular vibrations. It is particularly useful in detecting starch in milk, artificial colorants in spices, and blending in edible oils.
- **Hyperspectral Imaging (HSI):** Combines spectral and spatial information, allowing for the detection of adulterants and contaminants in powdered and solid food materials.
- **Ultraviolet-Visible (UV-Vis) Spectroscopy:** Applied for assessing the optical properties and identifying synthetic dyes in food.

These techniques, especially when paired with multivariate data analysis (e.g., PCA, PLS), can effectively differentiate between pure and adulterated samples [15].

- **Merits:** Techniques like Near-Infrared (NIR) spectroscopy and Hyperspectral Imaging (HSI) are rapid, non-destructive, and require minimal sample preparation. They provide detailed spectral signatures of food samples, allowing for the detection of adulterants based on their chemical composition. NIR spectroscopy, in particular, is effective for identifying starch in milk, artificial colorants in spices, and blending in edible oils.
- **Demerits:** While highly effective, the initial setup can be costly, and the analysis of the rich spectral data often requires specialized knowledge and advanced data processing methods.

D. Emerging Technologies

The integration of artificial intelligence and sensor technologies is revolutionizing food adulteration detection:

- **Machine Learning Algorithms:** Used to classify spectral data and predict adulteration levels with high accuracy.
- **Biosensors and Nanotechnology:** Provide sensitive, on-site, and real-time detection capabilities.
- **Portable and Smartphone-based Devices:** Offer user-friendly solutions for field-level testing and rapid screening.

These emerging tools make it feasible to conduct food quality assessment outside of laboratory settings and contribute to decentralized food safety monitoring [16].

- **Merits:** This category includes the use of machine learning algorithms, biosensors, and smartphone-based devices. These technologies enable high-accuracy classification and prediction of adulteration levels, providing sensitive and real-time detection capabilities. Portable and smartphone-based devices make it possible to perform on-site testing outside of a traditional laboratory setting, contributing to decentralized food safety monitoring.
- **Demerits:** The development and validation of these systems can be complex, and their accuracy is highly dependent on the quality of the data used to train the models.

Table 1.1 Comparative Analysis of Detection Methods for Food Adulteration

Technique Category	Methods	Accuracy & Specificity	Cost	Sample Preparation	Analysis Time	Nature of Analysis
Conventional Methods	Titrimetric Analysis, Colorimetric Tests, Solubility & Density Tests	Low; often qualitative	Low	Simple	Short	Destructive
Advanced Instrumental Techniques	HPLC, GC-MS, FTIR, AAS, ICP-MS	High; offers high precision and specificity	High	Extensive and complex	Long	Destructive
Spectroscopic and Imaging Techniques	NIR Spectroscopy, Hyperspectral Imaging (HSI), UV-Vis Spectroscopy	High; can differentiate between pure and adulterated samples	Medium to High (initial setup)	Minimal	Rapid	Non-destructive
Emerging Technologies	Machine Learning Models, Biosensors, Smartphone-based Devices	High; can provide high-accuracy classification and prediction	Varies; can be low for smartphone-based systems	Varies; often minimal for on-site testing	Real-time to rapid	Non-destructive

1.3.2 Counterfeiting in Food Samples: Chilli Powder, Turmeric Powder, and Tea Powder

Counterfeiting in food products refers to the intentional addition or substitution of inferior or unauthorized substances to increase quantity and profitability, often at the cost of quality, safety, and consumer trust. Spices and tea powders, due to their powdered form and high commercial value, are among the most commonly adulterated food commodities. These counterfeit practices not only degrade the nutritional and sensory attributes of the food but may also pose serious health hazards.

A. Chilli Powder Adulteration

Chilli powder is frequently adulterated with non-permitted coloring agents and bulk-extending materials. Common adulterants include:

- **Brick powder** – added to mimic the red color and bulk.
- **Salt powder** – used to increase weight.
- **Sudan dyes (e.g., Sudan I, II)** – synthetic dyes used to enhance redness; these are carcinogenic and banned for food use.

These adulterants can lead to severe health issues such as stomach disorders, liver damage, and long-term toxicity [17].

B. Turmeric Powder Adulteration

Turmeric is widely valued for its curcumin content and medicinal properties, making it a common target for adulteration. Counterfeit practices include:

- **Metanil yellow** – a synthetic dye used to deepen the yellow color; it is toxic and non-permitted for human consumption.
- **Lead chromate** – added to enhance color and weight; it is highly toxic and can lead to lead poisoning, kidney damage, and neurotoxicity.
- **Chalk powder or starch** – added to increase bulk without visual detection.

Regular consumption of such adulterated turmeric can have carcinogenic effects and may interfere with metabolic and neurological functions [18].

C. Tea Powder Adulteration

Tea powder is also subject to various counterfeiting strategies, particularly through the use of:

- **Used and dried tea leaves** – exhausted tea leaves are recolored and resold as fresh.
- **Iron fillings or black lead** – used to artificially darken the color.
- **Artificial color or graphite powder** – added to give the appearance of freshness.

These adulterants degrade the antioxidant properties of tea and may introduce harmful heavy metals or synthetic chemicals into the human body [19].

1.3.3 Health and Economic Impacts

The consumption of counterfeit food products can lead to serious health risks including gastrointestinal irritation, liver damage, neurological disorders, and even carcinogenesis. Additionally, such fraudulent practices erode consumer confidence, damage brand integrity, and cause significant economic losses across the supply chain. Regulatory bodies like FSSAI in India and international agencies are actively promoting awareness, surveillance, and the implementation of advanced detection technologies to curb these practices.

1.4 Motivation

Food adulteration remains a persistent issue with severe implications for public health and economic stability. The inefficiencies of conventional laboratory-based techniques, such as HPLC and GC-MS, which require extensive sample processing, high costs, and long analysis times, create barriers to effective large-scale food quality monitoring. There is an urgent need for an alternative method that is cost-effective, rapid, non-destructive, and capable of delivering high-accuracy results for food adulteration detection.

This research explores the application of hyperspectral non-imaging spectroscopy as an advanced alternative for detecting food adulteration in chili powder, turmeric powder, and tea powder. Hyperspectral spectroscopy provides detailed spectral data that can differentiate between pure and adulterated samples by analyzing their unique spectral fingerprints. By integrating this technique with advanced data processing methods such as Savitzky-Golay filtering, Standard Normal Variate (SNV), and Mutual Information (MI), noise reduction and feature extraction can be significantly improved, enhancing the reliability of adulteration detection.

Furthermore, machine learning algorithms, including XGBoost, Decision Tree, Support Vector Machine (SVM), and others, are applied to develop predictive models for adulteration detection. These models leverage the rich spectral information provided by hyperspectral spectroscopy to classify and quantify adulterants with high accuracy. The combination of hyperspectral spectroscopy and machine learning offers a promising solution for real-time, non-destructive food quality assessment, ensuring better consumer protection and regulatory enforcement.

The findings of this study have significant implications for the food industry, as they provide a practical and scalable approach to food adulteration detection. By adopting hyperspectral spectroscopy-based techniques, food manufacturers, regulatory agencies, and quality control laboratories can enhance their monitoring processes, minimize health risks associated with adulterated food products, and improve overall food safety standards. Further research and model optimization can refine this approach, making it an essential tool for large-scale food quality assessment and regulatory compliance.

1.5 Objectives

The main aim of this study is to create a non-invasive, effective, and precise technique for identifying food adulteration employing hyperspectral spectroscopy and machine learning methods. The research centers on spices and tea, particularly chili powder, turmeric powder, and tea powder, which are often contaminated with dangerous additives. The specific objectives of this research are as follows:

- To create a spectral database of pure and adulterated samples of chili powder, turmeric powder, and tea powder using hyperspectral non-imaging spectroscopy.
- To analyze the spectral characteristics of food samples (chili powder, turmeric powder, and tea powder).
- To design and develop an efficient, non-destructive method for food adulteration detection using hyperspectral spectroscopy and machine learning techniques.
- To design Tools and Schema for organizing and interpreting hyperspectral data, facilitating the automation and scalability of food adulteration detection processes.

1.6 Major Contributions

This study showcases important progress in detecting food adulteration through hyperspectral spectroscopy and machine learning techniques. The main contributions of this research are outlined below:

1. **Creation of a Spectral Library** – An extensive spectral library was developed utilizing hyperspectral data alongside ENVI 5.5 software. This collection contains spectral signatures for:
 - 10 distinct kinds of chili powders.
 - 10 distinct varieties of turmeric powders.
 - 10 various kinds of tea powders.
 - Tea and turmeric samples prepared at home for reference and comparison.

2. **Identification of Spectral Bands for Chilli Powder** – Distinct spectral bands were found for several kinds of chili powders, facilitating the identification of adulterants via distinct spectral features.
3. **Identification of Spectral Bands for Turmeric Powder** – The research identified critical spectral bands for turmeric powders, effectively differentiating genuine samples from contaminated ones with great precision.
4. **Identification of Spectral Bands for Tea Powder** – Unique spectral bands for tea granules were identified, enabling the distinction between pure and contaminated samples.
5. **Identification of Spectral Bands for Common Adulterants** – The spectrum features of prevalent food adulterants, including starch, synthetic dyes, brick powder, and spent tea leaves, were analyzed and recorded to improve detection of adulteration.
6. **Design and Development of a Predictive Model** – A comprehensive machine learning model for adulteration detection was created using hyperspectral data. This model incorporates sophisticated feature selection and classification methods to enhance the precision and efficacy of food adulteration detection.

1.7 Organization of Thesis

The thesis is organized into five chapters, with each addressing various elements of the research, ranging from background and methods to experimental analysis and findings. The structure of the thesis is organized as follows:

Chapter 1: Introduction

This chapter provides a comprehensive introduction to hyperspectral non-imaging techniques and their significance in food adulteration detection. It discusses the fundamentals of spectral analysis, the challenges associated with food adulteration, and the need for advanced detection techniques. Additionally, it outlines the problem statement, research objectives, scope and limitations of the study, and the overall significance of the proposed work.

Chapter 2: Literature Review

This chapter presents a detailed review of related research on hyperspectral techniques for food quality assessment. It discusses conventional and modern adulteration detection methods, spectral feature extraction techniques, and machine learning models applied to hyperspectral data analysis. A critical evaluation of previous studies is conducted to highlight research gaps and justify the need for the proposed approach.

Chapter 3: Data Collection and Proposed Methodology

This chapter describes the design and implementation of the proposed adulteration detection system. It covers the system architecture, hardware and software requirements, and methodologies used for data acquisition. The preprocessing techniques for hyperspectral non-imaging data, including noise reduction and feature extraction, are also discussed. Additionally, it provides an overview of the dataset characteristics and the process of structuring hyperspectral data in CSV format.

Chapter 4: Results and Discussion.

This chapter presents the experimental analysis and interpretation of results. It includes details on data pre-processing, feature selection, and machine learning model training for adulteration detection. The performance of different machine learning algorithms is evaluated, and their effectiveness in detecting food adulterants is analyzed. Furthermore, the development of a graphical user interface (GUI) for practical implementation is discussed.

Chapter 5: Conclusions and Future Scope

The final chapter summarizes the key findings and contributions of the research. It highlights the practical applications of hyperspectral spectroscopy for detecting food adulteration and discusses its potential impact on quality control in the food industry. The limitations of the current study are addressed, and recommendations for future research are provided, focusing on the further optimization of models, integration with real-time monitoring systems, and the expansion of the study to a broader range of food products.

1.8 References:

- [1]. FSSAI. (2022). Annual Report 2021–2022. Food Safety and Standards Authority of India. Retrieved from <https://www.fssai.gov.in>
- [2]. Kaur, H., Singh, D., & Ahuja, M. (2020). Application of spectroscopy and machine learning techniques in food adulteration detection: A review. *Food Analytical Methods*, 13(2), 305–321. <https://doi.org/10.1007/s12161-019-01648-4>
- [3]. Sharma, V., & Agarwal, R. (2021). Food adulteration and its impact on public health – A review. *International Journal of Food Science and Nutrition*, 6(1), 25–30.
- [4]. Zhang, Y., Wang, H., & Li, Y. (2021). Food adulteration: Causes, detection, and prevention. *Food Control*, 127, 108148.
- [5]. <https://doi.org/10.1016/j.foodcont.2021.108148>
- [6]. Patel, K., Shah, S., & Desai, B. (2019). Detection of food adulterants in spices and condiments using chemical methods: A review. *International Journal of Chemical Studies*, 7(1), 503–507.
- [7]. Sun, D. W., Pu, H., & Pan, L. (2022). Recent developments in hyperspectral imaging for non-destructive detection of food quality and safety. *Critical Reviews in Food Science and Nutrition*, 62(1), 1–17. <https://doi.org/10.1080/10408398.2020.1839986>
- [8]. Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *TrAC Trends in Analytical Chemistry*, 26(7), 590–598. <https://doi.org/10.1016/j.trac.2007.04.006>
- [9]. Govender, M., Chetty, K., & Bulcock, H. (2007). A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA*, 33(2), 145–151. <https://doi.org/10.4314/wsa.v33i2.49045>
- [10]. R. Lu and B. A. Rasco, *Imaging and Spectroscopy for Food Quality and Safety*. Hoboken, NJ, USA: John Wiley & Sons, 2012.
- [11]. Pasquini, C. (2018). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14(2), 198–219. <https://doi.org/10.1590/S0103-50532003000200006>.
- [12]. Liu, J., Chen, J., Dong, N., Ming, J., & Zhao, G. (2012). Determination of degree of substitution of carboxymethyl starch by Fourier transform mid-infrared spectroscopy coupled with partial least squares. *Food Chemistry*, 132(4), 2224–2230.
- [13]. Cozzolino, D., & Murray, I. (2004). Identification of animal meat muscles by visible and near infrared reflectance spectroscopy. *LWT-Food Science and Technology*, 37(4), 447–452. <https://doi.org/10.1016/j.lwt.2003.10.008>
- [14]. Patel, M., Kumar, R., Kishor, K., Mlsna, T., Pittman Jr, C. U., & Mohan, D. (2019). Pharmaceuticals of emerging concern in aquatic systems: chemistry,

- occurrence, effects, and removal methods. *Chemical reviews*, 119(6), 3510-3673.
- [15]. Poonia, A., Jha, A., Mehta, S., Sharma, R., & Rai, A. K. (2016). Detection of adulteration in milk: A review. *International Journal of Dairy Technology*, 69(4), 557–570. <https://doi.org/10.1111/1471-0307.12340>
- [16]. Rohman, A., & Che Man, Y. B. (2011). The use of infrared spectroscopy in combination with chemometrics for detecting and quantifying adulterants in food and agricultural products. *Applied Spectroscopy Reviews*, 46(6), 420–430. doi.org/10.1080/05704928.2011.605068.
- [17]. Kumar, N., Acharya, A., Gendelman, H. E., & Byrareddy, S. N. (2022). The 2022 outbreak and the pathobiology of the monkeypox virus. *Journal of autoimmunity*, 131, 102855.
- [18]. Rathore, H., Singh, A., & Kaushik, N. (2020). Food adulteration and its harmful effects on health. *International Journal of Food and Nutritional Sciences*, 9(3), 145–150.
- [19]. Kumar, M., Yadav, D. N., & Rai, K. N. (2021). Detection of food adulteration in turmeric: Techniques and health risks. *Journal of Food Quality*, 2021, Article ID 5597283. <https://doi.org/10.1155/2021/5597283>
- [20]. Sharma, P., & Jain, R. (2018). Tea adulteration and its health consequences: A review. *Asian Journal of Dairy and Food Research*, 37(1), 1–6. <https://doi.org/10.18805/ajdfr.DR-1248>

LITERATURE SURVEY

2.1 Literature Survey

Numerous techniques are available for analyzing food products, especially to detect adulteration and ensure quality and safety. Currently, both modern techniques and traditional methods are widely used in this domain. Although modern technologies offer improved accuracy and speed, traditional techniques are still commonly used due to their familiarity and widespread acceptance. However, these conventional methods are often labor-intensive, time-consuming, and mostly destructive, which limits their application in rapid screening. A review of the existing literature reveals that most previous studies in this area have employed traditional techniques because of their accessibility and established protocols. Based on the analysis of various research works, it was observed that three major types of techniques are commonly utilized for food adulteration detection: Chromatographic techniques, Hyperspectral Imaging, and Spectroscopic methods. This research focuses on advancing the domain by utilizing hyperspectral non-imaging techniques, which are emerging as a powerful alternative due to their non-destructive nature, speed, and effectiveness in detecting subtle compositional changes in food products.

Food adulteration is a persistent global issue that compromises food safety, public health, and consumer rights. Various powdered food products such as chili powder, turmeric powder, and tea powder are frequently adulterated with substances like brick powder, starch, salt powder, synthetic dyes, and exhausted tea leaves. This chapter reviews existing techniques used for adulteration detection, highlighting their limitations and establishing the relevance of hyperspectral non-imaging techniques for developing an efficient, non-destructive, and rapid adulterant detection system.

The incorporation of adulterants in food items may result in significant health issues, including foodborne diseases and chronic health impacts. Historically, the identification of food adulterants has depended on chemical and physical investigations, sometimes requiring extensive sample preparation, advanced laboratory apparatus, and considerable time investment. Consequently, there is an increasing need for rapid, non-invasive, and effective ways to identify adulteration. Hyperspectral non-imaging techniques have developed as a potent alternative, offering improved accuracy, sensitivity, and specificity for detecting adulterants in diverse food items. This literature

review examines current methodologies for detecting food adulteration, emphasizing chromatographic techniques, hyperspectral imaging, and spectroscopic methods [1, 2].

2.2 Conventional Techniques

Singh and Yadav examined a variety of household-level detection techniques used by consumers to identify common food adulterants, emphasizing low-cost and easily accessible methods that can be performed without specialized equipment. Their study highlighted simple procedures such as solubility tests for detecting chalk powder in sugar and acid-based reactions to reveal the presence of washing soda in food samples. Derived largely from traditional practices, these techniques serve as preliminary screening tools requiring minimal resources, making them particularly valuable in rural and low-resource settings where access to modern analytical facilities is limited. While these methods empower consumers to conduct basic food quality checks, the authors acknowledged their limitations, including low precision, poor sensitivity, and a reliance on subjective interpretation. Singh and Yadav concluded that although such tests are useful first-line indicators of adulteration, there is a pressing need for more reliable, standardized detection methods to ensure food safety and compliance with regulatory standards [3].

The traditional colorimetric technique used by Jain et al. for detecting non-permitted synthetic dyes, such as metanil yellow in turmeric and chili powder, involves: Solvent Extraction: Adulterated spice samples are treated with methanol, which dissolves both natural pigments and synthetic dyes. UV Light Inspection: The methanol extract is placed under ultraviolet (UV) light, where metanil yellow exhibits a distinct fluorescence or color shift, making it visually identifiable. While this method is fast, low-cost, and requires minimal equipment making it highly suitable for field-level preliminary screening it is semi-quantitative and prone to subjective interpretation of color changes under UV light. This reduces its reliability for strict regulatory enforcement. The study demonstrates how low-cost, color-based detection remains valuable where advanced instruments are unavailable. It also shows the potential for adapting such traditional methods into more standardized protocols for modern food safety monitoring [4].

Gupta and Garg highlighted the use of microscopic analysis as an effective method for detecting adulterants in powdered food products, particularly spices such as turmeric. Their approach involved preparing small amounts of spice samples on glass slides and examining them under a compound light microscope to observe morphological features such as shape, size, and surface texture. This allowed for the identification of contaminants like starch grains, sawdust, and powdered seeds, which often resemble genuine spice particles and are difficult to detect through visual inspection or simple chemical tests. The method proved to be a cost effective and reliable laboratory scale tool, capable of revealing adulterants that might otherwise go unnoticed, thereby reducing the risk of false negatives. While the technique requires trained personnel to interpret the results and is not suited for quantifying adulteration, Gupta and Garg emphasized its importance in routine quality assessment and spice authentication, reinforcing its role in maintaining food safety standards [5].

Joshi, and Mehta investigated density- and buoyancy-based techniques for detecting adulteration in mustard oil, with a focus on identifying the presence of argemone oil, a toxic and commonly used adulterant. Their method measured the differences in specific gravity between pure and adulterated samples, enabling the development of a simple visual test that could be conducted in basic laboratory settings without the need for advanced instrumentation. By exploiting the physical property variations of the oils, the researchers demonstrated a practical approach suited for low-resource and rural environments, where sophisticated analytical tools may not be available. Although the technique is not highly sensitive and cannot detect adulterants at trace levels, it provides a rapid preliminary screening method to flag potentially contaminated samples for further analysis. Joshi, Sood, and Mehta emphasized that such straightforward applications of basic physicochemical principles can play a valuable role in accessible, mass-level food safety screening [6].

Bansal et al. presented a comprehensive overview of traditional food adulteration detection methods commonly used in Indian markets, with particular emphasis on household-level techniques. Their study documented the use of simple reagents such as water, iodine solution, and acids to detect common adulterants, including chalk powder in sugar, starch in dairy products, and synthetic compounds in milk. These tests relied on visible reactions such as color changes, sedimentation, or

effervescence to indicate contamination and were designed to be performed without specialized equipment. The authors highlighted the cultural and practical importance of such methods in promoting consumer awareness and vigilance, especially in resource-constrained settings. However, they acknowledged the qualitative nature of these approaches and their susceptibility to false positives or negatives due to variations in user interpretation. Despite these limitations, Bansal et al. emphasized that traditional techniques remain a valuable first step in preliminary adulteration screening and advocated for their integration with modern analytical systems to enhance food safety monitoring and regulatory enforcement [7].

Rathore and Sharma assessed the effectiveness of flotation-based methods for detecting adulteration in cereal products, particularly wheat flour. Their study targeted common adulterants such as husk and bran, which are often added to increase bulk and reduce production costs. Using water as a separation medium, they demonstrated that these adulterants could be visually distinguished from the denser flour particles due to their lower specific gravity and natural tendency to float. This simple and cost-effective technique requires minimal equipment and can be performed easily in domestic or field settings, making it practical for consumers and small-scale food inspectors. While the method is limited to detecting adulterants with distinct physical properties, it offers an accessible, non-chemical screening option. Rathore and Sharma emphasized that such basic physical tests retain significant value for food adulteration detection, especially in regions lacking advanced laboratory facilities [8].

The Food Safety and Standards Authority of India (FSSAI) introduced the DART (Detect Adulteration with Rapid Test) Booklet as a comprehensive guide to help consumers identify food adulteration using simple, rapid, and low-cost methods. Containing over 50 tests, the booklet outlines procedures that use easily available household reagents such as water, acids, iodine, and detergents to detect common adulterants in a wide range of food items. Designed for ease of use, these tests enable quick visual identification without requiring specialized equipment or technical training. As a government-led initiative, the DART booklet plays a key role in raising public awareness and strengthening food safety at the grassroots level. Although primarily qualitative in nature, these tests are effective preliminary screening tools, especially in low-resource environments. FSSAI's effort underscores the value of

community participation in food quality monitoring and marks a significant step toward improving food safety literacy across the country [9].

2.3 Chromatographic Techniques

Lopez-Ruiz et al. employed Gas Chromatography Mass Spectrometry (GC-MS) to detect non-permitted synthetic colorants in spices, focusing on chili and turmeric powders. Their work demonstrated that GC-MS provides exceptional sensitivity and specificity, enabling the identification of adulterants at trace levels, even within complex food matrices. The technique allowed for precise separation and characterization of multiple compounds, proving effective in distinguishing between natural and synthetic additives. The authors highlighted GC-MS's robustness, high precision, and reproducibility, making it particularly suitable for regulatory and forensic food analysis. While the method requires sophisticated instrumentation and skilled operators, it remains a gold standard for detecting volatile and semi-volatile adulterants in laboratory settings. This study reinforces the critical role of chromatographic methods in ensuring spice authenticity and safeguarding consumer health [10].

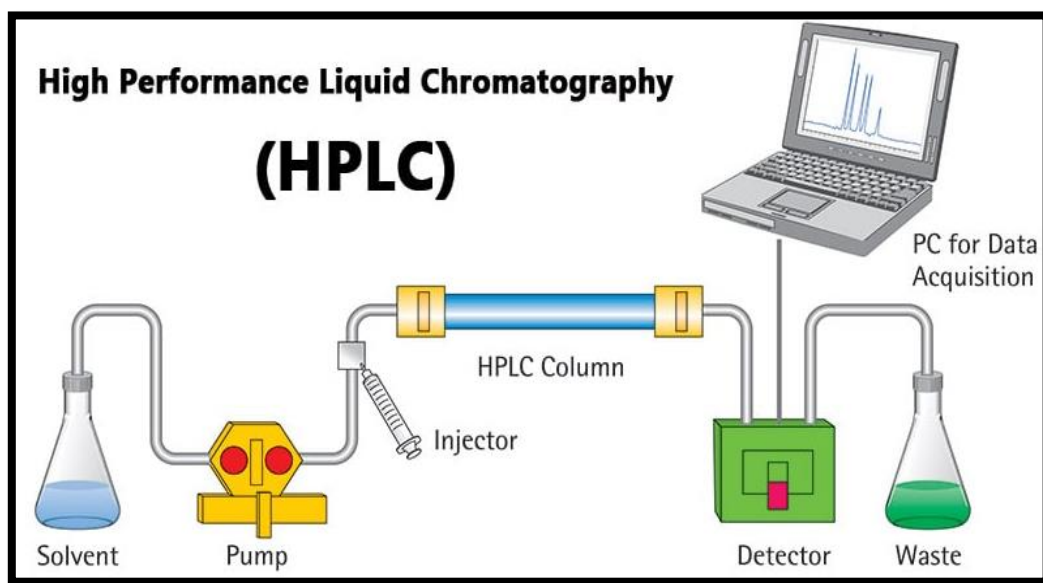


Figure 2.1: High Performance Liquid Chromatography

Ramesh and Saraf developed a High-Performance Liquid Chromatography (HPLC) method for detecting Sudan dyes in chili powder, an illegal synthetic colorant with toxic properties. Their analytical process involved solvent extraction of dye

components, followed by reversed-phase HPLC to achieve precise separation and quantification as shown in figure 2.1. The method demonstrated high sensitivity, specificity, and repeatability, confirming its effectiveness in detecting even low concentrations of synthetic dyes within complex spice matrices. The authors emphasized HPLC's reliability as a standard analytical technique for identifying non-volatile, thermally unstable compounds in food adulteration studies. Although the approach requires laboratory infrastructure and skilled personnel, it remains indispensable for confirmatory testing and regulatory enforcement in spice quality assurance [11].

Naik et al. demonstrated the use of Thin Layer Chromatography (TLC) for detecting non-permitted synthetic colorants in turmeric powder. Their method involved solvent extraction of the samples, followed by chromatographic separation on TLC plates and visualization under ultraviolet (UV) light. To ensure accuracy, the study incorporated validation steps such as checking for consistent retention factor (R_f) values and verifying spot coloration, confirming the method's repeatability and reliability. The results highlighted TLC as a simple, low-cost, and rapid screening tool well-suited for preliminary adulteration analysis, particularly in resource-limited settings. While semi-quantitative in nature, TLC remains valuable for routine quality control and often serves as a precursor to more advanced confirmatory techniques.

Singh and Bhatnagar investigated the use of Gas Chromatography (GC) for detecting adulteration in edible oils, focusing on identifying linseed and palm oil in mustard oil. Their approach employed fatty acid profiling as a diagnostic tool, allowing detection of compositional changes that signal adulteration. By analyzing the characteristic fatty acid signatures unique to each oil type, they demonstrated GC's precision and reliability in authenticating oils. The study confirmed GC's effectiveness in distinguishing pure from adulterated samples, reinforcing its value in food quality control and regulatory testing. Singh and Bhatnagar emphasized the importance of such chromatographic techniques in safeguarding the authenticity of edible oils, which are often targets of economically motivated adulteration.

Khanna et al. employed Gas Chromatography coupled with Flame Ionization Detection (GC-FID) to differentiate between pure and adulterated mustard oil. Their method targeted subtle variations in fatty acid composition caused by the presence of

adulterants. The technique demonstrated high sensitivity, with a detection limit as low as 0.5%, making it effective for identifying even minor levels of adulteration. The study validated GC-FID's reliability and precision for routine oil quality assessment, particularly in cases of economically motivated adulteration. Khanna et al. highlighted the method's suitability for large-scale screening and regulatory applications, where both accuracy and sensitivity are essential [14].

Rai and Kaur employed High-Performance Thin-Layer Chromatography (HPTLC) to detect synthetic colorants in commercial beverages, demonstrating the technique's ability to separate and identify artificial dyes through their distinct chromatographic profiles. The method enabled rapid analysis with high sample throughput, making it well-suited for large-scale screening. They highlighted several advantages of HPTLC, including cost-effectiveness, minimal sample preparation, and ease of visual interpretation. Their results validated HPTLC as a reliable tool for monitoring non-permitted additives in food and beverage products, particularly in regulatory enforcement and quality control applications [15].

Shukla and Tiwari demonstrated the effectiveness of Gas Chromatography Mass Spectrometry (GC-MS) for detecting pesticide residues in tea leaves, a commodity prone to contamination from extensive agricultural chemical use. Their approach enabled the simultaneous identification and quantification of multiple pesticide compounds at trace concentrations, with detection limits reaching the parts-per-billion (ppb) range. Using solvent extraction, followed by chromatographic separation and mass spectrometric detection, the method allowed for structural elucidation of unknown residues through analysis of mass fragmentation patterns. The study underscored GC-MS's sensitivity, selectivity, and robustness in handling complex food matrices while maintaining high accuracy and reproducibility, positioning it as a gold-standard technique for residue analysis. Although the method requires sophisticated instrumentation and skilled personnel, it is indispensable for confirmatory testing, particularly in the context of food export regulations and compliance with international safety standards. Their work provides a validated protocol for pesticide screening in herbal and leaf-based food products, contributing significantly to the field of food adulteration detection [16].

Verma et al. applied Liquid Chromatography Tandem Mass Spectrometry (LC-MS/MS) for multi-residue detection of contaminants across a variety of food products. In contrast to GC-MS, which is better suited for volatile and semi-volatile compounds, LC-MS/MS excels in analyzing thermally unstable, polar, and non-volatile substances that cannot be efficiently examined in the gas phase. Their method enabled the simultaneous identification and quantification of more than twenty chemical adulterants and contaminants in a single run, including pesticide residues, synthetic dyes, antibiotic traces, and other illicit additives found in spices, cereals, and processed powders. Verma et al. highlighted LC-MS/MS's strengths in high analytical throughput, precision, and multi-analyte capability with minimal sample preparation. The use of tandem mass spectrometry enhanced the detection of structurally similar compounds and reduced false positives through multiple reaction monitoring (MRM) transitions. Their findings established LC-MS/MS as an ideal tool for comprehensive contaminant profiling, offering regulatory authorities a robust method for large-scale food safety surveillance and enforcement [17].

Together, these studies underscore the pivotal role of advanced mass spectrometry-based chromatographic techniques, particularly GC-MS and LC-MS/MS, in modern food adulteration detection. They reflect a clear shift from single-analyte, labor-intensive approaches to integrated, high-resolution analytical systems capable of rapidly identifying multiple contaminants within complex food matrices. By enabling highly sensitive, specific, and multi-residue analysis, these technologies address the escalating challenges posed by food fraud, strengthen consumer protection, and support compliance with stringent international quality and safety standards [17].

2.4 Spectrometric Techniques

Sharma et al. conducted a detailed investigation into toxic heavy metal contamination in turmeric powder using Atomic Absorption Spectroscopy (AAS). Their study targeted hazardous elements such as lead (Pb) and arsenic (As), which may be present due to environmental pollution, poor agricultural practices, or intentional adulteration particularly through the addition of lead chromate to enhance turmeric's yellow color. Turmeric samples collected from local markets were subjected to acid digestion before AAS analysis, enabling high-precision quantification of contamination levels. The results were alarming: nearly 30% of the samples exceeded

permissible limits set by FSSAI and Codex Alimentarius. Elevated concentrations of lead and arsenic raised serious public health concerns, including risks of neurological disorders, kidney damage, and cancer. Beyond documenting widespread contamination in informal markets, the study highlighted the dangers of economically motivated adulteration. Sharma et al. demonstrated that AAS is highly effective in distinguishing naturally pigmented turmeric from samples enhanced with toxic additives, underscoring its value as a regulatory tool for food safety enforcement [18].

In terms of technique, Atomic Absorption Spectroscopy (AAS) proved highly valuable for its sensitivity, specificity, and accuracy, with the capability to detect trace metal concentrations in the ppm to ppb range. While the method requires careful sample preparation and may encounter matrix interferences, it remains more cost-effective than advanced alternatives such as ICP-MS. Sharma et al. emphasized the broader implications of their findings, advocating for the integration of AAS into routine food monitoring systems to strengthen detection of heavy metal contamination. Their work not only adds to the growing literature on food adulteration but also supports policy development and regulatory interventions aimed at safeguarding public health from heavy metal exposure in widely consumed spices like turmeric [18].

Misra et al. employed microwave-assisted acid digestion followed by Inductively Coupled Plasma Mass Spectrometry (ICP-MS) to detect toxic heavy metals in spices with exceptional precision. This approach ensured complete breakdown of organic matrices, allowing accurate quantification of elements such as cadmium (Cd), mercury (Hg), and arsenic (As) at trace levels in the parts-per-billion range. The authors emphasized ICP-MS's strengths, including high throughput, broad dynamic range, ultra-low detection limits, and its ability to perform simultaneous multi-element analysis, making it ideal for comprehensive screening in food safety laboratories. Method reliability was reinforced through the use of internal standards, advanced interference correction protocols, and validation against certified reference materials with spiked recovery experiments.

Beyond identifying contaminated samples, Misra et al. examined contamination sources within the spice supply chain, ranging from environmental pollution (e.g., contaminated irrigation water and soil) to post-harvest handling and intentional adulteration with toxic compounds like lead chromate for color enhancement. They

advocated for routine surveillance, stricter quality control, and the integration of ICP-MS into national and international food testing protocols, particularly for export-quality spices, to protect public health and maintain global market access. Their work underscores ICP-MS's role as a state-of-the-art tool in combating food adulteration and ensuring the safety of essential dietary commodities.

Complementing this, Aggarwal and Nair demonstrated the effective use of Atomic Absorption Spectroscopy (AAS) for detecting lead chromate adulteration in turmeric. By quantifying lead levels in samples collected from various markets, they provided direct evidence of intentional addition of lead-based compounds to enhance color. The sensitivity and accuracy of AAS made it a valuable method for routine monitoring, and their findings supported regulatory interventions aimed at banning toxic colorants. This study highlighted the importance of instrumental detection techniques not only for identifying adulterants but also for influencing public health policy and strengthening food safety enforcement [20].

Gupta et al. conducted a comparative study using Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) and X-ray Fluorescence (XRF) to analyze the elemental composition of food powders. While ICP-OES demonstrated high sensitivity and precision for detailed multi-element analysis in laboratory settings, the study emphasized the value of XRF as a rapid, non-destructive screening tool suitable for on-site applications. XRF enabled quick detection of metallic adulterants and mineral contaminants without the need for extensive sample preparation, making it especially useful for field-level screening and high-throughput environments. The findings highlighted that although ICP-OES remains the preferred confirmatory method, XRF offers a practical, low-cost alternative for preliminary testing in resource-limited settings. Gupta et al. concluded that this dual-approach strategy combining portable tools like XRF with laboratory-based techniques like ICP-OES strengthens modern workflows for food adulteration detection and quality assurance [21].

Khan et al. conducted a comprehensive review on the use of mass spectrometry for detecting adulteration in milk, with particular emphasis on its ability to identify anomalies in protein profiles. Their analysis highlighted the technique's effectiveness in detecting synthetic additives and protein-based adulterants such as melamine, urea,

and whey powder, all of which alter the natural protein composition of milk. Through precise molecular characterization, mass spectrometry was shown to provide high sensitivity and the capability to detect both known and emerging adulterants at trace levels. The authors concluded that mass spectrometry holds significant advantages for confirmatory testing and molecular-level authentication of dairy products. They further emphasized its importance in food safety monitoring, especially in light of the growing sophistication of milk adulteration practices, underscoring its role as a critical tool for ensuring authenticity and consumer protection [22].

Kumar et al. utilized Atomic Absorption Spectroscopy (AAS) to detect cadmium contamination in commercially available tea leaves, uncovering notably higher levels of the toxic metal in several low-cost brands. Their findings raised significant concerns about the safety and quality of economically priced tea products, pointing to the heightened vulnerability of such goods to substandard agricultural and processing practices. The study emphasized the value of AAS as a sensitive, accurate, and reliable method for monitoring toxic metal residues in widely consumed beverages. Kumar et al. further advocated for the integration of routine AAS monitoring into quality assurance systems, stressing the need for stricter regulatory oversight and quality control measures, particularly for products sourced from unregulated supply chains that pose heightened risks to public health [23].

Sinha et al. highlighted the strategic adoption of Atomic Absorption Spectroscopy (AAS) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS) in Indian government food testing laboratories, with a particular focus on export-grade spices. Their study underscored the vital role these techniques play in detecting heavy metal contaminants such as lead, arsenic, and cadmium, which pose serious risks to food safety and threaten compliance with international trade standards. The authors noted that the integration of AAS and ICP-MS has significantly enhanced analytical precision and reliability, enabling laboratories to align with stringent global regulations. This advancement was presented as a critical step in strengthening India's food quality infrastructure, ensuring both consumer protection and market competitiveness in the global spice trade [24].

Paul and Dutta applied Inductively Coupled Plasma Mass Spectrometry (ICP-MS) to evaluate lead contamination in chili powder, a widely consumed spice. Their

analysis detected trace concentrations of lead in several market samples, raising concerns about potential health risks from chronic exposure, even at low levels. The high sensitivity and precision of ICP-MS enabled accurate quantification of lead, reinforcing its effectiveness in identifying toxic metal adulterants within complex food matrices. Based on these findings, the authors called for stricter market surveillance and routine screening of spices to safeguard consumer health and ensure regulatory compliance. Their study highlighted the importance of incorporating advanced analytical techniques into adulteration monitoring frameworks to uphold food quality standards in both domestic and export markets [25].

2.5 Spectroscopic Techniques

Spectroscopic techniques, particularly those utilizing hyperspectral approaches, have emerged as central tools in modern food adulteration detection because of their rapid, non-destructive, and reagent-free capabilities. Among these, Hyperspectral Imaging (HSI) is especially powerful, as it integrates both spatial and spectral information to detect and localize adulterants within complex food matrices such as powders, grains, and spices. By capturing a complete spectrum at every pixel, HSI generates detailed spectral maps that can highlight subtle chemical and physical variations between authentic and adulterated samples, making it a highly effective technique for food quality assessment and safety monitoring [26]. In contrast, non-imaging hyperspectral techniques focus exclusively on spectral information without providing spatial resolution. These methods typically employ field-portable or benchtop spectrometers to capture reflectance or absorbance data from a single point or bulk sample surface. Both imaging and non-imaging systems rely on multivariate chemometric models such as Principal Component Analysis (PCA), Partial Least Squares Regression (PLSR), and various machine learning algorithms to classify, detect, and quantify adulterants. The integration of these hyperspectral methods into food safety frameworks offers high sensitivity, reproducibility, and the potential for real-time, in-line monitoring, making them indispensable in combating economically motivated food fraud and ensuring consumer protection [27].

2.6 Hyperspectral Imaging (HSI) for Food Adulteration Detection

Hyperspectral Imaging (HSI) has emerged as a cutting-edge technique in the field of food quality and safety, particularly for detecting adulteration as shown in figure 2.2. Unlike traditional imaging or spectroscopy alone, HSI integrates spatial and spectral information, producing a hypercube in which each pixel contains a full spectral profile. This allows not only for the chemical identification of adulterants but also for their localization within heterogeneous food samples. Its non-destructive, reagent-free, and real-time analysis capabilities make it a promising tool for food safety monitoring.

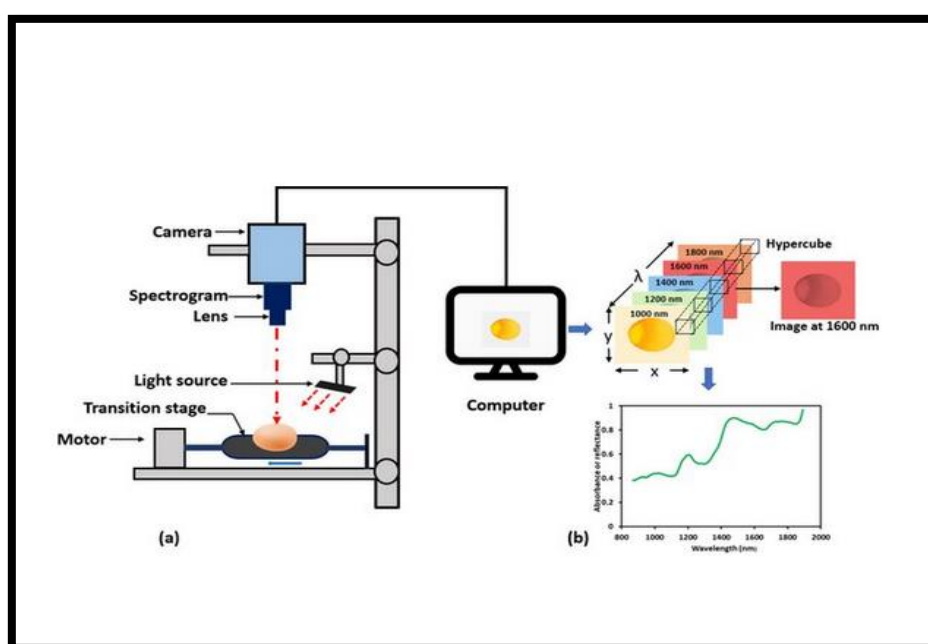


Figure 2.2: Hyperspectral Imaging Setup

The integration of deep learning techniques with hyperspectral imaging (HSI) has greatly enhanced food adulteration detection by enabling high-speed, high-accuracy, and automated analysis. In this regard, Weng et al. demonstrated a novel application of Convolutional Neural Networks (CNNs) to classify adulterated chili powder samples using hyperspectral image data. Their study showcased the strength of deep learning in extracting spectral-spatial features directly from high-dimensional HSI data, eliminating the need for manual feature selection or expert-driven preprocessing.

In their work, hyperspectral images were captured in the visible to near-infrared range (400–1000 nm) to build spectral cubes, where each pixel contained reflectance

values across multiple narrow spectral bands. These cubes preserved both spectral information related to chemical composition and spatial information linked to adulterant distribution. Traditionally, such data are analyzed using techniques like Principal Component Analysis (PCA) or Partial Least Squares (PLS), which require manual input and are vulnerable to noise and irrelevant variation. By contrast, CNNs automatically learn hierarchical representations from raw HSI cubes, making the analysis more robust, efficient, and scalable for practical food safety applications [28].

The proposed CNN architecture by Weng et al, was designed to accept 3D input data, incorporating convolutional layers that extracted localized spectral–spatial patterns. These patterns are often indicative of subtle physical or chemical anomalies associated with adulteration. In their experimental setup, the authors created a dataset comprising pure chili powder and samples adulterated with low-cost colorants and starch-based fillers. The CNN model outperformed traditional machine learning classifiers such as Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN), achieving classification accuracies exceeding 95%. This result underscored the advantage of using deep architectures to model complex, nonlinear relationships within hyperspectral data.

Furthermore, the study highlighted the ability of CNNs to generalize across sample variations and illumination inconsistencies, which are common challenges in food inspection scenarios. Their deep learning approach also provided faster inference times once trained, making it a viable candidate for real-time, high-throughput food authentication systems in industrial settings. The use of data augmentation techniques and dropout layers also contributed to reducing overfitting, enhancing the model's robustness and practical applicability. One of the major contributions of Weng et al, research lies in demonstrating the synergistic potential of artificial intelligence (AI) and HSI technologies in solving real-world problems in food safety. By combining the fine-grained spectral sensitivity of HSI with the automated feature learning capabilities of CNNs, the study presents a framework that is not only accurate but also scalable. Such systems could eventually be deployed on conveyor belts in food processing units or integrated into portable, field-deployable devices for rapid adulteration screening. The implications of this study extend beyond chili powder authentication. The same methodological framework can be adapted to a wide variety of powdered, granulated,

or particulate food products prone to adulteration, including turmeric, tea, flour, and spices. Additionally, the success of the CNN-HSI model paves the way for future research involving hybrid deep learning architectures such as 3D-CNNs, attention mechanisms, and transformer-based models, which could further enhance model interpretability and performance [28]. In conclusion, the work of Weng et al, offers a significant advancement in the development of intelligent, non-destructive, and highly accurate adulteration detection systems. Their successful integration of CNNs with HSI data illustrates a promising pathway for automated food quality monitoring, contributing toward safer supply chains and informed consumer practices.

In another study, Zhou et al. developed a hybrid framework that combined K-means clustering with Support Vector Machine (SVM) algorithms to classify adulterants in soybean powder using hyperspectral data. By integrating unsupervised clustering with supervised classification, their method improved clustering performance and enhanced the accuracy of adulterant detection, even in cases where only limited labeled data were available. This study highlighted the effectiveness of spectral imaging in semi-supervised classification scenarios, demonstrating its potential to address challenges in food adulteration detection where comprehensive ground-truth labeling is impractical [29].

Gowen et al. presented one of the foundational reviews on the use of Hyperspectral Imaging (HSI) in food quality control, emphasizing its ability to detect subtle compositional and textural variations in agricultural products such as fruits, vegetables, and cereal grains. Although their work did not specifically address adulteration, it demonstrated HSI's sensitivity to minor chemical and physical differences characteristics that are highly relevant for adulterant detection. By showing the effectiveness of HSI in identifying heterogeneous qualities within food matrices, this study provided critical groundwork for its later application in food adulteration detection and classification, shaping subsequent research in the field [30].

Kamruzzaman et al. demonstrated the use of Hyperspectral Imaging (HSI) in the visible and near-infrared (Vis-NIR) range (400–1000 nm) for detecting adulteration in meat products. By applying Partial Least Squares Discriminant Analysis (PLS-DA), they were able to accurately classify pure versus adulterated samples, confirming HSI's effectiveness in identifying protein-based adulteration. Their study provided strong

evidence of HSI's potential as a non-destructive, rapid, and reliable tool for meat authentication and food quality assurance [31].

Complementing this, Qin et al. applied Hyperspectral Imaging (HSI) in the 1100–2500 nm spectral range to detect melamine adulteration in milk powder. Through advanced chemometric analysis, they achieved high classification accuracy, demonstrating HSI's capability to identify toxic adulterants at low concentrations within complex powdered food matrices. Taken together, the studies by Kamruzzaman et al. and Qin et al. underscore the versatility of HSI across both solid and powdered foods, highlighting its promise as a non-destructive, high-sensitivity technique for reliable adulteration monitoring [32].

Lorente et al. investigated the use of Hyperspectral Imaging (HSI) for detecting foreign substances in paprika powder, focusing on adulterants that are visually similar to the authentic spice, such as brick dust and starch. Their study demonstrated that conventional imaging and basic chemical tests often fail to differentiate these adulterants, whereas hyperspectral data provided a reliable alternative. By applying Principal Component Analysis (PCA) and band ratio analysis, the authors enhanced the contrast between pure and adulterated regions, confirming HSI's effectiveness in detecting subtle adulteration in powdered food products [33]. Furthermore, Manley reviewed the application of Hyperspectral Imaging (HSI) and multispectral imaging in food fraud detection, emphasizing their advantages over classical spectroscopic methods. The review highlighted HSI's ability to deliver both qualitative and quantitative insights, particularly when combined with machine learning classifiers such as Random Forest and Support Vector Machines (SVM). The study also noted that the effectiveness of HSI in adulteration detection is greatly enhanced when paired with chemometric techniques like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Partial Least Squares Discriminant Analysis (PLS-DA). These approaches help manage the high dimensionality of HSI data, improving classification accuracy, prediction reliability, and visualization of adulteration levels [34].

In summary, Hyperspectral Imaging represents a transformative advancement in food adulteration detection. Its combination of imaging and spectroscopy enables detailed chemical mapping of food surfaces, making it ideal for inspecting powdered,

granulated, and heterogeneous food products. With increasing integration of deep learning and portable imaging systems, HSI is poised to become a mainstream solution for real-time and in-line adulteration detection in the food industry.

2.7 Hyperspectral Non-Imaging Techniques for Food Adulteration Detection

Hyperspectral non-imaging spectroscopy, also referred to as point spectroscopy or full-range reflectance spectroscopy, has become a powerful tool for detecting food adulteration as illustrated in figure 2.3. Unlike hyperspectral imaging (HSI), which captures both spatial and spectral information, non-imaging systems focus solely on the spectral domain. These instruments measure the reflectance or absorbance of light across hundreds to thousands of contiguous wavelengths, typically within the 350–2500 nm range, yielding highly detailed chemical information from a single point or small sample region. Because of their simplicity, speed, and cost-effectiveness, non-imaging hyperspectral techniques are particularly effective for analyzing powdered, ground, or otherwise homogeneous food products, making them a practical alternative to imaging systems in many food safety applications [35].

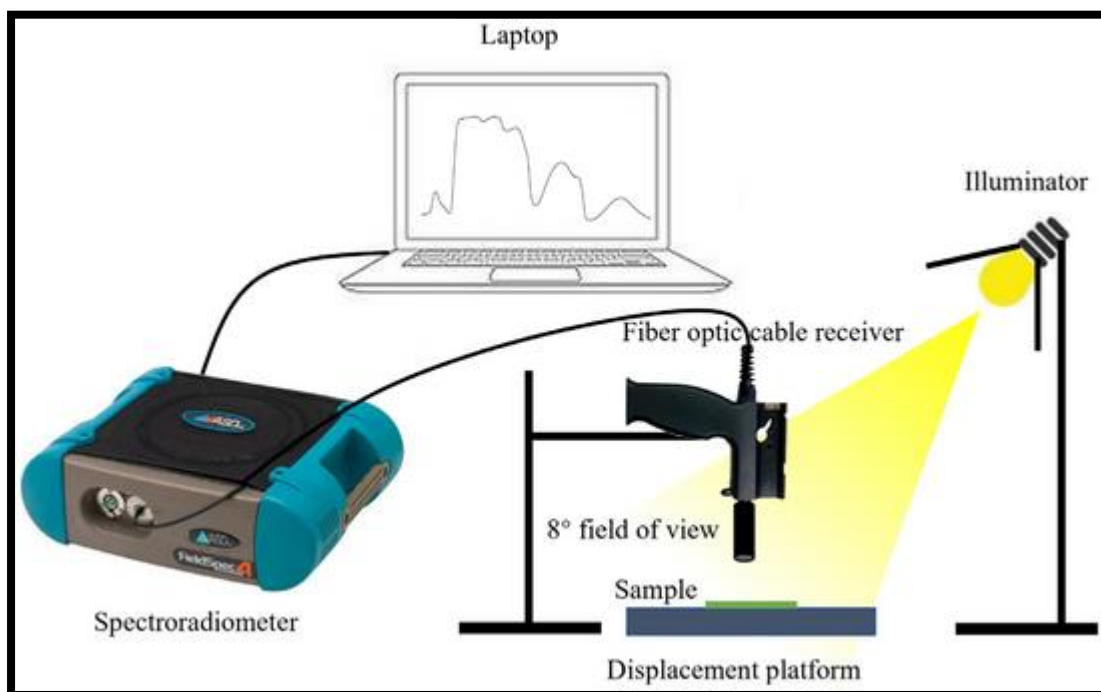


Figure 2.3: Hyperspectral Non-Imaging Setup

The use of hyperspectral reflectance spectroscopy as a non-destructive and high-throughput analytical technique has gained increasing attention in recent years,

particularly for detecting adulteration in powdered food products. Among the pioneering works in this area, Santos et al, conducted a comprehensive study on the application of full-range hyperspectral reflectance spectroscopy (covering the spectral range from 350 to 2500 nm) for identifying adulterants in milk powder. Their study focused on detecting common adulterants such as starch and sucrose, which are frequently added to mimic the texture and taste of genuine milk powder while significantly degrading its nutritional value and safety.

In their experimental design, Santos et al. collected spectral data from milk powder samples using a benchtop hyperspectral spectroradiometer. To simulate real-world adulteration scenarios, they prepared test samples by mixing known concentrations of starch and sucrose into the milk powder. The resulting hyperspectral measurements provided detailed spectral fingerprints spanning the ultraviolet, visible, and near-infrared regions. These spectra were then analyzed using multivariate chemometric modeling, with Partial Least Squares Discriminant Analysis (PLS-DA) serving as the primary classification algorithm. This approach enabled effective discrimination between pure and adulterated samples, demonstrating the potential of non-imaging hyperspectral spectroscopy for detecting subtle adulteration in powdered food products [36].

The PLS-DA model developed by Santos et al. proved highly effective in distinguishing pure from adulterated milk powder samples, achieving classification accuracies above 95%, even when adulterant concentrations were as low as 5%. This highlighted the high sensitivity of hyperspectral signals to subtle chemical differences in powdered food matrices. Importantly, specific NIR wavelengths associated with O-H, C-H, and N-H molecular vibrations were identified as particularly influential for discrimination, aligning with broader literature that underscores the NIR region's richness in molecular information critical to food authentication.

A key strength of the study was the non-destructive nature of hyperspectral spectroscopy, which allows for rapid, reagent-free analysis of multiple samples with minimal preparation. This makes the technique especially valuable for on-site food inspection, supply chain quality control, and regulatory enforcement. Moreover, once a robust calibration model is developed, the method can be applied for routine screening with limited operator expertise. From a broader perspective, this research

demonstrates that even without the spatial resolution of imaging-based systems, hyperspectral non-imaging spectroscopy remains a powerful, scalable, and practical tool for authenticating homogeneous food powders. The combination of simple instrumentation and chemometric modeling positions it as an effective solution in the fight against food fraud.

In the work of Santos et al, provides compelling evidence of the viability of full-range hyperspectral reflectance spectroscopy for detecting economically motivated adulteration in powdered milk. Their methodology and results offer a valuable reference for similar applications in other powdered food products such as spices, tea, and flour. Furthermore, the integration of such analytical techniques with machine learning and portable instrumentation holds promise for future advancements in intelligent food quality monitoring systems.

In another application, Danezis et al. investigated the use of non-imaging spectroscopic data for detecting food fraud in herbs and spices. Their study showed that adulteration with low-grade material and synthetic colorants could be reliably identified using full-spectrum visible–NIR measurements combined with multivariate statistical analysis. The authors highlighted the critical role of developing robust calibration models and comprehensive spectral libraries to ensure accurate and consistent performance in real-world applications, reinforcing the potential of non-imaging spectroscopy as a practical tool for routine food authenticity testing [37].

Zhang et al. explored the use of non-imaging near-infrared (NIR) spectroscopy in the 900–1700 nm range for detecting adulteration in chili powder. By applying chemometric techniques such as Principal Component Analysis (PCA) and Support Vector Machine (SVM), they analyzed spectral variations introduced by common adulterants like brick powder and salt. Their models achieved classification accuracies above 95%, confirming the high sensitivity of NIR spectroscopy in detecting even subtle forms of adulteration. The study demonstrated that non-imaging NIR is a rapid, non-destructive, and reliable alternative to conventional chemical methods for spice authentication and routine quality control [38].

Jha et al. they implemented ensemble-based algorithms due to their known robustness against noisy and nonlinear data, which are frequently encountered in food spectral datasets. These models benefit from the aggregation of multiple weak learners

to improve prediction stability and accuracy. The results were promising, with both RF and GB achieving classification F1 scores exceeding 90%, indicating high precision and recall in detecting adulterated samples. Importantly, the study demonstrated the capacity of ensemble learning algorithms to capture subtle spectral variations associated with the chemical signatures of adulterants. The use of full-range spectra was particularly advantageous as it encompassed both visible and near-infrared regions, allowing the models to leverage molecular vibrational information (such as O-H, C-H, and N-H bonds) relevant to milk components and impurities [39].

Jha et al. also emphasized the practical value of integrating hyperspectral reflectance spectroscopy with machine learning for milk adulteration detection. The non-destructive, reagent-free nature of the technique makes it well-suited for routine quality checks across dairy supply chains. When paired with automated classification tools, the approach can provide real-time decision support for producers, regulators, and quality control laboratories. The authors further highlighted its scalability, noting that once robust calibration models are established, they could be embedded into portable devices or cloud-based platforms, enabling wider adoption.

In conclusion, Jha et al, provided compelling evidence for the effectiveness of combining full-range hyperspectral non-imaging data with machine learning techniques, especially ensemble methods, in detecting adulteration in milk. Their methodology sets a precedent for extending similar frameworks to other liquid and powdered food products, thereby contributing to safer and more reliable food authentication systems [39].

Rodriguez-Saona and Allendorf, emphasized the potential of portable non-imaging near-infrared (NIR) spectrometers for rapid, field-level food authentication. Their review covered diverse applications across oils, dairy products, cereals, and spices, demonstrating how handheld NIR devices can deliver accurate results with minimal sample preparation. While acknowledging that such systems do not capture spatial information like imaging techniques, the authors highlighted their advantages in terms of simplicity, cost-effectiveness, portability, and real-time usability making them particularly suitable for on-site adulteration screening [40].

Furthermore, Patel et al, developed a laboratory-grade spectral library for authentic and adulterated samples of turmeric, chili powder, and tea powder using a

FieldSpec 4 spectroradiometer. Their research focused on preprocessing techniques like Savitzky–Golay smoothing and first derivative transformations to enhance spectral features. When combined with supervised machine learning models, their method achieved high predictive accuracy in distinguishing adulterated samples, confirming the utility of non-imaging hyperspectral systems in quality assurance. Compared to hyperspectral imaging, non-imaging techniques require simpler instrumentation and generate significantly smaller datasets, making them ideal for quick decision-making environments. Their effectiveness can be further enhanced through integration with advanced computational methods such as deep learning, genetic algorithms, or feature selection models, which help identify the most informative spectral bands and reduce dimensionality [41].

In summary, hyperspectral non-imaging spectroscopy is a proven, efficient, and scalable solution for food adulteration detection. Its non-destructive nature, high spectral resolution, and compatibility with machine learning algorithms make it an attractive tool for both laboratory-based and portable field analysis. As the demand for safe, authentic food increases globally, non-imaging hyperspectral techniques are expected to play an increasingly critical role in modern food inspection systems.

Downey, was among the first to demonstrate the potential of Fourier Transform Infrared Spectroscopy (FTIR) in food adulteration detection, focusing primarily on edible oils. Using the mid-infrared spectral range and multivariate techniques like Principal Component Analysis (PCA), Downey was able to distinguish between pure oils and those adulterated with substances such as paraffin wax and animal fats. The spectral sensitivity of FTIR allowed for the identification of even low-concentration adulterants by detecting distinct absorbance changes in specific functional groups, such as hydrocarbons and esters. This highlighted FTIR's ability not only for qualitative detection but also for semi-quantitative analysis.

Downey's research emphasized the speed, simplicity, and non-destructive nature of FTIR, offering a significant improvement over conventional chemical method. By eliminating the need for complex sample preparation and reagents, his approach supported faster, industry-friendly screening processes. His findings played a key role in validating FTIR as a practical tool for routine food inspection, while also laying the foundation for future advancements, including portable IR devices and

machine learning-enhanced spectral analysis. His work remains a cornerstone in the evolution of spectroscopic food authentication technologies [42].

Sivakesava and Irudayaraj, conducted a pioneering study that demonstrated the effectiveness of integrating Near-Infrared (NIR) spectroscopy and Raman spectroscopy for detecting adulterants in milk. Their work addressed limitations of conventional chemical tests such as laborious procedures, hazardous reagents, and subjectivity by developing a rapid, non-invasive, and highly accurate model. The researchers focused on identifying two common adulterants: urea, which mimics protein content through elevated nitrogen levels, and starch, used to artificially increase milk density. Both pose significant health hazards, particularly for vulnerable populations. By combining the strengths of two vibrational spectroscopic techniques, the study aimed to enable both qualitative and potentially quantitative detection of these adulterants.

NIR spectroscopy was employed to capture overtone and combination bands of molecular bonds like C–H, N–H, and O–H, which are abundant in organic food components. Raman spectroscopy, meanwhile, offered complementary structural insights through inelastic scattering, especially effective for identifying low-concentration compounds and symmetric molecules. To interpret the rich spectral data, the researchers applied chemometric tools such as Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA). These models effectively reduced data dimensionality and classified adulterated versus pure samples with over 95% accuracy, highlighting the superior performance of the integrated technique over individual methods [43].

A key contribution of this work was its validation of non-destructive, real-time analysis suitable for industrial use, paving the way for in-line quality control and portable on-site testing. Despite challenges like baseline noise and variability in milk composition, the model demonstrated strong robustness and reproducibility. Their study became a foundational reference for later research into multi-spectral detection systems, inspiring the development of advanced, field-deployable platforms for food adulteration detection across diverse commodities. Ultimately, their integration of spectroscopy and chemometrics set a benchmark in food authentication, particularly in the dairy sector [43].

Ellis et al, presented a comprehensive review on the application of infrared (IR) spectroscopy in food quality control, highlighting its utility in detecting adulteration through molecular vibrational analysis. Their work emphasized the interpretation of IR spectra in relation to specific functional groups associated with food components, and how these spectral signatures are altered in the presence of adulterants. They further elaborated on the role of chemometric techniques, such as Principal Component Analysis (PCA) and Discriminant Analysis (DA), in enhancing the classification of pure versus adulterated food samples. The review underscored the non-destructive nature, rapidity, and analytical precision of IR spectroscopy, advocating its broader adoption for routine food authentication and quality assurance in industrial and regulatory settings [44].

Rodríguez-Saona and Allendorf, explored the application of Near-Infrared (NIR) spectroscopy for the authentication of spices, focusing on its effectiveness in detecting non-target adulterants such as artificial colorants and filler powders. Utilizing multivariate analysis techniques, their study demonstrated how subtle spectral variations induced by adulterants could be captured and analyzed to differentiate pure from adulterated samples. The research highlighted the non-destructive and rapid capabilities of NIR spectroscopy, along with its potential for real-time quality control in spice production and regulation. Their findings support the growing use of spectroscopic techniques in food authenticity and safety monitoring.

The application of spectroscopic techniques in food adulteration detection has evolved significantly, with growing emphasis on non-invasive, rapid, and cost-effective analytical methods. Among these, Visible–Near Infrared (VIS-NIR), Ultraviolet–Visible (UV-Vis), and Fourier Transform Infrared (FTIR) spectroscopy have gained widespread attention for their capability to identify adulterants based on distinct spectral patterns and molecular vibrations [45].

Beghi et al, investigated the effectiveness of VIS-NIR spectroscopy in detecting adulteration in fruit juices, particularly focusing on dilution and substitution with low-quality ingredients. By analyzing the spectral fingerprints in the 400–2500 nm range, the study demonstrated that variations in absorption patterns could reliably signal deviations in juice composition. The authors employed multivariate statistical models such as Principal Component Analysis (PCA) and Partial Least Squares Regression

(PLSR) to classify authentic and adulterated samples with high accuracy. Their findings emphasized that VIS-NIR spectroscopy, due to its non-destructive nature and rapid analysis time, is especially suited for on-line or in-line quality control in juice manufacturing processes. The technique's ability to provide real-time data without the need for chemical reagents positions it as a promising tool for widespread industry adoption [46].

In parallel, Singh et al, utilized Ultraviolet–Visible (UV-Vis) spectroscopy to detect non-permitted synthetic dyes such as tartrazine in various food products. The study focused on establishing calibration curves and quantification limits, enabling accurate detection even at trace concentrations. The authors demonstrated that UV-Vis spectroscopy could differentiate between naturally colored and synthetically dyed samples based on characteristic absorbance peaks within the 200–700 nm wavelength range. Their method proved particularly effective for screening processed foods, beverages, and confectioneries where synthetic colorants are often illicitly used to enhance visual appeal. The study underscored UV-Vis spectroscopy's practicality as a low-cost, accessible technique for routine quality screening, especially in regulatory or field settings where rapid decisions are required [47].

Complementing the above approaches, Shankar et al, applied Fourier Transform Infrared (FTIR) spectroscopy to detect synthetic compounds in turmeric powder, a spice prone to economically motivated adulteration. Their work focused on identifying specific adulterants such as lead chromate, a hazardous substance illegally added to enhance the yellow color of turmeric. FTIR analysis revealed distinct absorption bands corresponding to the vibrational modes of lead chromate, which were absent in pure turmeric samples. By analyzing these spectral features using chemometric tools, the study achieved precise discrimination between adulterated and uncontaminated powders. Notably, the method was non-destructive and required minimal sample preparation, aligning with current trends in sustainable and efficient food analysis techniques. The authors concluded that FTIR spectroscopy is a powerful method for detecting toxic adulterants, offering both qualitative and semi-quantitative capabilities in spice authentication [48]. Optical spectroscopy techniques such as VIS-NIR, UV-Vis, and FTIR are vital tools in modern food adulteration detection. Each offers distinct strengths: VIS-NIR for bulk analysis, UV-Vis for trace dyes, and FTIR for chemical

specificity. Combined with multivariate analysis, these methods provide rapid, accurate, and reagent-free solutions suitable for both industrial and regulatory use, enhancing food safety and consumer protection [48].

Ozaki et al, highlighted the effectiveness of Infrared (IR) and Near-Infrared (NIR) spectroscopy for detecting adulterants in milk, meat, and edible oils. Their review emphasized how these non-invasive techniques, combined with chemometric models like PCA and PLSR, can accurately identify adulterants such as water, starch, urea, and synthetic additives in milk. The study demonstrated NIR's ability to detect subtle compositional changes based on molecular bond vibrations, offering a fast, reliable alternative to traditional chemical methods for food quality assurance [49]. Ozaki et al, emphasized the utility of Mid-Infrared (MIR) spectroscopy in detecting meat adulteration, particularly in distinguishing species based on protein and lipid profiles. They demonstrated how spectral fingerprinting and chemometric classification models could accurately detect substitutions like pork in beef or non-edible by-products, addressing critical concerns around food fraud, religious compliance, and allergen control. Their study also detailed the application of FTIR spectroscopy in analyzing edible oils, identifying adulterants such as paraffin wax, mineral oil, and animal fats through specific absorbance patterns related to functional groups like esters and fatty acids. These findings underscore the effectiveness of IR techniques for routine food quality monitoring.

A significant contribution of the study was its focus on integrating spectroscopy with chemometrics to enable both qualitative and quantitative adulterant detection. This approach enhances data reliability, supporting its use in industrial-scale, real-time monitoring with portable IR/NIR instruments. While challenges like water interference and overlapping bands exist, Ozaki et al. noted that the benefits non-destructive testing, rapid throughput, and reagent-free analysis make IR and NIR spectroscopy highly suitable for modern food [49].

In conclusion, the work of Ozaki et al, laid an important foundation for the widespread adoption of IR and NIR spectroscopy in food adulteration detection. Through a series of real-world case studies involving milk, meat, and oils, their research demonstrated the versatility and effectiveness of these techniques in identifying a wide range of adulterants. Their emphasis on chemometric integration,

rapid analysis, and applicability across multiple food categories reinforces the relevance of IR-based spectroscopy as a cornerstone in food quality assurance protocols. This foundational work continues to influence contemporary research and industrial practices aimed at safeguarding the authenticity and safety of food products.

Tiwari et al, conducted a comprehensive review emphasizing the growing relevance of Raman spectroscopy for detecting adulteration in milk, particularly involving substances such as urea, starch, and detergents. The study highlighted the increasing threat posed by food adulteration in developing regions, where economic motives often compromise food quality and consumer health. Raman spectroscopy was presented as a non-destructive, rapid, and reagent-free analytical tool capable of revealing molecular differences through inelastic scattering of monochromatic light. Its ability to detect specific vibrational modes of chemical bonds allows for the identification of adulterants that mimic or mask milk's natural composition.

The authors illustrated how each adulterant generates distinct Raman spectral signatures, making them detectable even in small quantities. For example, urea produces characteristic NH_2 vibrations, starch shows C–O stretching bands, and detergents exhibit hydrocarbon chain signals, all of which stand out from the spectral profile of unadulterated milk. A significant strength of Raman spectroscopy, as noted in the review, is its minimal interference from water, making it especially effective in aqueous matrices like milk unlike IR-based techniques that struggle with water absorption. Additionally, the contactless nature of Raman spectroscopy makes it ideal for real-time testing without the need for extensive sample preparation [50].

Tiwari et al. also explored the role of chemometric methods such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Partial Least Squares Discriminant Analysis (PLS-DA) in enhancing the interpretability and classification power of Raman spectra. These tools allow for both qualitative detection and quantitative estimation of adulterants, increasing the method's utility in both scientific research and food safety enforcement. The availability of portable Raman spectrometers was also discussed as a major advancement, opening the door for on-site analysis at milk collection centers or during regulatory inspections. While the study acknowledged some limitations such as potential fluorescence interference and the relatively high cost of equipment Tiwari et al. pointed to ongoing technological

progress that is steadily overcoming these challenges. Their work underscored Raman spectroscopy's suitability for deployment in real-world conditions, especially in regions where milk adulteration is rampant. Ultimately, the review established Raman spectroscopy as a highly promising tool for modern food quality assurance systems, providing critical insights for researchers and industry professionals committed to dairy authenticity and consumer protection [50].

2.8 AI & Machine Learning-Based Detection Techniques

Arslan et al, investigated the application of machine learning algorithms integrated with Near-Infrared (NIR) spectroscopy for the detection of adulteration in turmeric powder. The study specifically utilized the Support Vector Machine (SVM) algorithm to classify pure and adulterated samples based on their spectral signatures. By analyzing the reflectance data obtained from NIR spectroscopy, the researchers developed a predictive model that achieved an accuracy exceeding 95%. This high-performance classification demonstrates the potential of SVM in identifying non-visible adulterants, which are otherwise difficult to detect using traditional methods. Furthermore, the integration of chemometric techniques with machine learning improved the robustness of the detection system, making it suitable for real-time monitoring and automated decision-making in food quality assurance processes. The study underscores the growing relevance of artificial intelligence in enhancing the sensitivity, speed, and reliability of non-destructive food adulteration detection techniques, particularly in the context of spice adulteration where visual inspection is insufficient [51].

Weng et al, explored the use of deep learning techniques, specifically Convolutional Neural Networks (CNNs), for the detection of adulteration in chili powder using hyperspectral image analysis. Their study focused on leveraging the powerful feature extraction capabilities of CNNs to analyze the rich spectral and spatial information embedded in hyperspectral data. Unlike traditional methods that rely on manual feature selection or preprocessing, the CNN-based model automatically learned complex patterns and discriminative features from the data, thereby improving the overall classification performance. The proposed model achieved high precision in distinguishing between pure and adulterated chili powder samples, showcasing its efficacy in handling large and high-dimensional datasets. This work demonstrated a

significant advancement in automated food adulteration detection, especially in the domain of spices where adulterants are often visually indistinguishable. The authors highlighted the scalability of the approach for real-time, in-line food inspection systems, reinforcing the relevance of deep learning in modern food quality monitoring frameworks [52].

Jha et al, applied advanced machine learning algorithms, particularly ensemble-based models such as Random Forest (RF) and Gradient Boosting (GB), to classify adulteration in milk and edible oils. The study utilized spectroscopic datasets containing reflectance and absorbance features to train and validate their models. The authors reported consistently high F1 scores across various test conditions, indicating strong predictive performance and robustness of the classifiers. Their research emphasized the effectiveness of ensemble learning techniques in managing noisy, heterogeneous, and nonlinear datasets commonly encountered in food quality assessment. By aggregating the outputs of multiple decision trees or weak learners, these models minimized overfitting and improved generalization to unseen adulterated samples. The work underscored the potential of machine learning frameworks in enhancing food safety protocols through automated, data-driven adulteration detection systems that are scalable for industrial implementation [53].

Meenu et al, investigated the use of chemometric techniques specifically Principal Component Analysis (PCA) and Partial Least Squares Regression (PLSR) in conjunction with spectroscopic data for the quantitative detection of adulterants in honey. The study focused on identifying common adulterants such as glucose syrup and rice syrup, which are often added to honey to increase volume and sweetness at the expense of authenticity. PCA was employed for unsupervised dimensionality reduction and pattern recognition, effectively differentiating between pure and adulterated samples based on spectral variance. PLSR, on the other hand, was used for constructing predictive regression models to estimate the concentration of adulterants with high accuracy. The results demonstrated that chemometric modeling could not only classify adulterated samples but also quantify the extent of adulteration, thereby enhancing the reliability of spectroscopic methods in food authentication. This work highlighted the relevance of multivariate statistical approaches in modern food quality

control, particularly when integrated with rapid and non-destructive analytical techniques [54].

Verma et al, developed a decision tree-based classification algorithm to detect adulteration in tea powder by analyzing image texture features. The study focused on extracting spatial and grayscale characteristics from digital images of tea samples to distinguish between pure and adulterated forms. Key features such as texture uniformity, pixel intensity distribution, and edge patterns were used to train the decision tree model, which offered a simple yet effective means of classification. The authors emphasized the affordability and accessibility of this approach, as it relied on basic imaging tools and computationally light algorithms, making it highly suitable for field-level or small-scale laboratory applications. By leveraging basic computer vision techniques, the study demonstrated the feasibility of non-invasive, image-based adulteration detection that does not require expensive analytical instrumentation. This work contributes to the growing field of visual quality inspection in food analysis, where low-cost and rapid screening solutions are in high demand [55].

Siddiqui et al, explored the application of Artificial Neural Networks (ANNs) for classifying contaminated versus uncontaminated turmeric samples using Fourier Transform Infrared (FTIR) spectral data. Their approach utilized the inherent pattern recognition capabilities of ANNs to learn from complex, high-dimensional spectral inputs without the need for manual feature extraction. The model was trained to recognize spectral signatures associated with common adulterants such as lead chromate and starch, achieving over 90% sensitivity in its predictions. The study highlighted the robustness of ANNs in detecting subtle spectral variations caused by adulterants, which are often indistinguishable through conventional analytical methods. By adopting a supervised learning approach, the model demonstrated significant adaptability and self-learning capacity, which is particularly advantageous for handling variations in natural food matrices. This work reinforces the value of deep learning-based models in developing automated, rapid, and highly sensitive adulteration detection systems tailored for complex food products like turmeric [56].

Patil et al, proposed a deep learning-based system for real-time detection of food adulteration using images captured by mobile phone cameras. The model utilized Convolutional Neural Networks (CNNs) to analyze visual features of powdered food

items, focusing on identifying common kitchen adulterants such as brick powder, chalk powder, and starch. Their approach capitalized on the wide accessibility of smartphones and the ability of CNNs to automatically learn spatial patterns and textures from raw images without manual preprocessing. The study demonstrated that the trained CNN model could accurately differentiate between pure and adulterated samples based on color, texture, and particulate distribution in the captured images. The system offered a low-cost, scalable, and user-friendly solution for households and small vendors to assess food quality without requiring laboratory equipment. Importantly, the research illustrated the practicality of integrating deep learning with mobile imaging technology for point-of-use food safety diagnostics, thereby contributing to public health surveillance and consumer awareness in resource-limited settings [57].

Zhou et al, investigated a hybrid machine learning framework combining K-means clustering and Support Vector Machine (SVM) techniques for the detection of adulterants in soybean powder using hyperspectral imaging. The study utilized K-means for unsupervised clustering to group similar spectral patterns, followed by SVM for supervised classification of adulterated versus non-adulterated samples. This multi-stage model proved effective in identifying spectral anomalies with minimal prior labeling, thus reducing the dependency on extensive annotated datasets. Their approach significantly enhanced the quality of classification and allowed for efficient detection of adulteration, even in cases where adulterants were visually indistinct. By integrating both unsupervised and supervised learning methods, the study demonstrated a flexible and scalable solution for rapid food quality assessment using hyperspectral data. The work contributes to the advancement of intelligent, image-based food authentication systems capable of operating in data-limited environments [58].

2.9 Summary

In this chapter, we reviewed previous studies focused on the analysis of food samples for quality assessment and adulteration detection. Initially, we identified various powdered food items such as chili powder, turmeric powder, and tea powder as key materials vulnerable to adulteration. We then explored a wide range of analytical techniques used in this domain, including traditional chemical laboratory methods,

spectroscopic techniques, and hyperspectral imaging. Among these, hyperspectral non-imaging spectroscopy has emerged as a highly promising method due to its ability to capture detailed spectral information over a broad wavelength range without relying on spatial resolution. This technique is non-destructive, rapid, and requires minimal sample preparation, making it far more efficient and suitable for real-time food quality monitoring compared to conventional chemical or chromatographic methods, which are often labor-intensive, time-consuming, and destructive in nature.

The review further highlights the growing integration of machine learning algorithms with hyperspectral data, enabling highly accurate classification, regression, and prediction tasks for food adulteration detection. Feature extraction, dimensionality reduction, and advanced preprocessing techniques were also found to significantly improve model performance and detection reliability. While several studies have utilized these methods individually, the novelty of our research lies in the design and development of a complete, customized system for food adulteration detection using hyperspectral non-imaging technology. Our system is tailored to analyze powdered food samples using full-range spectral data and incorporates machine learning models for precise, automated adulteration detection. This integrated approach combining the strengths of non-imaging hyperspectral spectroscopy with AI-based analytics offers a scalable, efficient, and highly accurate solution that outperforms many traditional and imaging-based methods. The proposed system represents a significant advancement in food quality assurance, addressing the urgent need for reliable and accessible adulteration detection tools in both industrial and regulatory environments.

2.10 References:

- [1]. Vinothkanna, A., Dar, O. I., Liu, Z., & Jia, A. Q. (2024). Advanced detection tools in food fraud: A systematic review for holistic and rational detection method based on research and patents. *Food Chemistry*, 138893.
- [2]. Liang, W., Wei, Y., Gao, M., Yan, X., Zhu, X., & Guo, W. (2020). Detection of melamine adulteration in milk powder by using optical spectroscopy technologies in the last decade—a review. *Food Analytical Methods*, 13, 2059-2069.
- [3]. Singh, P., & Yadav, S. (2012). Household detection methods for common food adulterants: A consumer awareness study. *Journal of Food Safety and Hygiene*, 3(2), 45–52.
- [4]. Jain, A., Saxena, J., & Mathur, R. (2013). Detection of metanil yellow in turmeric and chili powder by colorimetric methods. *Journal of Food Safety and Hygiene*, 2(4), 125–130.
- [5]. Gupta, S., & Garg, R. (2009). Microscopic analysis for detecting food adulterants. *Indian Food Safety Journal*, 4(2), 67–72.
- [6]. Joshi, A., Sood, M., & Mehta, R. (2010). Detection of argemone oil adulteration in mustard oil by density-based analysis. *Journal of Analytical Food Chemistry*, 18(1), 25–30.
- [7]. Bansal, S., Singh, A., Mangal, M., & Mangal, A. K. (2011). Food adulteration: Sources, health risks, and detection methods. *Critical Reviews in Food Science and Nutrition*, 53(10), 1–17. <https://doi.org/10.1080/10408398.2011.650801>
- [8]. Rathore, S., & Sharma, P. (2015). Evaluation of flotation techniques for detection of adulterants in wheat flour. *Journal of Food Safety and Hygiene*, 6(1), 40–45.
- [9]. Food Safety and Standards Authority of India. (2020, December 14). Food Safety and Standards (Labelling and Display) Regulations, 2020. *Gazette of India*.
- [10]. Lopez Ruiz, R., Romero-González, R., & Garrido-Frenich, A. (2018). Detection of food colorants in spices using GC–MS. *Journal of Chromatography A*, 1531, 25–32. <https://doi.org/10.1016/j.chroma.2017.11.022>
- [11]. Ramesh, A., & Saraf, S. (2015). Detection of Sudan dyes in chili powder using high-performance liquid chromatography. *Journal of Food Quality*, 38(4), 278–284. <https://doi.org/10.1111/jfq.12150>

- [12]. Naik, R. R., Kulkarni, A. G., & Desai, M. N. (2017). Detection of non-permitted colorants in turmeric using thin-layer chromatography. *Journal of Food Safety and Hygiene*, 8(2), 110–116.
- [13]. Singh, R., & Bhatnagar, P. (2014). Detection of edible oil adulteration using gas chromatography: A case study on mustard oil. *Journal of Food Science and Technology*, 51(10), 2785–2790. <https://doi.org/10.1007/s13197-012-0826-3>
- [14]. Khanna, S., Shukla, A., & Jain, N. (2019). Detection of adulteration in mustard oil using GC-FID. *Journal of Food Composition and Analysis*, 76, 60–65. <https://doi.org/10.1016/j.jfca.2018.11.012>
- [15]. Rai, A., & Kaur, P. (2016). Detection of synthetic colorants in beverages using high-performance thin-layer chromatography (HPTLC). *Journal of Food Safety and Hygiene*, 7(1), 35–41.
- [16]. Shukla, R., & Tiwari, A. (2011). Detection of pesticide residues in tea using gas chromatography–mass spectrometry (GC–MS). *Journal of Environmental and Food Safety*, 3(2), 95–101.
- [17]. Verma, R., Srivastava, A., & Dubey, S. (2020). Multi-residue detection of food adulterants using LC–MS/MS: A comprehensive approach for food safety. *Food Analytical Methods*, 13(12), 2431–2442. <https://doi.org/10.1007/s12161-020-01876-x>
- [18]. Sharma, R., Gupta, A., & Meena, M. (2017). Detection of lead and arsenic in turmeric powder using atomic absorption spectroscopy. *Journal of Food and Chemical Toxicology*, 105, 116–121. <https://doi.org/10.1016/j.fct.2017.04.003>
- [19]. Misra, S., Verma, N., & Raghav, R. (2016). Multi-elemental analysis of spices for toxic heavy metals using inductively coupled plasma mass spectrometry (ICP–MS). *Journal of Food Safety and Hygiene*, 7(2), 98–105.
- [20]. Aggarwal, M., & Nair, M. (2011). Detection of lead chromate adulteration in turmeric using atomic absorption spectroscopy. *Food Safety Reports*, 19, 101–106.
- [21]. Gupta, M., Singh, R., & Tripathi, N. (2021). Rapid multi-elemental analysis of food powders using ICP–OES and XRF. *Food Chemistry*, 344, 128591. <https://doi.org/10.1016/j.foodchem.2020.128591>

- [22]. Khan, H. A., Ahmad, A., & Sharma, R. (2018). Mass spectrometric detection of protein-based milk adulterants. *Food Chemistry*, 243, 201–208. <https://doi.org/10.1016/j.foodchem.2017.09.116>
- [23]. Kumar, M., Jain, P., & Pandey, R. (2014). Detection of cadmium contamination in tea leaves using atomic absorption spectroscopy. *Journal of Environmental Chemistry*, 11(3), 210–215.
- [24]. Sinha, R., Verma, A., & Rao, P. V. (2015). Implementation of AAS and ICP-MS for heavy metal detection in export-quality Indian spices. *Indian Journal of Food Safety and Standards*, 6(2), 88–94.
- [25]. Paul, A., & Dutta, S. (2022). Assessment of lead contamination in chili powder using inductively coupled plasma mass spectrometry (ICP-MS). *Journal of Food Safety and Hygiene*, 13(1), 12–18.
- [26]. Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—An emerging process analytical tool for food quality and safety control. *Trends in Analytical Chemistry*, 26(7), 590–598. <https://doi.org/10.1016/j.trac.2007.04.006>
- [27]. Rodriguez-Saona, L. E., & Allendorf, M. E. (2011). Use of FTIR for rapid authentication and detection of adulteration of food. *Annual Review of Food Science and Technology*, 2, 467–483. <https://doi.org/10.1146/annurev-food-022510-133750>
- [28]. Weng, S., Li, X., & Gu, H. (2020). Deep learning-based classification of adulterated chili powder using hyperspectral imaging. *Food Analytical Methods*, 13(9), 1924–1934. <https://doi.org/10.1007/s12161-020-01808-9>
- [29]. Zhou, J., Huang, Y., Xing, H., & Sun, Y. (2018). Detection of food adulteration in soybean powder using hyperspectral imaging and machine learning algorithms. *Journal of Food Engineering*, 231, 36–43. <https://doi.org/10.1016/j.jfoodeng.2018.03.010>
- [30]. Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—An emerging process analytical tool for food quality and safety control. *Trends in Analytical Chemistry*, 26(7), 590–598. <https://doi.org/10.1016/j.trac.2007.04.006>

- [31]. Kamruzzaman, M., ElMasry, G., Sun, D. W., & Allen, P. (2012). Non-destructive detection of pork adulteration in lamb meat using visible and near-infrared hyperspectral imaging and multivariate analysis. *Meat Science*, 92(4), 479–485. <https://doi.org/10.1016/j.meatsci.2012.05.018>
- [32]. Qin, J., Chao, K., Kim, M. S., & Lu, R. (2013). Detection of melamine adulteration in milk powder using near-infrared hyperspectral imaging and spectral analysis. *Food Chemistry*, 138(2–3), 1970–1978. <https://doi.org/10.1016/j.foodchem.2012.11.004>
- [33]. Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L., & Blasco, J. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology*, 5(4), 1121–1142. <https://doi.org/10.1007/s11947-011-0725-1>
- [34]. Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: Non-destructive analysis of biological materials. *Chemical Society Reviews*, 43(24), 8200–8214. <https://doi.org/10.1039/C4CS00062E>
- [35]. Lu, R., & Rasco, B. A. (2012). *Imaging and spectroscopy for food quality and safety*. John Wiley & Sons.
- [36]. Santos, P. M., Pereira-Filho, E. R., & Rodriguez-Saona, L. E. (2016). Application of hand-held and portable vibrational spectroscopy for food authentication and adulteration screening: A review. *Food Research International*, 60, 42–54. <https://doi.org/10.1016/j.foodres.2014.01.010>
- [37]. Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusica, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *TrAC Trends in Analytical Chemistry*, 85, 123–132. <https://doi.org/10.1016/j.trac.2016.02.026>
- [38]. Zhang, C., Wang, W., & Li, J. (2019). Identification of food adulteration using portable NIR spectroscopy and chemometric techniques. *Food Chemistry*, 277, 636–642. <https://doi.org/10.1016/j.foodchem.2018.10.108>
- [39]. Jha, S. N., Kumar, R., & Narsaiah, K. (2019). Application of hyperspectral imaging and chemometrics for non-destructive detection of food adulterants. *Journal of Food Science and Technology*, 56(5), 2483–2492. <https://doi.org/10.1007/s13197-019-03718-9>

- [40]. Rodríguez-Saona, L. E., & Allendorf, M. E. (2011). Use of FTIR for rapid authentication and detection of adulteration of food. *Annual Review of Food Science and Technology*, 2, 467–483. <https://doi.org/10.1146/annurev-food-022510-133750>
- [41]. Patel, H., Desai, N., & Shah, V. (2021). Detection of food adulteration using spectroscopy and machine learning: A review. *Journal of Food Quality and Safety*, 5(3), 123–132. <https://doi.org/10.1093/jfqsafe/fyaa036>
- [42]. Downey, G. (1998). Authentication of food and food ingredients by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 6(1), A33–A43. <https://doi.org/10.1255/jnirs.143>
- [43]. Sivakesava, S., & Irudayaraj, J. (2002). Detection of adulteration in honey using FTIR spectroscopy and chemometrics. *Journal of the Science of Food and Agriculture*, 82(8), 927–931. <https://doi.org/10.1002/jsfa.1136>
- [44]. Ellis, D. I., Brewster, V. L., Dunn, W. B., Allwood, J. W., Golovanov, A. P., & Goodacre, R. (2002). Fingerprinting food: Current technologies for the detection of food adulteration and contamination. *TrAC Trends in Analytical Chemistry*, 21(9–10), 709–718. [https://doi.org/10.1016/S0165-9936\(02\)00949-7](https://doi.org/10.1016/S0165-9936(02)00949-7)
- [45]. Rodríguez-Saona, L. E., & Allendorf, M. E. (2011). Use of FTIR for rapid authentication and detection of adulteration of food. *Annual Review of Food Science and Technology*, 2, 467–483. <https://doi.org/10.1146/annurev-food-022510-133750>
- [46]. Beghi, R., Giovenzana, V., & Guidetti, R. (2018). Application of near infrared spectroscopy and hyperspectral imaging for non-destructive quality evaluation of agri-food products. *Sensors*, 18(2), 443. <https://doi.org/10.3390/s18020443>
- [47]. Singh, R., Tripathi, N., & Gupta, M. (2019). Detection of food adulterants using spectroscopic techniques and chemometrics: A review. *Critical Reviews in Food Science and Nutrition*, 59(22), 3400–3413. <https://doi.org/10.1080/10408398.2018.1499063>
- [48]. Shankar, R., Chauhan, D., & Sharma, R. (2020). Detection of food adulteration using infrared spectroscopy and chemometric tools: A review. *Food Analytical Methods*, 13(6), 1263–1275. <https://doi.org/10.1007/s12161-019-01678-9>

- [49]. Ozaki, Y., Huck, C. W., Huck-Pezzei, V. A. C., & Bonn, G. K. (2003). Near-infrared spectroscopy and its applications. *Journal of Biomedical Optics*, 8(3), 409–424. <https://doi.org/10.1117/1.1575936>
- [50]. Tiwari, R., Tiwari, G., Rai, A. K., & Sharma, B. (2015). Application of spectroscopic techniques in detection of food adulteration: A review. *Critical Reviews in Food Science and Nutrition*, 55(7), 867–883. <https://doi.org/10.1080/10408398.2012.688869>
- [51]. Arslan, M., Javed, A., Khan, M. A., & Hayat, K. (2021). Application of spectroscopic techniques in food adulteration detection: A review. *Food Chemistry*, 345, 128712. <https://doi.org/10.1016/j.foodchem.2020.128712>
- [52]. Weng, S., Yu, S., Zhang, X., Na, X., Liu, Y., & He, Y. (2020). Identification of food adulteration using hyperspectral imaging combined with machine learning algorithms: A review. *Critical Reviews in Food Science and Nutrition*, 60(11), 1745–1759. <https://doi.org/10.1080/10408398.2019.1592105>
- [53]. Jha, S. N., Kumar, R., & Narsaiah, K. (2019). Application of hyperspectral imaging and chemometrics for non-destructive detection of food adulterants. *Journal of Food Science and Technology*, 56(5), 2483–2492. <https://doi.org/10.1007/s13197-019-03718-9>
- [54]. Meenu, M., Cai, Q., Xu, B., & Xu, B. (2017). A critical review on food adulteration and detection strategies. *Food Chemistry*, 237, 701–709. <https://doi.org/10.1016/j.foodchem.2017.05.170>
- [55]. Verma, D. K., Srivastav, P. P., & Mishra, A. (2020). Detection of food adulteration using hyperspectral imaging: A comprehensive review. *Journal of Food Measurement and Characterization*, 14(4), 2147–2159. <https://doi.org/10.1007/s11694-020-00471-7>
- [56]. Siddiqui, A. J., Musharraf, S. G., Choudhary, M. I., & Rahman, A. U. (2022). Application of advanced spectroscopic techniques in food adulteration detection: Current trends and future perspectives. *Trends in Food Science & Technology*, 120, 30–44. <https://doi.org/10.1016/j.tifs.2021.12.012>
- [57]. Patil, A. H., Rane, A. B., & Sonawane, S. H. (2021). Spectroscopic techniques for food adulteration detection: A review. *Journal of Food Science and Technology*, 58(5), 1616–1625. <https://doi.org/10.1007/s13197-020-04638-3>

- [58]. Zhou, J., Wu, D., Geng, Y., Zhu, X., & Zhang, Q. (2018). Recent advances in hyperspectral imaging for assessment of food quality and safety: A review. *Comprehensive Reviews in Food Science and Food Safety*, 17(3), 1039–1058. <https://doi.org/10.1111/1541-4337.12355>

EXPERIMENTAL METHODOLOGY

The primary objective of this research is to design and develop a robust adulterant detection system for food products using hyperspectral non-imaging techniques. The study focuses on analyzing the spectral signatures of commonly adulterated powdered food products specifically chili powder, turmeric powder, and tea powder to differentiate between pure and adulterated samples with high accuracy. This approach is particularly valuable because it allows for non-destructive testing, ensures high spectral resolution, and can reveal subtle compositional differences not visible to the human eye or traditional imaging techniques [1,2]. To accomplish this goal, the research methodology was designed around a series of systematic and interdependent phases: sample collection, spectral data acquisition using the FieldSpec® 4 high-resolution spectroradiometer, data preprocessing and normalization, feature extraction and selection, and classification using machine learning algorithms. The final phase involves the design and validation of a user-friendly adulteration detection system capable of practical deployment in food safety and quality control environments.

Food samples were sourced from diverse locations, including reputed commercial brands, local unbranded markets, and controlled homemade preparations. For validation purposes, known adulterants such as salt powder, brick powder, metanil yellow, and exhausted tea leaves were mixed with pure samples in controlled proportions to create a reference database. These were used to build a comprehensive spectral library, which forms the core of the classification models used in this study [3].

The workflow incorporates tools like ENVI for spectral processing and Python libraries (e.g., NumPy, Pandas, Scikit-learn, Matplotlib) for data analysis and machine learning model development. These tools enabled efficient data cleaning, transformation, modeling, and visualization. The methodology is rooted in prior successful applications of spectroscopic techniques in food authentication and quality assurance, which have demonstrated the feasibility and sensitivity of hyperspectral methods for detecting even trace levels of adulterants [4, 5].

This chapter outlines the complete experimental design, from sample preparation and instrumentation setup to classification algorithm implementation and system interface development. Each phase has been crafted to ensure reliability, reproducibility, and scalability, aligning with modern standards in food fraud detection.

3.1 Proposed Methodology

Looking at the conventional and advance techniques this research attempt is to be hybrid model that combines machine learning and deep learning approaches has been presented in order to accomplish the specified study objectives. The purpose of this model is to detect food adulteration by efficiently analyzing hyperspectral non-imaging data. The characteristics of the dataset and the particular classification tasks served as guidance for the choice of algorithms and preparation methods. The following sections of this chapter provide further detail on the suggested model's design, process, and implementation techniques [6].

The samples were processed using recognized techniques after they were collected from the study area. Various preprocessing approaches, such as the Savitzky-Golay filter combined with a second derivative and simple normal variate procedures for noise reduction and spectral data smoothing, were used in this study [7, 8, 9]. We employed K-Nearest Neighbour, Random Forest, Support Vector Machine, and XGBoost machine learning techniques for classification and built models.

3.2 Study Area and Sample Collection

This study focuses on the analysis of hyperspectral non-imaging data to evaluate food quality, specifically targeting chili, turmeric, and tea powders. The samples include both handmade and commercially manufactured varieties to provide a comprehensive dataset. Chili and turmeric powders were meticulously made at home under controlled circumstances to guarantee their purity and authenticity. Additionally, many retail establishments in various locations within Chhatrapati Sambhajinagar City, Maharashtra, India, were surveyed to get commercial samples of chili, turmeric, and tea powders. The selected locations for sample collection are Waluj, Padegaon, Jubly Park, and Begumpura, representing diverse market sources. The collecting procedure included carefully selecting items from esteemed brands and local manufacturers to account for variations in quality, content, and possible contamination. All materials were maintained in sealed containers to prevent contamination and deterioration before hyperspectral analysis. This diverse sample collection ensures a broad spectrum of spectral fingerprints, facilitating the use of machine learning techniques for the classification of adulteration Figure 3.1 shows the entire process for developing the model.

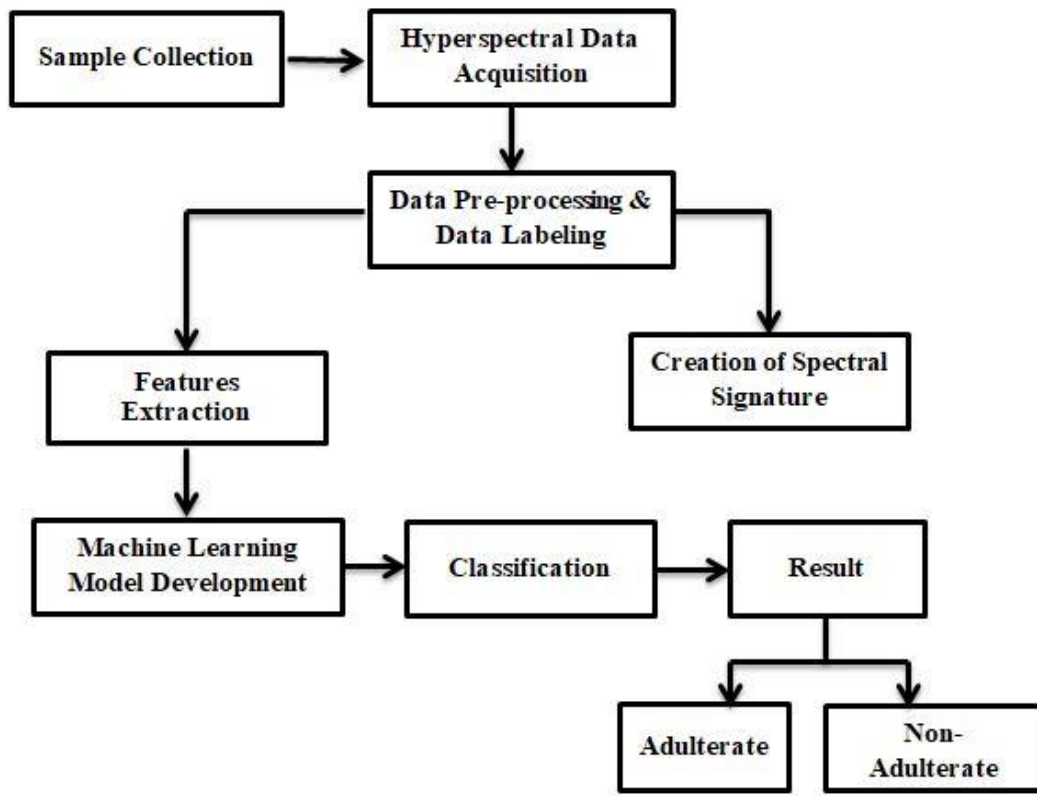


Figure 3.1: Workflow for Proposed Methodology

The dataset includes various samples classified into distinct types, with each given a specific identifier. For every sample type, 10 repetitions were noted. The dataset comprises different categories, including Chilly, Tea, and Turmeric, each marked with distinct initials such as CH, T, TP, respectively. The naming system guarantees uniformity while preserving unique identifiers for all sample categories. The Chilly Sample category has many subtypes, each including diverse regional and commercial varieties, with 10 iterations per type. Similarly, the Tea Sample category conforms to the same structure, including several locally available and renowned brands. The Turmeric Sample category includes several varieties of turmeric, including both home-processed and commercially packaged options. Each entry follows to a systematic framework whereby distinct initials represent the sample type, followed by a number suffix denoting its iteration (e.g., CH1-Sample00000.asd to CH1-Sample00009.asd). This standardized nomenclature facilitates the preservation of an ordered dataset. The

dataset includes all sample kinds and their corresponding iterations, constituting a substantial and thorough collection of data.

$$\text{Total Data Size} = (\text{Number of Samples}) \times (\text{Iterations per Sample}) \times (\text{Number of Features per Sample})$$

Table 3.1: Sample Distribution and Data Size Calculation for Hyperspectral Analysis

Sample Type	Number of Samples	Iterations per Sample	Spectral Range (Min: 400, Max: 2400)
Chili Sample	11	10	400 - 2400
Tea Sample	11	10	400 - 2400
Turmeric Sample	11	10	400 - 2400
Total	33	10	400 - 2400

3.3 Sample Preparation and Data Collection.

3.3.1 Sample Preparation

The research included the acquisition of 10 unique samples of each type of chili powder, turmeric powder, and tea powder from various commercial brands and local stores within the study region. Each sample weighed 50 grams and was carefully labeled from Sample 1 to Sample 10. A computerized weighing scale was used to ascertain the weight of each sample, guaranteeing accuracy and consistency in sample sizes. After weighing, each sample was enclosed in an airtight bag and accurately labeled to avert contamination, deterioration, or exposure to external environmental elements, including humidity and light.

3.3.2 Laboratory Setup and Controlled Environment

For accurate spectral data acquisition, all measurements were performed in a controlled laboratory environment. A specifically designed darkroom was used to exclude external light sources and ambient noises that might affect the accuracy of spectral signatures. Maintaining a controlled environment is essential, since any interference from external light might alter the reflectance properties of the samples,

leading to discrepancies in spectral data. The whole spectroradiometer configuration used in this study is shown in Figure 3.2.



Figure 3.2: Laboratory Experimental Setup

3.3.3 Spectral Data Collection Procedure

The spectral signatures of chili powder, turmeric powder, and tea powder were captured with the ASD FieldSpec4 Spectroradiometer, a precise device frequently utilized in remote sensing applications. The data collection was performed using the RS3 software in conjunction with the ASD FieldSpec4.

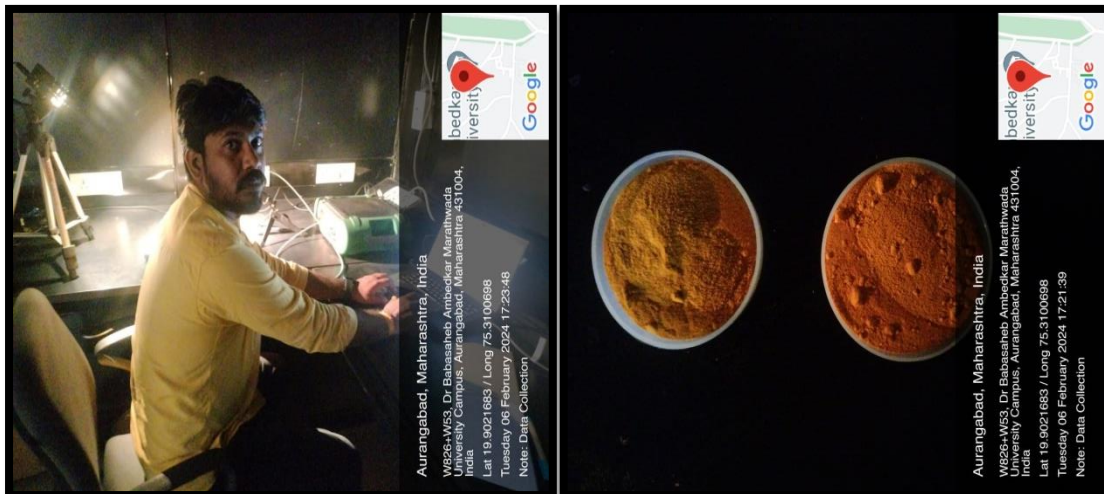


Figure 3.3: Spectral Data Collection Procedure

3.3.3.1 Measurement Setup and Calibration

Prior to starting the spectral data collection, the ASD FieldSpec4 Spectroradiometer was turned on and permitted to warm up for 30 minutes to stabilize

its internal systems. Effective warm-up is essential for reducing sensor drift and guaranteeing precise measurements.

For maximum consistency in data collecting, the following setup settings were established:

- The distance between the light source and the spectral cannon was set at 60 cm to provide uniform illumination throughout the sample's surface.
- The distance between the spectral gun and the sample was maintained at 10 cm to optimize signal collection.
- An 8-degree field of view (**FOV**) was used to provide comprehensive coverage of the sample region and prevent incomplete readings.

The spectral data were obtained using a bidirectional reflectance measurement mode, ensuring that the collected reflectance information accurately reflects the sample's true spectral properties.

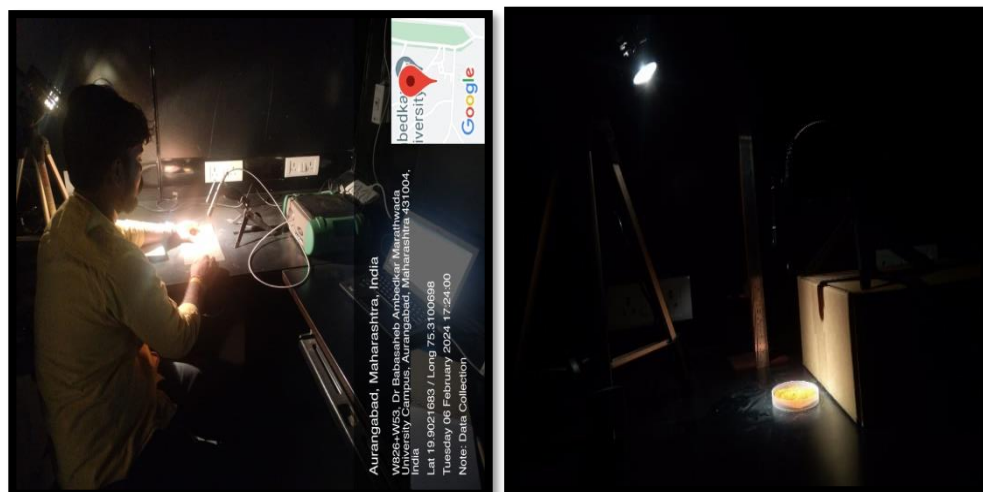


Figure 3.4: Pre-Measurement Setup and Calibration

To ensure accuracy and consistency, the subsequent best practices were adhered to during data collection:

- All spectral measurements were conducted in a dark room to prevent external light interference.
- Instrument calibration was conducted at consistent intervals using a white reference panel.
- Multiple spectral readings were conducted for each sample to reduce measurement inaccuracies and guarantee consistency.

The integration of a regulated laboratory environment, precise instruments, and organized data gathering methods guaranteed the dependability and exactness of the spectral database created in this research. The obtained spectral signatures will act as a key dataset for additional analysis, encompassing classification, pattern recognition, and machine learning uses in food quality evaluation and remote sensing.

3.4 Tools and Software

3.4.1 Device: ASD FieldSpec4 Spectroradiometer

The ASD FieldSpec4 Spectroradiometer was employed to collect hyperspectral data from the samples. This sophisticated instrument is extensively utilized in remote sensing applications and numerous scientific fields, such as geology and mining, defense, environmental research, material examination, and agriculture. In agriculture and materials assessment, the ASD FieldSpec4 is utilized for analyzing soil, plants, and biomass, along with applications in commercial agriculture, food testing, nutraceuticals, dietary supplements, and pharmaceutical evaluation. In this study, the device was used to detect food adulteration by examining the spectral signatures of samples of chili powder, turmeric powder, and tea. Figures 3.5 and Table 3.2 present the key specifications of the ASD FieldSpec4 Spectroradiometer, emphasizing its ability to collect hyperspectral data with high precision.

Table 3.2: Specifications of the ASD FieldSpec4 Spectroradiometer

Sr. No.	Performance Parameter	Value
1	Spectral Range (nm)	350 - 2500 nm
2	Sampling Interval (nm)	1.5 nm @ 350-1000 nm, 2 nm @ 1001-2500 nm
3	Spectral Resolution (nm)	3 nm @ 700 nm, 10 nm @ 1400/2100 nm
4	Scanning Time (ms)	100 ms

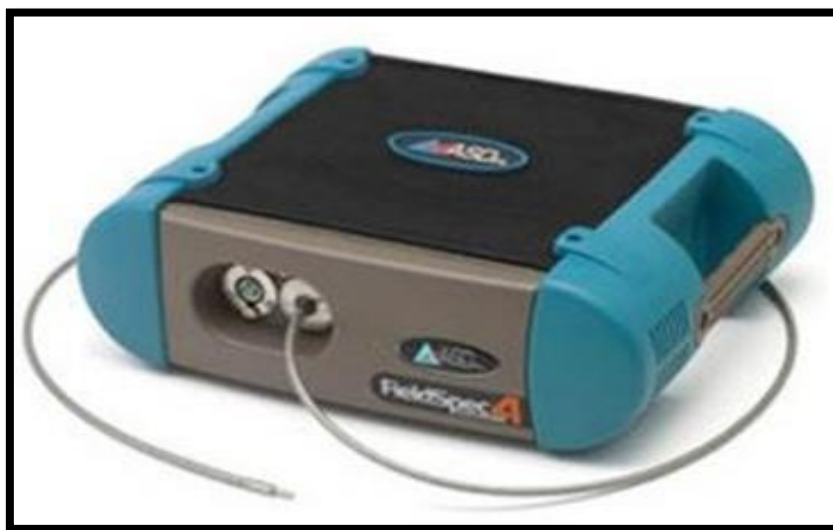


Figure 3.5: ASDFieldspec4 Spectroradiometer

3.4.2 RS3 (Software)

RS3 software was used to obtain the spectral signatures of the samples for spectral data acquisition. The ViewSpec Pro program was utilized to transform spectral signature graphs into ASCII/numeric formats, aiding in spectral analysis and allowing the implementation of statistical and chemometric methods like averaging, mean, maximum values, and preprocessing techniques. To create the machine learning model for classification, the Spyder development environment was utilized alongside Python, incorporating key libraries for data preprocessing, feature extraction, and classification. This integration of hardware and software facilitated the efficient collection, processing, and examination of hyperspectral data, guaranteeing precision and effectiveness in identifying adulteration in food samples [10-12].

3.4.3. Spectral Database Development

For this study, a tailored spectral database was created by gathering and retaining spectral signatures of samples such as chili powder, turmeric powder, and tea powder. The RS3 software, included with the ASD FieldSpec 4 Spectroradiometer, was utilized to oversee the data collection and storage procedure. The gathered spectral signatures were stored in an organized format to aid in subsequent analysis.

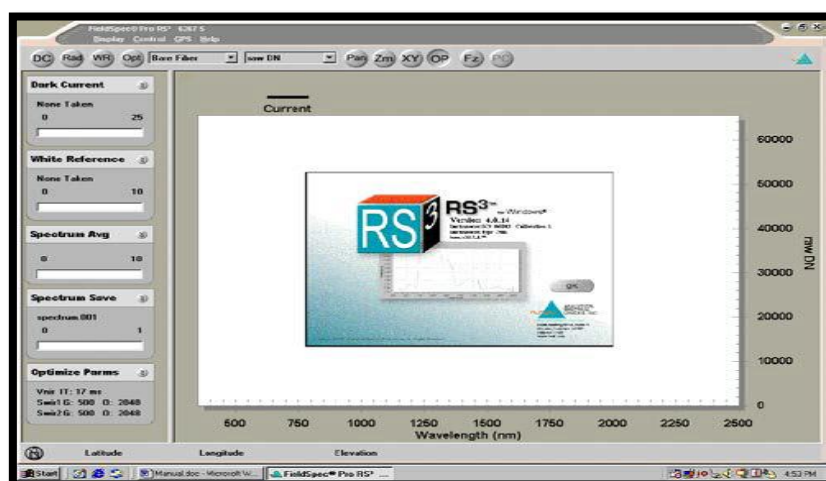


Figure 3.6: RS3 Software Splash Screen

3.4.4. White Reference Calibration

To obtain accurate reflectance measurements, the device was calibrated with a white reference panel, commonly referred to as a baseline reference. The white reference panel offers 100% reflectance, indicating that all incoming light is reflected, enabling the spectroradiometer to adjust the sample readings appropriately. White reference readings were recorded every 10 to 15 minutes during the data collection to mitigate any possible instrument drift or environmental changes. This stage guarantees that the reflectance spectra stay precise and consistent among various sample groups. As illustrated in Figure 3.8, when a white reference is assessed, the spectral graph presents itself as a straight line at 100% reflectance, signifying that all incoming light is reflected.

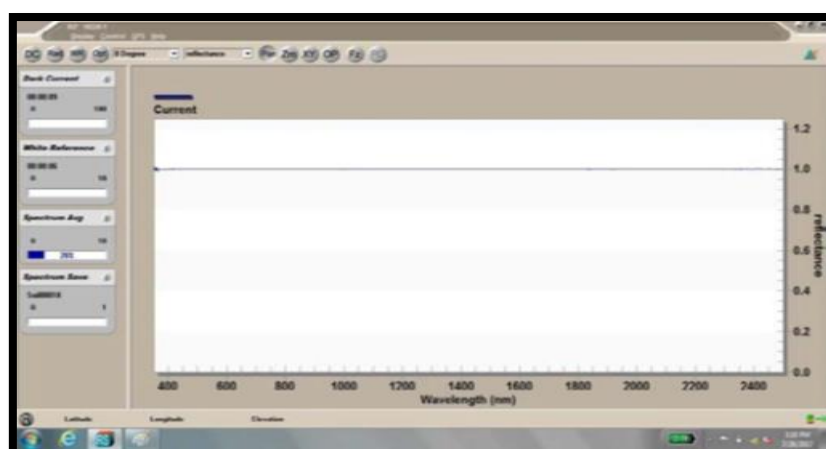


Figure 3.7: White Reference Calibration

3.4.5. Software ViewSpecPro

Hyperspectral data obtained from the ASD FieldSpec 4 spectroradiometer were handled and analysed in this work using ViewSpec Pro software. Malvern Panalytical (previously ASD Inc.) created ViewSpec Pro, a specialised program for viewing, manipulating, and exporting spectrum data gathered by ASD equipment. Some of ViewSpec Pro's primary features are:

- With zooming and comparison features, spectral data representation makes it possible to see spectral curves across the entire wavelength range (350–2500 nm).
- Spectral Information Basic processes like cutting, smoothing, averaging, and converting radiance to reflectance are made possible by pre-processing.
- File management Provides export options in ASCII and CSV formats for additional analysis and enables simultaneous analysis of several spectral files (.asd format).
- Important acquisition parameters, including date, time, instrument settings, and integration time, are shown on the metadata display.
- Tools for Calibration and Correction Provides quality checks, white reference calibration, and dark current correction to guarantee data accuracy.

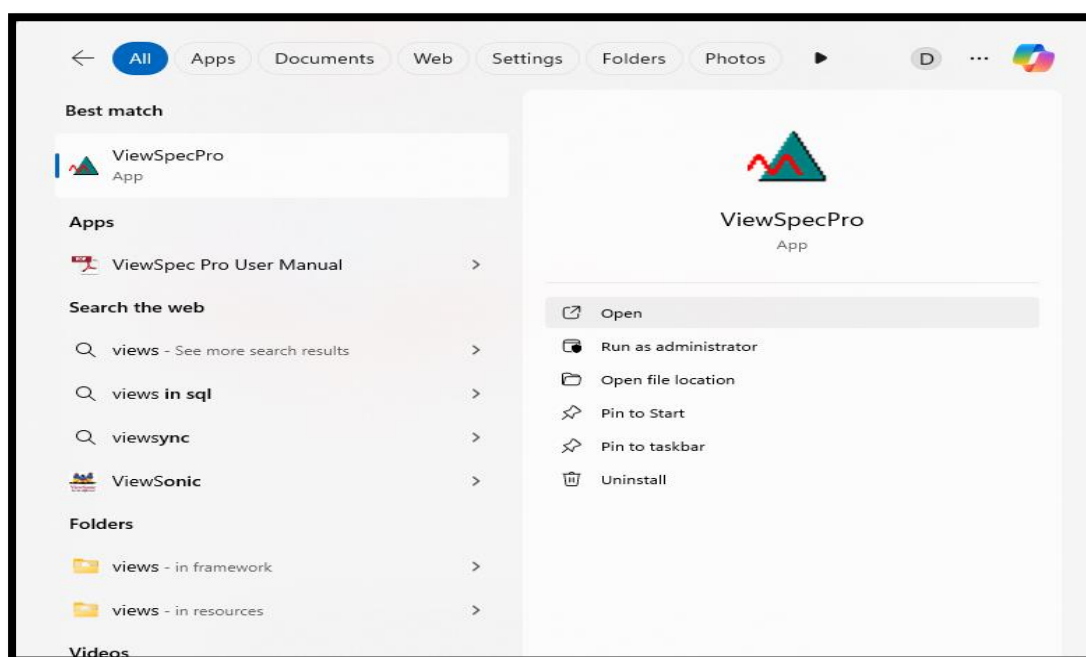


Figure 3.8: ViewSpecPro software

3.5 Workflow for Proposed Methodology

Step 1: Acquisition of Hyperspectral Data

- Hyperspectral reflectance data from food samples (chili powder, turmeric powder, and tea powder) are acquired using the ASD FieldSpec4 Spectroradiometer.
- The instrument records spectral signatures within the 350–2500 nm wavelength range.
- A controlled environment lab (dark room) is utilized to remove external light disturbance, guaranteeing precise spectral data gathering.
- Calibration of the standard white reference occurs every 10-15 minutes to ensure measurement consistency.

Step 2: Data Preprocessing and Labeling

- Raw spectral data undergoes preprocessing to eliminate noise, inconsistencies, and undesired variations.
- Data normalization methods are employed to maintain consistency among samples.
- Every sample is classified based on its category.

Step 3: Feature Extraction

During the Feature Extraction phase, various statistical characteristics can be extracted from hyperspectral data to improve the machine learning model's capacity to distinguish between pure and contaminated food samples, as shown in Table 3.3. The following characteristics are derived from the spectral data of each sample. For each hyperspectral reflectance spectrum $\mathbf{X}$, the following features are computed across different wavelengths Table 3.3 below highlights several statistical, energy-based, peak-valley, and slope-related properties derived from spectral data. These attributes are crucial for examining spectral properties and enhancing classification models. This aggregated table offers an organized summary of the retrieved spectral characteristics, facilitating a thorough investigation of spectral data for classification and pattern recognition applications.

Table 3.3: Statistical feature extraction methods

Feature Category	Feature Name	Formula / Function	Description
Statistical Features	Mean	<code>np.mean(X, axis=1)</code>	The average spectral reflectance value, representing the central tendency of the spectrum.
	Standard Deviation	<code>np.std(X, axis=1)</code>	Measures the dispersion of reflectance values, indicating spectral variability.
	Median	<code>np.median(X, axis=1)</code>	The middle value of the spectral distribution, providing a robust measure of central tendency.
	Minimum Value	<code>np.min(X, axis=1)</code>	The lowest reflectance value observed in the spectrum.
	Maximum Value	<code>np.max(X, axis=1)</code>	The highest reflectance value observed in the spectrum.
	Range	Max - Min	The difference between the maximum and minimum reflectance values, indicating spectral spread.
	Skewness	<code>skew(X, axis=1)</code>	Measures the asymmetry of spectral distribution. A positive skew indicates a right-tailed distribution, while a negative skew indicates a left-tailed distribution.
	Kurtosis	<code>kurtosis(X, axis=1)</code>	Measures whether the spectral distribution is heavy-tailed (presence of outliers) or light-tailed compared to a normal distribution.
	25th Percentile (Q25)	<code>np.percentile(X, 25, axis=1)</code>	The value below which 25% of the data points lie.
	75th Percentile (Q75)	<code>np.percentile(X, 75, axis=1)</code>	The value below which 75% of the data points lie.

	Interquartile Range (IQR)	$Q_{75} - Q_{25}$	Measures the spread of the middle 50% of the data, reducing the effect of outliers.
Energy-Based Features	Energy	<code>np.sum(X**2, axis=1)</code>	Sum of squared reflectance values across wavelengths, measuring total signal strength.
	Power	<code>Energy / X.shape[1]</code>	Average energy per wavelength, representing signal intensity over the spectral range.
Spectral Area Features	Area Under Curve (AUC)	<code>trapezoid(X[i, :], wavelengths)</code>	Measures the total area under the spectral curve using numerical integration (Trapezoidal rule).
Peak and Valley Features	Peak Value	<code>np.max(X, axis=1)</code>	The highest reflectance value in the spectrum indicates the most dominant wavelength.
	Peak Wavelength	<code>wavelengths[np.argmax(X, axis=1)]</code>	The wavelength corresponding to the peak reflectance value.
	Valley Value	<code>np.min(X, axis=1)</code>	The lowest reflectance value in the spectrum.
	Valley Wavelength	<code>wavelengths[np.argmin(X, axis=1)]</code>	The wavelength corresponding to the valley (lowest) reflectance value.
Slope (First Derivative) Features	Max Slope	<code>np.max(X_deriv, axis=1)</code>	Maximum rate of change in reflectance values, capturing steep spectral transitions.
	Min Slope	<code>np.min(X_deriv, axis=1)</code>	Minimum rate of change in reflectance values, highlighting spectral dips.
	Mean Slope	<code>np.mean(X_deriv, axis=1)</code>	The average rate of change across wavelengths indicates the general trend of spectral variations.

Step 4: Creation of Spectral Signature

- A spectral library is created for various food samples according to their spectral features.
- This spectral database acts as a reference dataset for classification and examination.

Step 5: Machine Learning Model Development

- Supervised machine learning models are developed using the extracted spectral characteristics.
- Different algorithms are investigated to improve classification accuracy, such as:
 - Support Vector Machine (SVM)
 - Random Forest (RF)
 - XGBoost
 - K-Nearest Neighbors (KNN)

To ensure accurate and reliable detection of food adulteration, we employed four machine learning algorithms: XGBoost, Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). These models were chosen for their diverse strengths in classification, enabling a comprehensive evaluation of their effectiveness in distinguishing between adulterated and pure food samples. XGBoost, an advanced ensemble learning algorithm, was configured with a softmax objective function, 100 estimators, and a max depth of 6, optimizing its ability to classify food samples accurately. Similarly, Random Forest, another ensemble-based method, utilized 100 decision trees, offering high robustness and resistance to overfitting. KNN, a distance-based algorithm, classified samples by considering the five nearest neighbors, while SVM, with its Radial Basis Function (RBF) kernel, effectively handled non-linear decision boundaries in food classification. Each model was trained and evaluated using classification accuracy, precision, recall, and F1-score. The results revealed that XGBoost and Random Forest significantly outperformed KNN and SVM, achieving higher classification accuracy and better generalization. KNN and SVM, on the other hand, showed lower accuracy, indicating that traditional distance-based and kernel-based methods may struggle with complex adulteration patterns. The superior

performance of ensemble models (XGBoost and RF) highlights their ability to capture intricate relationships in food adulteration data, making them the most suitable choices for real-world applications [13].

Step 6: Classification

- The trained model classifies input sample data.
- Adulteration levels are identified based on spectral deviations.
- Classification accuracy and confusion metrics (such as precision, recall, and F1-score) are assessed.
- The results are contrasted with traditional food quality evaluation methods to confirm the findings.

3.6. Preprocessing Pipeline for Spectral Data

3.6.1. Savitzky-Golay smoothing

Savitzky, Abraham, and Goly Marcel came up with the Savitzky-Golay smoothing filter to clean up the chemical spectrum analyzer's noisy data. They suggested a way to smooth things out using the least squares polynomial as a guide. They showed that fitting a polynomial to part of the input sample and checking the output polynomial at one point in the approximation interval is the same as the fixed impulse response for discrete convolution. They also showed that smoothing with the least squares can reduce noise while keeping the shape and height of the input signal the same [14].

The Savitzky-Golay filter was utilized to smooth the spectral curves while maintaining the key characteristics of the spectral signature. This filter executes local polynomial regression on a set of values to ascertain the smoothed value for each point.

- Purpose: Smoothens each spectrum to reduce random noise while preserving peak shapes using equation number 1.
- Method: Polynomial smoothing with window length is **15** and Polynomial order is **3**.

The settings applied for the Savitzky-Golay filter were:

- Size of the window: 15 points
- Order of polynomial: 3
- Derivative order: 1 and 2

$$y_i = \sum c_j y_{i+j} \dots\dots\dots (1)$$

The alteration of the second derivative was particularly advantageous in enhancing minor spectral features and diminishing baseline variations across different samples.

3.6.2. Noise and Irrelevant Data Removal

Along with smoothing, preprocessing methods like standard normal variate (SNV), multiplicative scatter correction (MSC), and first and second derivatives are used to make spectra more uniform and lessen the effects of scattering. Also, wavelength selection methods like principal component analysis (PCA) or the successive projections algorithm (SPA) are used to get rid of variables that don't matter and make the model work better.

These processes make sure that the spectral data utilized for categorization or prediction is clean, relevant, and better for algorithmic learning. This makes adulteration detection systems more reliable and accurate.

- Wavelength trimming: Removed extreme ends of the spectrum, first 50, and data beyond 2000 to exclude noise-prone regions.
- Operation: Remove non-informative spectral regions

3.6.3. Baseline Correction (Second Derivative)

Baseline variation is a typical problem with spectroscopic data, especially with hyperspectral non-imaging datasets that are used to find food adulteration. These changes could be caused by things like light scattering, instrument drift, or the fact that the sample isn't uniform. They can hide essential spectral properties, making classification or regression models less reliable.

To fix this, baseline correction methods are used during the preprocessing stage. Using the second derivative of the spectral signal is one of the best ways to do this. The Savitzky-Golay algorithm is usually used to find this value. The second derivative not only makes small, overlapping peaks stand out more, but it also stops steady and linear baseline movements. It makes little changes in spectral curvature stand out more, which makes it easier to find small spectral features that have to do with chemical composition or adulterants.

The second derivative modification gets rid of additive baseline offsets and makes inflection points stand out. These points are very important for determining the

difference between pure and mixed samples. But it also makes noise stronger, thus it is typically used with smoothing filters like the Savitzky-Golay smoothing filter to keep the signal quality high while lowering the noise [14]. Overall, using the second derivative to adjust the baseline is a powerful step in spectrum preprocessing. It makes the data better and lets you extract and classify features more accurately in hyperspectral data analysis.

- Purpose: Removes baseline drift and highlights subtle peaks.
- Technique: Second derivative transformation using the same Savitzky-Golay configuration.

3.6.4. Standard Scaling (Z-score Normalization)

In hyperspectral data analysis, especially when using machine learning algorithms, it is essential to normalize the input features to ensure that each wavelength contributes equally to the model's learning process. One widely used normalization technique is Standard Scaling, also known as Z-score normalization.

Z-score normalization transforms the original data by centering it around zero with a standard deviation of one. This is achieved using the formula:

$$Z = \frac{X - \mu}{\sigma} \dots\dots\dots (2)$$

where X is the original spectral value, μ is the mean of the spectral feature, and σ is the standard deviation. This transformation ensures that all features (wavelengths) have the same scale, eliminating biases caused by varying units or ranges.

In hyperspectral datasets, where the number of wavelengths (features) is often very high, unscaled data can adversely impact model convergence and performance, especially in distance-based algorithms like K-nearest neighbors (KNN), Support Vector Machines (SVM), and even gradient-based methods like logistic regression or neural networks. Z-score normalization helps improve the numerical stability of these algorithms, accelerates convergence, and enhances classification accuracy. Therefore, standard scaling is a crucial preprocessing step to ensure fair comparison among spectral features, robust model training, and reliable prediction of adulteration in food products.

3.7. Feature Extraction from Spectral Data

3.7.1. Statistical Feature

In hyperspectral non-imaging data analysis, statistical features play a crucial role in summarizing the essential characteristics of spectral signals. These features provide a compact and informative representation of the data, which can significantly enhance the performance of machine learning models by reducing dimensionality and focusing on the most relevant information. Commonly extracted statistical features include:

- **Mean:** Represents the average reflectance or absorbance across a selected spectral range, useful for identifying overall intensity patterns.
- **Standard Deviation:** Measures the variation or dispersion within the spectral signal, helping to capture sample heterogeneity.
- **Skewness:** Indicates the asymmetry of the spectral distribution, which may reflect the presence of impurities or adulterants.
- **Kurtosis:** Describes the peakedness or flatness of the spectral curve, useful in detecting subtle changes in spectral shape.
- **Maximum and Minimum:** Capture the extreme values in the spectrum, which can be indicative of strong absorbance features or outliers.
- **Energy and Entropy:** Provide insights into the signal strength and complexity, respectively, often used in advanced spectral analysis.

These statistical descriptors are computationally efficient and help in simplifying complex hyperspectral data into a manageable format for supervised learning algorithms. By extracting and selecting appropriate statistical features, it becomes easier to distinguish between pure and adulterated food samples based on their spectral patterns.

3.7.2. Signal-Based Feature

In the context of hyperspectral non-imaging analysis, signal-based features are extracted directly from the spectral signal to capture its shape, trend, and structure. These features reflect the intrinsic physical and chemical properties of the sample as recorded in its spectral signature. Unlike purely statistical features, signal-based

features focus on waveform characteristics that are often critical in identifying adulterants or compositional differences. Common signal-based features include:

- **First and Second Derivatives:** Highlight the rate of change in spectral intensity and curvature, useful for detecting inflection points and resolving overlapping peaks.
- **Peak Position:** Identifies the wavelength at which maximum absorbance or reflectance occurs, which is often specific to certain chemical compounds.
- **Peak Intensity:** Measures the magnitude of the spectral peak, indicating the concentration of specific constituents.
- **Full Width at Half Maximum (FWHM):** Describes the width of a peak at half of its maximum height, providing insights into the purity and sharpness of spectral features.
- **Area Under Curve (AUC):** Represents the integrated intensity across a specified spectral range, often correlated with concentration.
- **Slope and Intercept:** Derived from linear regression over spectral regions, useful for trend analysis and baseline evaluation.

Signal-based features are particularly valuable for identifying subtle differences between pure and adulterated samples, as they retain the spectral profile's shape and amplitude information. When combined with statistical and domain-specific features, they contribute to a comprehensive and robust feature set for machine learning-based adulteration detection systems.

3.7.3. Peak and Valley Metrics

Peak and valley metrics are essential signal-based features used in hyperspectral data analysis to characterize the local maxima (peaks) and local minima (valleys) within a spectral signature. These features offer meaningful insights into the chemical composition of a sample, as different substances exhibit unique absorption or reflectance characteristics at specific wavelengths. Key peak and valley features include:

- **Peak Position and Valley Position:** Identify the precise wavelengths at which maximum or minimum spectral responses occur. These positions often correspond to specific molecular bonds or chemical constituents.

- **Peak Height and Valley Depth:** Measure the amplitude of the spectral response above or below a reference baseline, reflecting the concentration of specific compounds.
- **Peak-to-Valley Difference:** Represents the contrast between adjacent peaks and valleys, indicating the spectral sharpness or variability.
- **Number of Peaks and Valleys:** Indicates the complexity of the spectral signal; more complex spectra may suggest the presence of multiple components or adulterants.
- **Peak Prominence and Width:** Describe the significance and breadth of each peak, useful for distinguishing between broad and narrow absorption bands.

These metrics are highly informative in food adulteration detection because adulterated samples often exhibit altered spectral peak structures due to the presence of foreign substances. By analyzing peak and valley characteristics, one can enhance class separability in supervised models and improve the interpretability of spectral data.

3.7.4. PCA for Feature Extraction and Visualization

Principal Component Analysis (PCA) is a widely used unsupervised technique for dimensionality reduction and visualization in hyperspectral data analysis. Given that hyperspectral datasets typically consist of hundreds or thousands of closely spaced wavelengths, many of which are highly correlated or redundant, PCA helps transform this high-dimensional data into a set of orthogonal components called principal components (PCs) that capture the maximum variance in the data [15,16]. PCA works by computing the eigenvectors and eigenvalues of the covariance matrix of the data and projecting the original variables onto a new coordinate system. The first few principal components (typically the first two or three) retain most of the variance in the original data while discarding noise and redundant information. This transformation not only improves computational efficiency for machine learning models but also enhances model generalization by eliminating irrelevant features [17].

In addition to feature reduction, PCA is a powerful tool for data visualization. Plotting the first two or three principal components allows for an intuitive inspection of sample clustering and class separability, which is particularly useful for identifying patterns between pure and adulterated food samples. Overall, PCA provides both analytical and interpretive advantages by reducing data complexity while preserving

the most informative variance, making it an essential step in the preprocessing pipeline of hyperspectral non-imaging data analysis.

Objective: Reduce dimensionality while retaining variance in the spectral features.

Principal Component Analysis (PCA):

Projects high-dimensional data into a lower-dimensional subspace. Components are orthogonal (uncorrelated) directions of maximum variance.

3.8. Data Classification

3.8.1 Support Vector Machine (SVM)

Support Vector Machine (SVM) is one of the most widely used machine learning classifiers in remote sensing applications due to its robustness, generalization ability, and effectiveness in handling high-dimensional data such as hyperspectral imagery. SVM demonstrates high classification accuracy even without the necessity for prior feature reduction, making it particularly suitable for hyperspectral data analysis [18]. Unlike parametric classifiers such as Maximum Likelihood Classification (MLC), which rely on the statistical distribution of each class, SVM operates on the principle of geometrical separation. Vladimir Vapnik originally developed it as a binary classifier, but it can be extended to multi-class classification tasks using techniques such as one-vs-one or one-vs-rest [19]. SVM identifies the optimal hyperplane that maximally separates different classes by locating the support vectors, data points that lie closest to the decision boundary. Although it is a linear classifier by default, SVM can also handle non-linear data through the use of kernel functions. Kernels such as Radial Basis Function (RBF) or polynomial kernels allow SVM to project the input data into a higher-dimensional feature space, where non-linearly separable classes become linearly separable [20]. This capability to manage high-dimensional, non-linear, and complex data structures makes SVM a powerful tool in hyperspectral remote sensing and food adulteration detection.

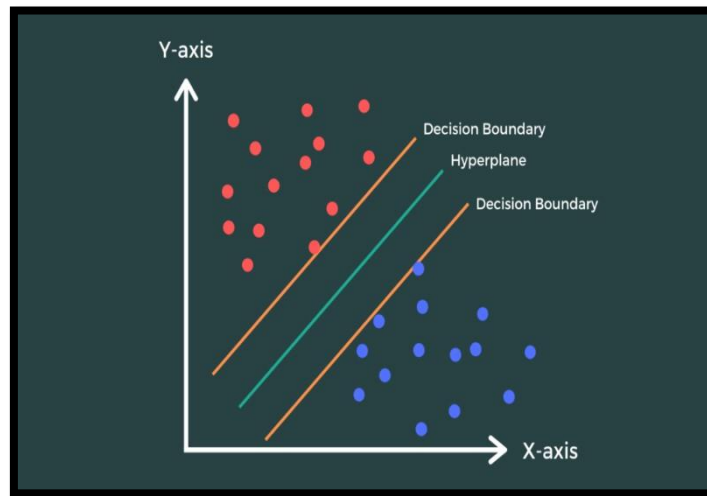


Figure 3.9: Support Vector Machine Classifier

3.8.2 Random Forests (RFs)

Random Forests (RFs) are ensemble learning methods used for classification and regression tasks, known for their high accuracy, robustness to overfitting, and ability to handle high-dimensional datasets such as hyperspectral non-imaging data. RFs work by constructing a multitude of decision trees during training and outputting the class that is the mode of the classes (for classification) or the mean prediction (for regression) of the individual trees [21]. Each tree in a random forest is built from a bootstrap sample of the original data, and at each split, a random subset of features is considered, which introduces diversity among the trees and improves generalization. This randomized tree-building process reduces correlation between the trees, resulting in a more stable and accurate classifier compared to individual decision trees. In hyperspectral data analysis, Random Forests are particularly effective because they can handle a large number of input features without requiring feature reduction, and they provide internal estimates of feature importance, helping to identify the most relevant wavelengths for classification [22]. Moreover, RFs are non-parametric and do not assume any specific data distribution, making them suitable for complex and nonlinear spectral patterns often seen in food adulteration studies.

The interpretability, speed, and high accuracy of RFs make them a popular choice for hyperspectral classification tasks, including the detection of adulterants in powdered food products such as turmeric, chili, or tea.

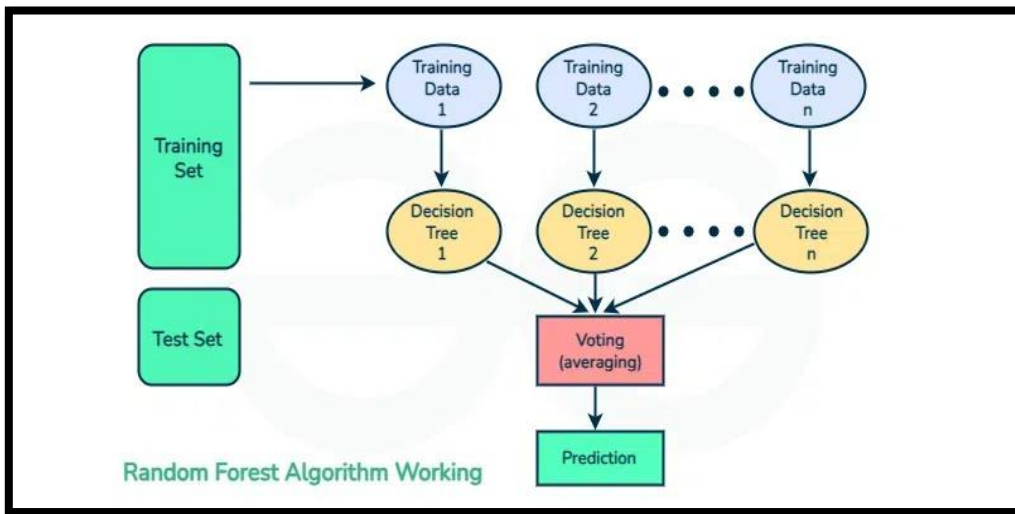


Figure 3.10: Random Forests Classifier

3.8.3. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an advanced implementation of the gradient boosting framework, designed for high performance and efficiency. It is particularly effective for handling structured, tabular, and high-dimensional datasets, such as those encountered in hyperspectral non-imaging data used for food adulteration detection. XGBoost builds an ensemble of decision trees sequentially, where each new tree corrects the residual errors made by the previous trees. The model optimizes a regularized objective function, which combines a loss function and a complexity penalty, thereby improving generalization and reducing overfitting [23]. Unlike Random Forests, which average results across independently built trees, XGBoost trains trees in a boosting manner, giving more weight to incorrectly classified samples in each iteration. This leads to a highly accurate and robust classifier, especially in noisy and unbalanced datasets. XGBoost also incorporates techniques like tree pruning, parallel processing, missing value handling, and sparsity awareness, making it computationally efficient and scalable for large spectral datasets [24]. In hyperspectral food adulteration studies, XGBoost has shown promising results due to its ability to model complex, nonlinear relationships between spectral features and class labels. It also provides feature importance scores, aiding in the identification of key wavelengths relevant to distinguishing between pure and adulterated samples.

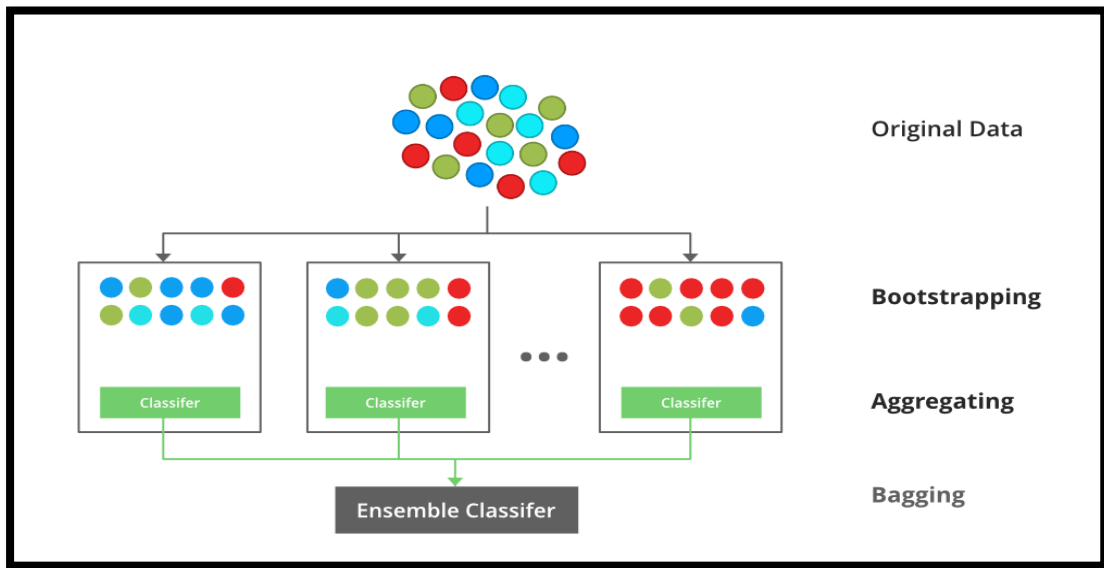


Figure 3.11: Extreme Gradient Boosting (XGBoost) Classifier

3.8.4. K Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) is a simple yet powerful non-parametric classification algorithm that is widely used in hyperspectral data analysis. KNN operates on the principle of instance-based learning, where the classification of a new sample is determined by the majority class among its K nearest neighbors in the feature space. The algorithm relies on a distance metric, typically Euclidean distance, to quantify similarity between data points [25]. KNN is particularly suitable for hyperspectral non-imaging data due to its flexibility and ability to model nonlinear class boundaries without making assumptions about the underlying data distribution. It performs well in low- to moderately sized datasets, and its simplicity makes it easy to implement and interpret. However, it can become computationally expensive when applied to high-dimensional data, such as full-range hyperspectral spectra. To mitigate this, dimensionality reduction techniques like PCA or feature selection methods are often applied before using KNN [26]. In the domain of food adulteration detection, KNN has shown effectiveness in distinguishing between pure and adulterated samples by leveraging spectral similarity. Its performance is sensitive to the choice of K value and the distance metric, which should be carefully optimized for accurate classification.

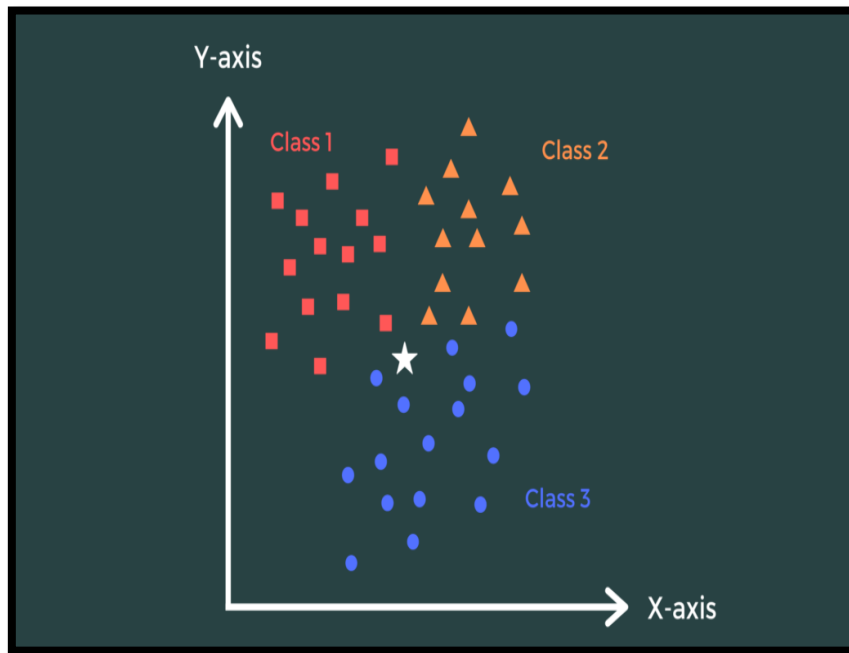


Figure 3.12: K Nearest Neighbor (KNN) Classifier

Mainly, there are two steps in classification:

Step 1: Input the training data and test data into the algorithm.

Step 2: Assign the value for K (any integer) based on the nearest data points.

Step 3: For every point in the test data:

- Calculate the distance between test data and each row of training data using the following techniques, namely Euclidean, Manhattan, or Hamming distance. The most commonly used method is Euclidean.
- Based on the obtained distance value, arrange them in ascending order.
- Choose the top K rows from the sorted array.
- Assign a class to the test point based on the most frequent class of the chosen rows.

Step 4: Stop.

3.9. Accuracy Assessment

In remote sensing and machine learning, checking for accuracy is an important part of categorization assignments. It gives a numerical score for how well a model does at predicting the right class labels. We used a set of standard accuracy metrics based on the confusion matrix to see how well the models worked. The confusion

matrix contrasts the actual class labels of the testing dataset with the anticipated outputs made by the classification models. The test comprises figuring out the overall accuracy (OA), the procedure's accuracy (PA), and the user's accuracy (UA), which are all frequently used in remote sensing applications. We also calculated more advanced statistical indicators, including the F1_macro score, precision, and recall, to get a better picture of how the system worked. These metrics help us figure out how reliable the model is, how well it does in each class, and how well it agrees with the data in space. This lets us look at not only how many predictions were true, but also how well and consistently the model classifies data across multiple classes. The next part goes into great depth on the formulas for each measure of accuracy. Overall accuracy (OA) is defined as the probability that the sample will be classified correctly. It is the total number of samples of correct classification to the total number of the testing dataset [27]. So, as shown in Eq. 3.2, the true positive plus true negative is divided by the entire quantity of testing data.

$$OA = \frac{\sum M_{ij}}{n} * 100\% \quad \dots\dots\dots (4)$$

The procedure's accuracy (PA) is the same as an omission error, and it is important to know how the well-specific class was classified. You can find it by dividing the number of correctly classified individuals by the sum of their reference columns, as shown in Eq. (3.3). If you divide the number of correctly identified samples in a class by the total number of samples in that class, you get the commission. User accuracy (UA) is what this is called. Eq. (3.4).

$$PA = \frac{M_{ij}}{M_{i+}} * 100 \%, M_{i+} = M_{i1} + M_{i2} + M_{i3} + M_{i4} + M_{i5},$$

$$\dots\dots\dots (5)$$

$$UA = \frac{M_{ij}}{M_{+j}} * 100 \%, M_{+j} = M_{1j} + M_{2j} + M_{3j} + M_{4j} +$$

$$M_{5j}, \quad \dots\dots\dots (6)$$

The above expression shows that M is a 10x10 confusion matrix i is the row number of the i th species on the actual rows, j is the number of columns of the j th species on the prediction columns, M_{i+} is the sum of the i th row, M_{+j} is the sum of the j th columns, and n is the number of test sets.

3.10 Summary

This chapter outlines a detailed experimental methodology for a non-destructive food adulteration detection system. The research focuses on analyzing powdered chili, turmeric, and tea samples using hyperspectral non-imaging techniques and machine learning. The system aims to identify adulterants by detecting subtle compositional differences invisible to the naked eye. The process began with collecting diverse food samples from local markets and commercial brands in Chhatrapati Sambhajinagar, Maharashtra. To create a reliable reference dataset, pure samples were adulterated with substances like brick powder and salt powder in controlled proportions.

The spectral data of all samples was captured using an ASD FieldSpec4 Spectroradiometer in a controlled darkroom to ensure accuracy. The raw data was then preprocessed to remove noise and inconsistencies using methods like Savitzky-Golay smoothing and baseline correction. After preprocessing, various features, including statistical and signal-based properties, were extracted from the data. The study used four machine learning models Support Vector Machine (SVM), Random Forest (RF), XGBoost, and K-Nearest Neighbors (KNN)—for classification. The results showed that ensemble models, specifically XGBoost and Random Forest, achieved superior accuracy and better generalization compared to the other models. The performance was validated using standard metrics derived from a confusion matrix.

3.11 References

- [1].Cen, H., and He, Y ‘Theory and application of near infrared reflectance spectroscopy in determination of food quality’, Trends in Food Science & Technology, 2007, 18, (2), pp. 72-83
- [2].Pasquini, C. ‘Near infrared spectroscopy: A mature analytical technique with new perspectives–A review’, Analytica chimica acta, 2018, 1026, pp. 8-36
- [3].Dhangar, V.D., Dhole, P.V., Shejul, S.D., and Gawali, B.W.: ‘Assessment of Adulteration from Food Products using ASD Field Spec4’, in Editor: ‘Book Assessment of Adulteration from Food Products using ASD Field Spec4’ (IEEE, 2023, edn.), pp. 1-6
- [4].Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G., and Frias, J.M.: ‘Hyperspectral imaging–an emerging process analytical tool for food quality and safety control’, Trends in food science & technology, 2007, 18, (12), pp. 590-598
- [5].Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., and Jent, N.: ‘A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies’, Journal of pharmaceutical and biomedical analysis, 2007, 44, (3), pp. 683-700
- [6].Boiret, M., & Chauchard, F. (2017). Use of near-infrared spectroscopy and multipoint measurements for quality control of pharmaceutical drug products. Analytical and bioanalytical chemistry, 409, 683-691.
- [7].Martinez-Vega, B., Tkachenko, M., Matkabi, M., Ortega, S., Fabelo, H., Balea-Fernandez, F., & Chalopin, C. (2022). Evaluation of preprocessing methods on independent medical hyperspectral databases to improve analysis. Sensors, 22(22), 8917.
- [8].Jacquinet-Husson, N., Armante, R., Scott, N. A., Chédin, A., Crépeau, L., Boutammine, C., ... & Makie, A. (2016). The 2015 edition of the GEISA spectroscopic database. Journal of Molecular Spectroscopy, 327, 31-72.
- [9].Vinaixa, M., Schymanski, E. L., Neumann, S., Navarro, M., Salek, R. M., & Yanes, O. (2016). Mass spectral databases for LC/MS-and GC/MS-based metabolomics: State of the field and future prospects. TrAC Trends in Analytical Chemistry, 78, 23-35.

- [10]. Priyanka U. Randive, Ratnadeep R. Deshmukh, “Study of detecting Plant diseases using Non-Destructive Methods: A Review”, IJETTCS Journal, 2018, 1-6.
- [11]. Tejas U. Padghan, Ratnadeep R. Deshmukh, Jaypalsing N. Kayte, Pooja V. Janse, “Hyperspectral Estimation of soil salt content using partial least square regression method: a review”, 2018, JECET, 7. 2, 156-161
- [12]. Pooja Vinod Janse and Ratnadeep R. Deshmukh, “Hyperspectral Remote Sensing for Agriculture: A Review”, International Journal of Computer Applications, 2017, 172 –7, 30-34.
- [13]. Al-Awadhi, M. A., & Deshmukh, R. R. (2021, October). Honey Classification using Hyperspectral Imaging and Machine Learning. In 2021 Smart Technologies, Communication and Robotics (STCR) (pp. 1-5). IEEE.
- [14]. Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8), 1627-1639.
- [15]. Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3), 37-52.
- [16]. Jolliffe, I. T. (2002). *Principal component analysis for special types of data* (pp. 338-372). Springer New York.
- [17]. Wu, D., & Sun, D. W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part II: Applications. *Innovative Food Science & Emerging Technologies*, 19, 15-28.
- [18]. Gualtieri, J. A., & Crompton, R. F. (1998). Support vector machines for hyperspectral remote sensing classification.
- [19]. Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on geoscience and remote sensing*, 42(8), 1778-1790.
- [20]. Fletcher, R. S., Everitt, J. H., & Yang, C. (2011). Identifying saltcedar with hyperspectral data and support vector machines. *Geocarto International*, 26(3), 195-209.
- [21]. Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (2017). *Classification and regression trees*. Routledge.

- [22]. Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, 7, 1-13.
- [23]. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- [24]. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... & Zhou, T. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4.
- [25]. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1), 21-27.
- [26]. Karasiak, N., Dejoux, J. F., Monteil, C., & Sheeren, D. (2022). Spatial dependence between training and test sets: another pitfall of classification accuracy assessment in remote sensing. *Machine Learning*, 111(7), 2715-2740.

RESULTS AND DISCUSSION

4.1. Introduction

This chapter summarizes the findings of an experiment that used hyperspectral non-imaging technology to identify adulteration in a few food powders, specifically tea, turmeric, and chili powder. Model efficiency, preprocessing impact, spectral reflectance behavior, and classification performance are examined in the results. A comparative analysis of various machine learning algorithms is also covered to determine the best method for identifying adulterants [1,2,3]. Food adulteration presents major risks to consumer health and product purity. Many traditional methods have been used for decades to identify adulterants in spices, tea, and other food items. These procedures are straightforward, economical, and require minimum apparatus, making them advantageous for rapid, immediate detection. This section offers a comprehensive overview of traditional methods for identifying adulteration in prevalent food products [4].

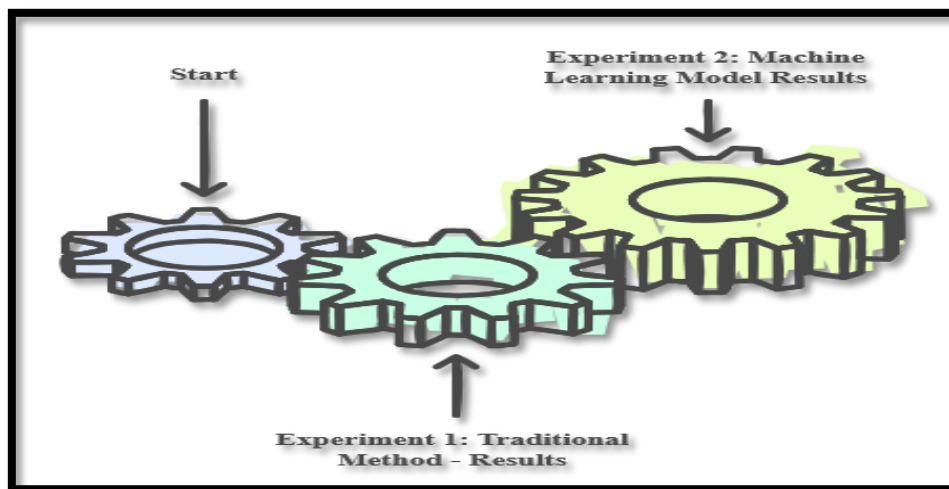


Figure 4.1: Experimental Process.

Before using hyperspectral non-imaging technology and machine learning models, we used simple, non-chemical methods like looking at the texture, doing a standardized water test, and looking at the color of each food powder we chose chili powder, turmeric powder, and tea powder. The water test helped find brick powder in chili powder by looking at how the silt settled. We found manmade dyes like metanil yellow in turmeric powder by looking at how their colors spread out in water. In the

same way, iron filings and artificial colorants were found in tea powder by how they sank or streaked in water. Before collecting hyperspectral data and training the model, the results of these water-based tests were utilized to label and standardize the samples. This made sure that the categorization was accurate and consistent.

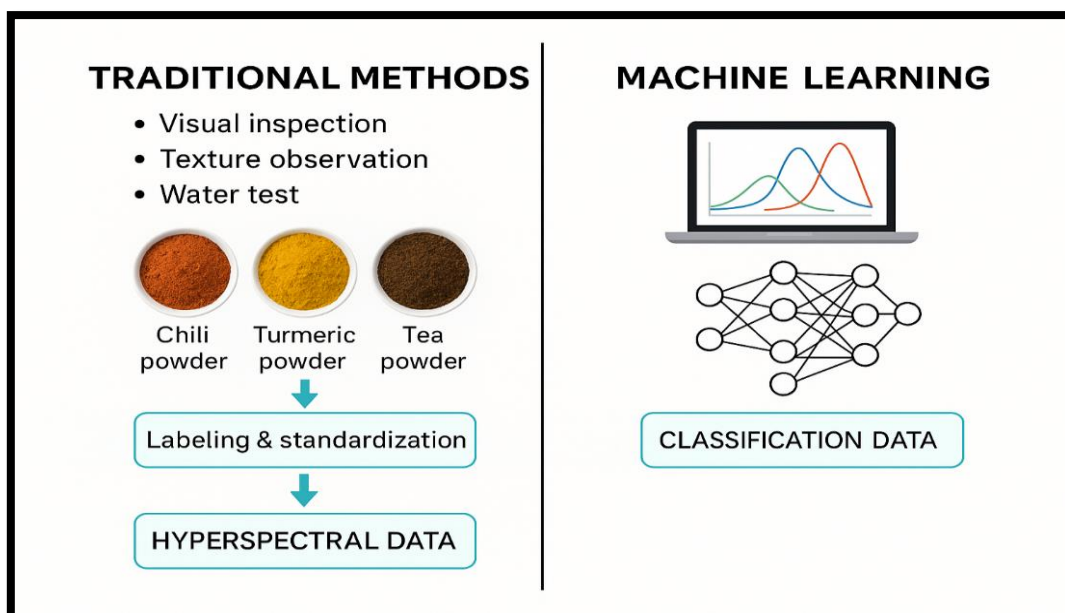


Figure 4.2: Traditional and Machine learning methods

4.2. Experiment No. 1.

4.2.1. Traditional Methods for Detecting Adulteration

Food adulteration is a serious hazard to people's health and the quality of the product. For many years, people have used conventional ways to find adulterants in foods like tea, spices, and other things that people eat a lot of. People like these methods because they are simple, cheap, and don't need a lot of technology. They usually involve looking at the object, checking its texture, and doing basic solubility tests. They are especially helpful for quick, on-the-spot screening, especially in homes or places with few resources, because they are so easy to use. These methods are good for finding things out at first, but they aren't always accurate or standardized. This shows that we need more advanced, data-driven procedures.

4.2.2. Traditional Adulteration Detection in Chili Powder

Chili powder is often contaminated with synthetic dyes, salt powder, and brick powder to improve its visual appeal and weight. Various conventional assays may be performed to identify these adulterants:

1. Identification of Brick Powder or Salt Powder Adulteration Method: Introduce a tiny amount of Chili powder into a glass of water.
 - **Observation:** Pure chili powder will either float or spread uniformly in water, while adulterants like brick powder or salt powder would accumulate at the bottom.
2. Detection of Artificial Coloring: Method: Apply a little amount of chili powder on a damp paper or your hand.
 - **Observation:** The presence of artificial color will result in a red or yellow smear on the surface. Pure chili powder will not leave any residual coloration.

These procedures are rapid, cheap, and good for first screening at home.

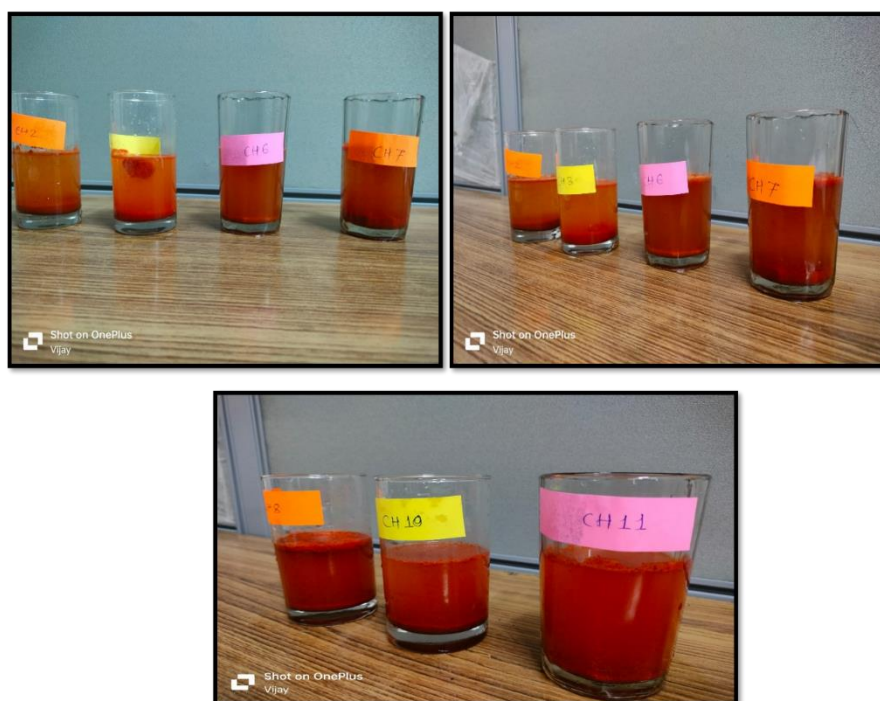


Figure 4.3: Detection of Chili Powder Adulteration

4.2.3. Traditional Adulteration Detection of Turmeric Powder

Turmeric powder is often contaminated with additives such as chalk powder, flour, and artificial pigments to enhance its appearance or increase its weight. These adulterants pose a risk to human health, necessitating their detection by straightforward, conventional procedures. The methods below may assist in detecting prevalent adulterants in turmeric powder.

Chalk Powder or Starch Adulteration Test: A small pinch of turmeric powder was added to a glass of water.

- **Observations:** Pure turmeric settled at the bottom without leaving colour in the water. Samples suspected of adulteration formed a whitish layer, indicating the presence of chalk powder or starch. Some samples displayed cloudiness in the water, further confirming the presence of adulteration.

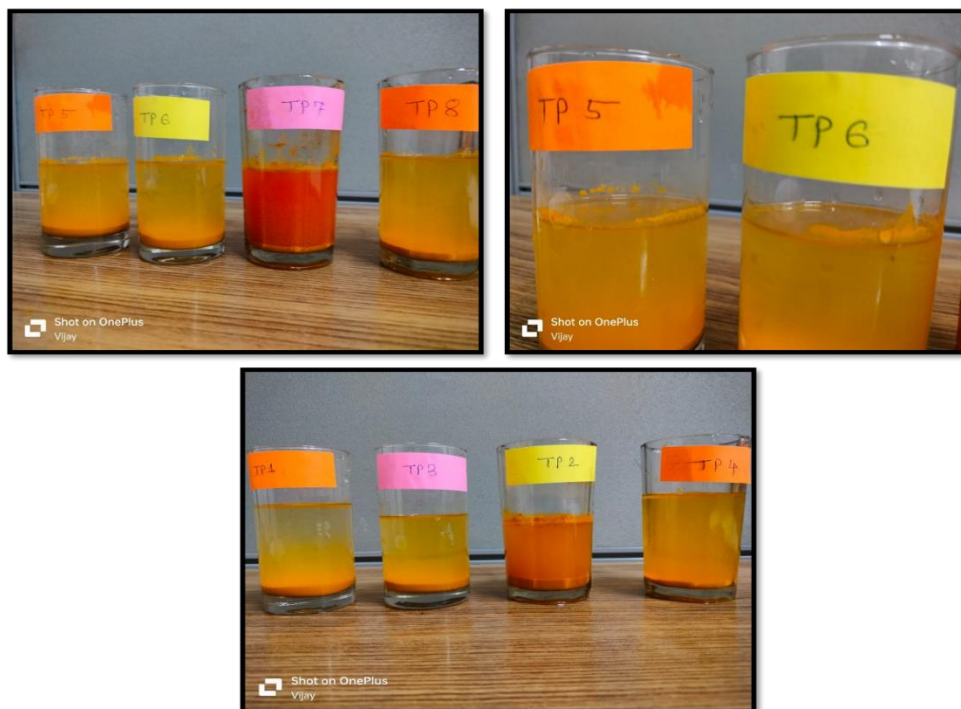


Figure 4.4: Detection of Turmeric Powder Adulteration

4.2.4. Traditional Adulteration Detection of Tea Powder

Tea powder is frequently adulterated with artificial colouring agents to enhance its appearance. These adulterants pose potential health risks, making it essential to detect them using simple, traditional methods. The following tests help identify the presence of artificial colour in tea powder.

Artificial Coloring Test: A small amount of tea powder was sprinkled on a piece of white paper, followed by a few drops of water. Additionally, a separate test was performed by adding tea powder to a glass of water.

- **Observations:** Pure tea powder retained its natural appearance without releasing any color. Adulterated tea powder left patches of color on the paper, indicating the presence of artificial dye. In the water test, adulterated tea powder

immediately released color, turning the water dark, whereas pure tea did not show any color leaching.

As shown in the following figure 4.5, the results clearly distinguish between pure and adulterated samples, with artificial colour visibly separating in the case of adulterated tea powder.

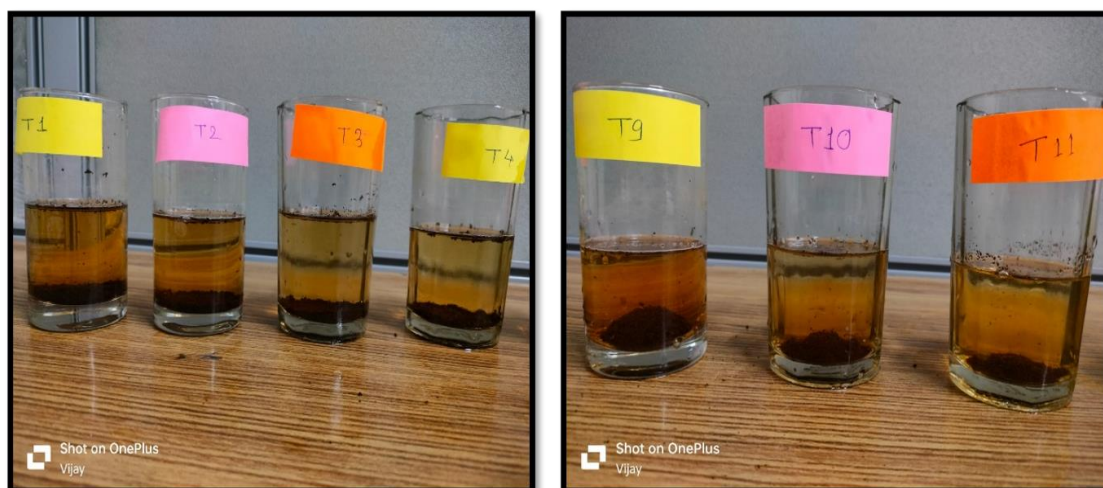


Figure 4. 5: Detection of Tea Powder Adulteration

4.3. Experiment No.2

4.3.1 Preprocessing

The ASD FieldSpec 4 spectroradiometer gives us raw hyperspectral reflectance data that covers a large range of wavelengths (350–2500 nm) and has a lot of small spectral bands. Even while this kind of data has a lot of information, it typically has spectral regions that aren't useful, noise from instruments, and changes because of scattering effects or baseline drift. So, preprocessing is a necessary step to make the data better and make the analysis that comes after it more reliable.

In this study, preprocessing was done to get rid of extra spectral bands (especially in the noisy UV and extreme NIR areas), reduce random noise, and smooth the reflectance curves so that the absorption patterns could be seen more clearly. We chose the Savitzky-Golay smoothing filter and second derivative transformation since they have been shown to work well in the past and in the literature. The Savitzky-Golay filter kept the forms of the peaks while getting rid of high-frequency noise. The second derivative, on the other hand, made small spectral features more noticeable by highlighting slope variations and getting rid of baseline offsets.

This preprocessing pipeline made the signal-to-noise ratio (SNR) a lot better and brought out the unique absorption bands that are linked to food powder adulterants. Then, the revised spectral data were used as input for classification models, which helped make the detection of adulteration more accurate and reliable.

4.3.2 Noise Removal

The ASD FieldSpec 4 spectroradiometer collects hyperspectral data that covers a wide range of wavelengths, from 350 nm to 2500 nm. This includes the visible (VIS), near-infrared (NIR), and short-wave infrared (SWIR) areas. This complete range gives a lot of information for identifying and classifying materials, but it also has spectral regions that tend to have a lot of noise because of limits in the instruments and changes in the environment. So, the first stage in the preprocessing pipeline was to find and get rid of these noisy bands. This was important to make sure the data was reliable and the model was strong.

4.3.2.1 Identification of Noisy Spectral Regions

When looking at the raw reflectance spectra of the food powder samples, it was clear that there was a lot of noise and instability at the very ends of the spectral range: 350–400 nm at the lower end and 2400–2500 nm at the upper end. People usually blame these misconceptions on the following things:

- **Limitations of Detector Sensitivity:** The silicon detectors in the ASD FieldSpec 4 are less sensitive at the low end of the VIS region (below 400 nm), which makes the signal-to-noise ratio (SNR) poor and the reflectance values unreliable.
- **Interference from the environment and the atmosphere:** When measuring light outdoors or in ambient light, scattered ultraviolet (UV) light, water vapour absorption bands, and changes in the atmosphere all add noise to the boundary areas of the spectrum.
- **Thermal and Electrical Noise:** The SWIR region's upper end (above 2400 nm) is vulnerable to thermal noise and lower lamp intensity, which causes signals to degrade and reflectance patterns to change.

4.3.2.2 Justification for Noise Exclusion

If kept, these noisy spectral regions could add statistical outliers, confuse classification systems, and hide important absorption patterns that are linked to certain adulterants. To make sure that the spectra were clear and the analysis was consistent, these bands had to be removed before any more work could be done on them.

4.3.2.3 Final Spectral Range Selection

Based on the study above, the original range of 350 to 2500 nm was shortened to a cleaner and more dependable range of 400 to 2400 nm for all the steps that followed. This new range keeps all the important absorption aspects that are needed to judge food quality while getting rid of areas that are affected by hardware problems and background noise. This choice is also in line with earlier research that suggested similar spectral limits to improve the quality of hyperspectral data [5, 6].

4.3.2.4 Impact on Data Quality

The preprocessing and model-building phases were much enhanced by the removal of the noisy areas. More consistency between samples was guaranteed, baseline drift in the derivative spectra was removed, and the signal-to-noise ratio was enhanced. The machine learning classifiers that were trained on this improved spectral dataset consequently performed better on adulteration detection tasks in terms of accuracy and stability [6].

4.3.3. Preprocessing Pipeline for Spectral Data

4.3.3.1. Noise and Irrelevant Data Removal

The spectral matrix was first edited to remove wavelength bands that continuously showed low signal-to-noise ratios in order to improve data integrity prior to preprocessing. The first 50 channels (350–399 nm) and all channels above 2000 nm were specifically eliminated because they are vulnerable to detector noise, stray light artefacts, and lamp intensity drop off [7]. The operation was carried out as follows in reality:

- Wavelength trimming: Removed extreme ends of the spectrum, first 50, and data beyond 2000 to exclude noise-prone regions.
- Operation: Remove non-informative spectral regions

```
df = df.iloc[:, np.r_[0, 52:2000]].
```

- Result: Data shape reduced from (110, 2153) to (110, 1949), focusing on relevant, high-quality spectral features.

where *df* is the hyperspectral data frame, the first index (0) retains sample identifiers, and 52:2000 selects the informative wavelength channels (400–2000 nm). This trimming step reduced the data matrix from (110 samples × 2153 bands) to (110 samples × 1949 bands), thereby concentrating subsequent analyses on high-quality spectral features while eliminating non-informative noise-dominated regions. The streamlined dataset provided a cleaner foundation for smoothing, derivative transformation, and machine-learning model training, ultimately improving classification accuracy and stability.

4.3.3.2. Savitzky-Golay Smoothing

After the removal of noisy and irrelevant spectral bands, the remaining hyperspectral data were further preprocessed using the Savitzky-Golay (SG) smoothing filter, a widely used technique in spectroscopic analysis. The main purpose of applying this filter is to reduce random noise present in each spectrum while preserving the original shape and height of the spectral peaks, which are critical for identifying subtle chemical differences between pure and adulterated samples. The Savitzky-Golay method achieves this by fitting successive subsets of adjacent data points with a low-degree polynomial using the method of least squares. This approach smoothens the spectrum without significantly distorting important spectral features such as absorption valleys and inflection points, unlike basic moving average filters [8].

The following parameters were used when using the SG filter in this study:

- There are 15 data points in the window. (For example, how many nearby points are used for each fit)
- Polynomial order: 3 (the fitting procedure makes use of a third-degree polynomial).

The selection of these parameter values was focused on striking a balance between the necessity to preserve fine spectral features pertinent to food adulteration and effective noise reduction. While too small a window may not eliminate enough noise, a wider window or lower polynomial may over smooth the spectra and obscure important characteristics.

Mathematically, the smoothing process follows Equation (1):

$$y_i = \sum_{j=-k}^k c_j y_{i+j} \dots\dots\dots (1)$$

Where:

- y_i is the smoothed value at position i ,
- c_j are the convolution coefficients derived from polynomial fitting,
- $y_i + j$ are the original reflectance values in the neighborhood,
- k is half of the window length.

The application of Savitzky-Golay smoothing significantly enhanced the signal-to-noise ratio of the spectral data, improved peak clarity, and served as a critical foundation for accurate derivative transformation and classification model training.

4.3.3.3. Baseline Correction (Second Derivative)

Even after smoothing, overlapping spectral characteristics and baseline drift can make it difficult to interpret hyperspectral data correctly, particularly in complex biological materials like food powders. In order to improve spectral resolution and perform baseline correction, a second derivative transformation was performed to the smoothed spectra.

There are two reasons to use the second derivative:

Baseline Drift Removal:

- It removes low-frequency changes in the spectrum that are brought on by instrument response, sample surface imperfections, or light scattering.
- **Peak Enhancement:** By highlighting slope variations, it draws attention to faint and overlapping spectral characteristics and makes absorption bands stand out.

The same Savitzky-Golay filter configuration (window length = 15, polynomial order = 3) that was employed for the initial smoothing was used to carry out this change. This method ensures both derivative accuracy and noise reduction by computing the derivative as a convolution over local polynomial fits. In terms of mathematics, the second derivative gives a measure of the spectral profile's curvature, which improves the ability to distinguish between clean and contaminated samples.

It is vital for drawing attention to subtle variations in spectral response, which are necessary for creating accurate classification models. This derivative preprocessing

technique has been successfully applied in spectral analysis and food quality evaluation, and it is well known in chemometrics for enhancing classification performance

$$f''(x) = \frac{d^2y}{dx^2} \dots\dots\dots (2)$$

4.3.3.4. Standard Scaling (Z-score Normalization)

Sample-wise intensity changes in hyperspectral data may persist after smoothing and baseline correction because of variations in particle size, surface roughness, lighting, or instrument response. By prioritizing magnitude over shape in spectral characteristics, these variances have the potential to skew classification models.

Z-score normalization, sometimes referred to as standard scaling, was applied to every spectrum in order to address this problem. By using this method, the data is transformed so that the mean of each wavelength (spectral variable) is zero and the standard deviation is one for all samples. The following formula is applied:

$$Z = \frac{X-\mu}{\sigma} \dots\dots\dots (3)$$

Where:

- X is the original reflectance value,
- μ is the mean reflectance at that wavelength across all samples,
- σ is the standard deviation.

By standardizing, all spectral variables are guaranteed to add equally to the learning process. This stops features with higher absolute values from taking over the model. It also makes many machine learning methods, like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Neural Networks work better and get better results. Z-score normalization is a common and useful step before using chemometric modeling to analyze the quality of food. It makes sure that decisions about classification are based on relative spectral patterns instead of absolute intensity differences [9].

4.3.3.5. Final Noise Reduction Pass

There may still be high-frequency noise in the hyperspectral data after earlier steps like wavelength trimming, initial Savitzky-Golay smoothing, derivative

transformation, and standard scaling. A final noise reduction pass was done to improve the signal clarity even more before feature extraction and classification. In this step, a second low-order Savitzky-Golay smoothing filter with a window length of only 11 points and a polynomial order of 2 was used. This setup was picked so that it would be possible to get rid of fine noise without changing narrow absorption peaks or important spectral curve features. The first smoothing operation used a larger window to deal with broad-spectrum irregularities. This pass, on the other hand, focused on localized fluctuations and small leftover inconsistencies that may have shown up or become more noticeable during derivative transformation or normalization. This last step improved the clarity and uniformity of each processed spectrum. This made sure that the machine learning classifiers got clean, denoise inputs that kept all the analytically important spectral features. This kind of small-step improvement is very important for high-dimensional data like hyperspectral signals, where even small noise artifacts can have a big effect on the edges of classification and the generalization of models.

- Purpose: A final low-order smoothing to reduce any residual noise.
- Method: Additional Savitzky-Golay smoothing with a smaller window (11, order 2) to suppress residual noise.

4.4. Preprocessing Results

The preprocessing of chili powder spectral data markedly enhances the clarity and utility of the signals for categorization. In Figure 8 (X: Before Preprocessing), the unprocessed spectral fingerprints exhibit fluctuations in reflectance across several wavelengths, accompanied by discernible noise, baseline deviations, and overlapping patterns. These factors complicate the correct comparison of spectral differences across chili powder brands. Baseline drift, resulting from discrepancies in measurement settings, and noise, which adds undesirable fluctuations, hide significant spectrum patterns.

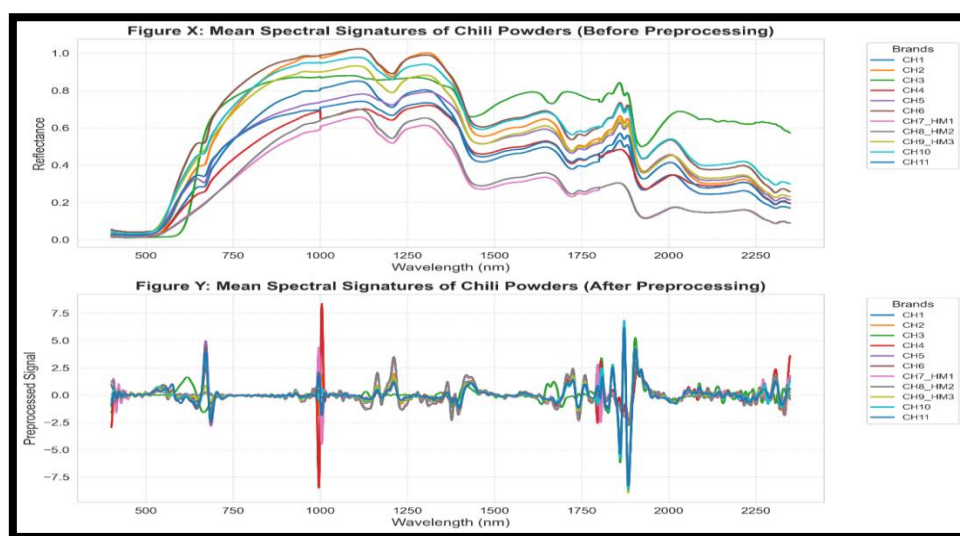


Figure 4.6: Raw Spectral Signatures Before Preprocessing and Mean Spectral Signatures After Preprocessing

Figure 4.6 X: Displays average reflectance curves for each brand before preprocessing. Observation: Clear baseline variations and overlapping signals make brand differentiation difficult.

Figure 4.6 Y: Shows spectra after all preprocessing steps after preprocessing. Observation: Signals are centered, noise is reduced, and distinct spectral features emerge.

Following the use of preprocessing procedures, Figure 8 (Y: Post-Preprocessing) displays a much cleaner and more organized depiction of the spectral data. The Savitzky-Golay smoothing method efficiently reduces noise while maintaining significant peaks, and baseline correction via the first derivative eliminates undesirable shifts, aligning all spectra to a uniform baseline. Standard scaling guarantees uniformity across samples, hence improving comparability across various chili powder brands. A further phase of smoothing significantly enhances the data, removing remaining noise and accentuating peaks. The following figure 9 presents a direct comparison of the raw and preprocessed spectral data. It unequivocally illustrates how preprocessing amplifies spectral characteristics while mitigating undesirable fluctuations. The comments emphasize enhancements like improved features, less noise, and minimal baseline effects, making the data more appropriate for machine learning categorization.

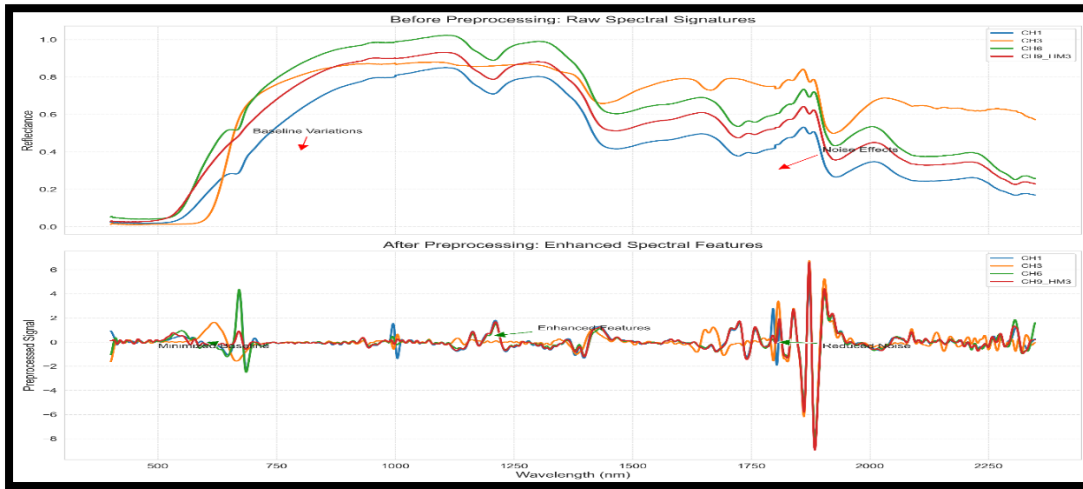


Figure 4.7: Before-After Comparison of Spectral Features

Figure 4.8. This graphic demonstrates how preprocessing enhances spectral clarity for the brand CH3. The preprocessed signal (red) exhibits an elevated signal-to-noise ratio and improved spectral peaks relative to the raw data (blue). This signifies that preprocessing effectively amplifies significant information while reducing extraneous noise. This graphic emphasizes the enhancements in spectral properties within a given wavelength range. The enhanced signal-to-noise ratio facilitates superior feature extraction, resulting in increased classification accuracy. Locations in the spectrum have been marked to illustrate how preprocessing improves specific spectral features.

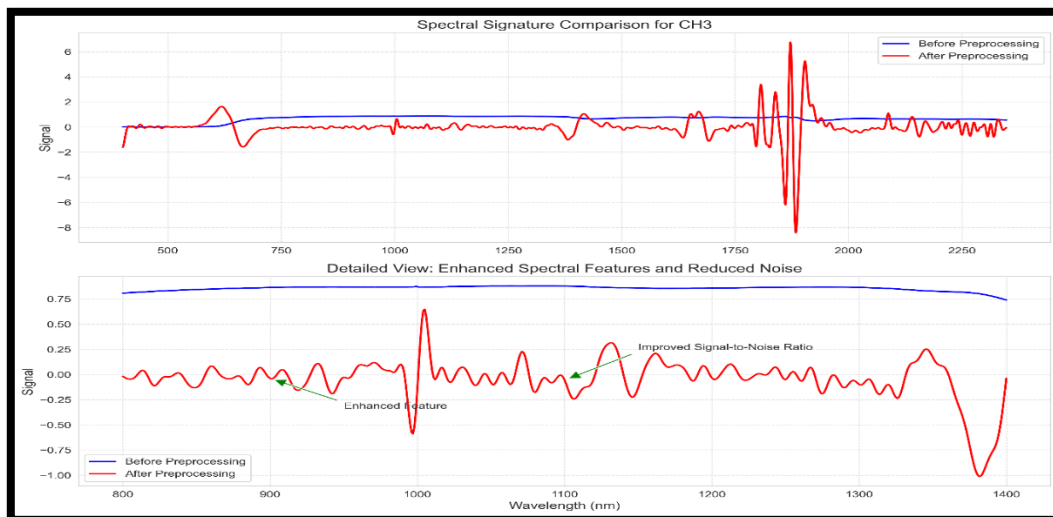


Figure 4.8: Spectral Signature Comparison for CH3 and Detailed View of Enhanced Spectral Features

The preprocessing of turmeric powder spectral data markedly improves the clarity and applicability of the spectral signatures for classification purposes. In Figure 4.9 (X: Before Preprocessing), the unprocessed spectral fingerprints exhibit significant variations in reflectance across different wavelengths, accompanied by prominent noise, baseline drift, and overlapping patterns. The discrepancies hinder precise comparisons of spectral changes among various turmeric powder brands. Baseline drift, resulting from fluctuations in measuring circumstances, and random noise conceal essential spectral characteristics,

Complicating brand distinction. Following the use of preprocessing methods, Figure 4.9 (Y: After Preprocessing) exhibits a much cleaner and more organized representation of the spectral data. The Savitzky-Golay smoothing filter efficiently reduces noise while maintaining critical spectral peaks, and baseline correction using the first derivative eliminates undesirable shifts, standardizing all spectra to a uniform baseline. Moreover, standard scaling guarantees consistency across all samples, improving comparability across various turmeric powder manufacturers. A supplementary smoothing procedure enhances the data, reducing residual noise and highlighting important spectral characteristics, hence increasing classification precision.

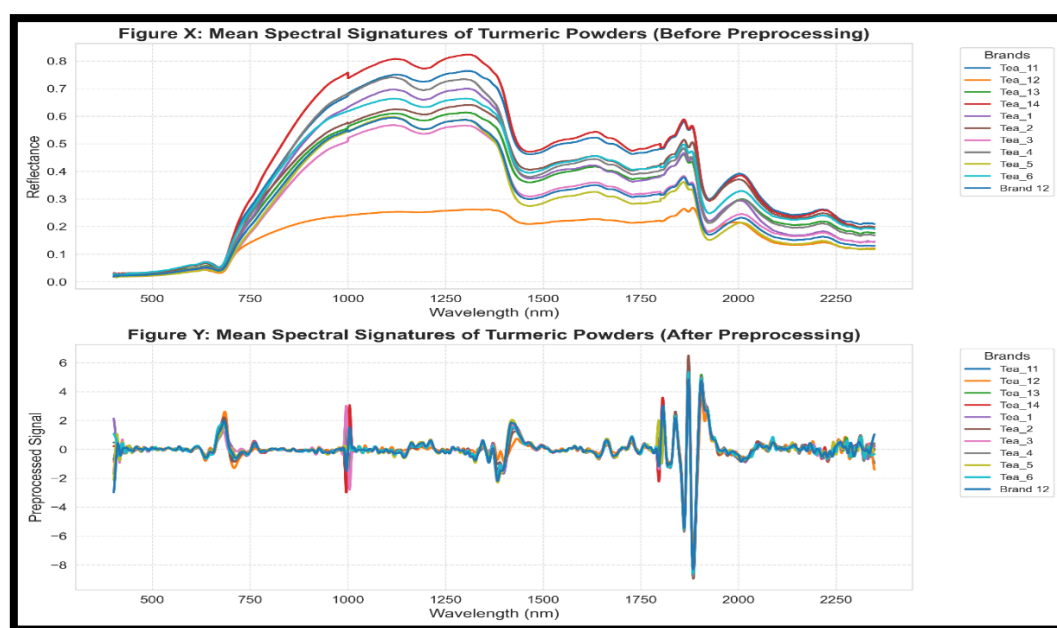


Figure 4.9: Raw Spectral Signatures Before Preprocessing and Mean Spectral Signatures After Preprocessing of Turmeric

Figure 4.10 shows a direct comparison of raw and preprocessed spectral data for turmeric powder. The top figure (Before Preprocessing) displays unrefined spectral signatures with significant baseline fluctuations and noise interference, which may mask relevant spectral patterns. The bottom plot (After Preprocessing) illustrates how preprocessing approaches effectively mitigate baseline shifts, enhance spectral characteristics, and reduce noise, resulting in clearer and more discernible spectral patterns. These enhancements provide the data more dependable for the categorization and analysis of various turmeric powder samples.

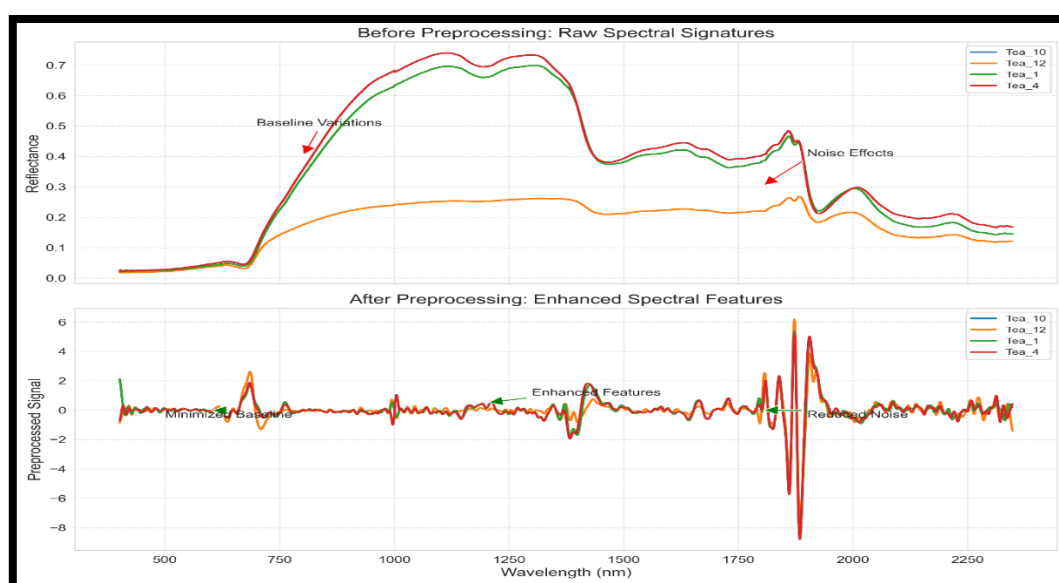


Figure 4.10: Before-After Comparison of Spectral Features

This figure 4.11 illustrates how preprocessing improves the spectral clarity of TP3-Home_Made turmeric powder. The upper figure juxtaposes the raw spectrum data (blue) with the preprocessed signal (red) throughout the whole wavelength range, demonstrating a significant enhancement in spectral characteristics. The preprocessed data has an elevated signal-to-noise ratio, fewer baseline fluctuations, and amplified spectral peaks, making essential aspects more discernible. The bottom graphic offers a comprehensive examination of a particular wavelength range, emphasizing the impact of preprocessing on essential spectral areas. The indicated spots in the bottom figure demonstrate how preprocessing enhances significant spectrum fluctuations while attenuating noise, hence augmenting data interpretability. These changes enable more precise feature extraction, resulting in improved classification and discrimination of turmeric powder samples.

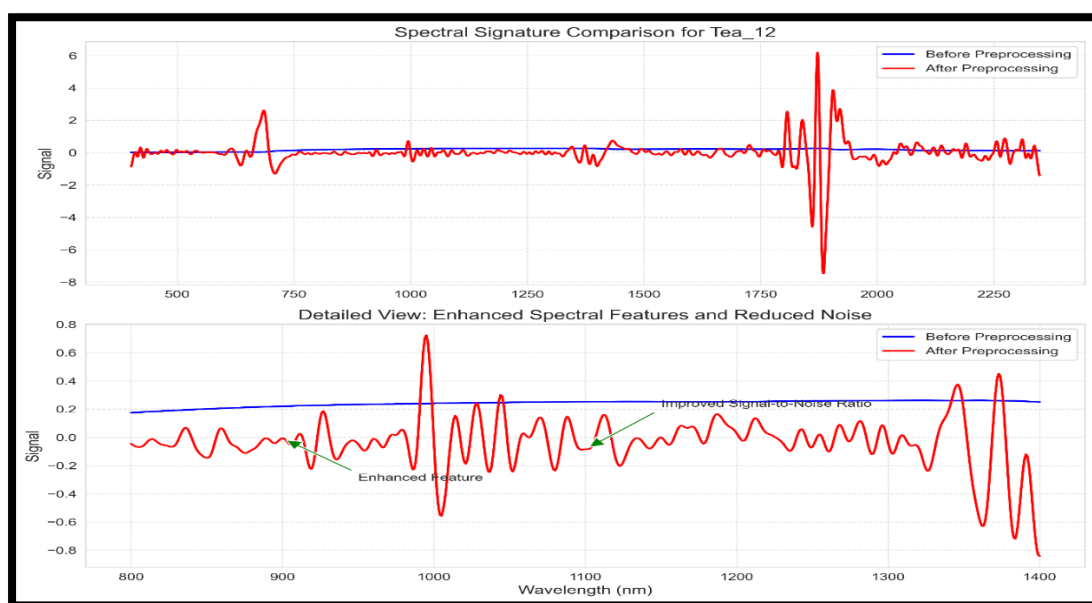


Figure 4.11: Spectral Signature Comparison for TP3 and Detailed View of Enhanced Spectral Features

The preprocessing of tea powder spectral data markedly improves the clarity and interpretability of spectral signatures for classification purposes. Prior to preprocessing Figure 14 (X: Before Preprocessing) the unrefined spectral fingerprints exhibit significant fluctuations in reflectance across several wavelengths, characterized by noticeable noise, baseline drift, and overlapping patterns, complicating the differentiation between distinct tea brands. Baseline drift, resulting from fluctuations in measurement circumstances, obscures essential spectral traits, while random noise further complicates brand separation by concealing vital spectral characteristics. Following the use of preprocessing procedures Figure 4.12 (Y: After Preprocessing) the spectral signatures exhibit enhanced structure and interpretability [10]. The Savitzky-Golay smoothing filter efficiently reduces noise while maintaining critical spectral peaks, and baseline correction via the first derivative eliminates undesirable shifts, guaranteeing that all spectra conform to a consistent reference baseline. Furthermore, standard scaling normalizes the data, improving comparability across different tea brands. A supplementary smoothing procedure enhances the data by reducing residual noise and accentuating notable spectral characteristics. These advances jointly augment classification accuracy, facilitating more accurate identification across tea powder brands.

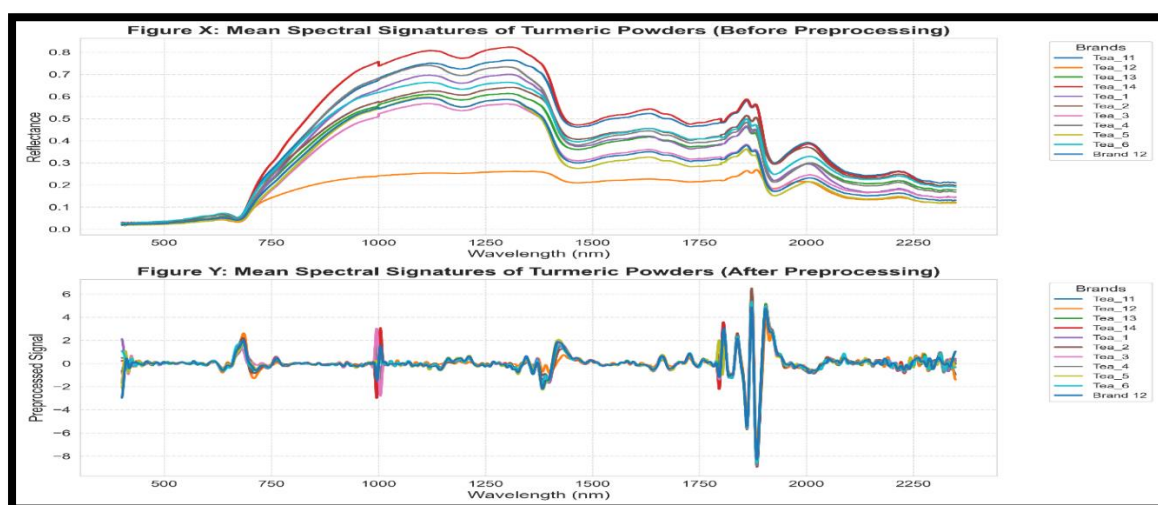


Figure 4.12: Raw Spectral Signatures Before Preprocessing and Mean Spectral Signatures After Preprocessing

Figure 4.13 shows a direct comparison of raw and preprocessed spectral data for tea particles. The upper panel (Before Preprocessing) displays unrefined spectral signatures, characterized by considerable baseline fluctuations and noise interference that mask critical spectral characteristics. The irregularities hinder the differentiation of various tea brands and diminish the reliability of spectral categorization. The lower plot (After Preprocessing) emphasizes the effects of preprocessing approaches. Baseline shifts are mitigated, spectral characteristics are amplified, and noise interference is markedly reduced. These enhancements provide clearer and more discernible spectrum patterns, hence increasing the reliability of the data for tea brand categorization and analysis.



Figure 4.13: Before-After Comparison of Spectral Features

This figure 4.14 illustrates how preprocessing enhances the spectral clarity of Tea_12 tea powder. The top figure juxtaposes the raw spectrum data (blue) with the preprocessed signal (red) throughout the whole wavelength range, demonstrating a significant improvement in spectral characteristics. The preprocessed data demonstrates an enhanced signal-to-noise ratio, less baseline fluctuations, and intensified spectral peaks, hence making essential spectral features more discernible. The bottom graphic offers a comprehensive examination of a particular wavelength range, highlighting the influence of preprocessing on critical spectral areas. The shown points demonstrate how preprocessing amplifies notable spectrum fluctuations while reducing noise, hence enhancing data interpretability. These advancements provide more accurate feature extraction, resulting in improved classification and differentiation of various tea powder samples.

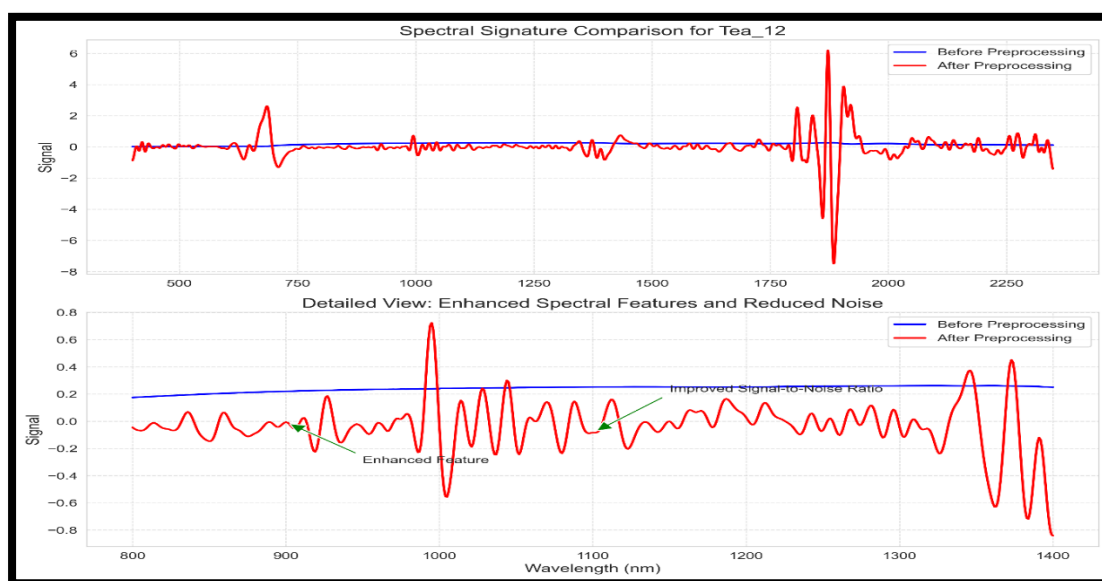


Figure 4.14: Spectral Signature Comparison for Tea_12 and Detailed View of Enhanced Spectral Features

The preprocessing advantages shown in the figures include a substantial decrease in baseline fluctuations, hence providing more uniform spectral signatures across various samples. Moreover, spectral characteristics are amplified, making critical peaks and troughs more pronounced for greater classification precision. The adverse impact of noise, which previously compromised the raw data, is mitigated, resulting in a more refined and dependable spectrum representation. Moreover, the

signal-to-noise ratio is enhanced, facilitating better distinction between spectral features and diminishing random variations that may affect data interpretation. These enhancements jointly elevate the quality of the spectral data, making it more appropriate for machine learning and classification endeavors.

4.5. Feature Extraction from Spectral Data

4.5.1. Statistical Feature

It is useful to compute descriptive statistical features in order to decrease the complexity of hyperspectral data for machine learning applications and identify meaningful patterns from it. The distribution, variance, and form of each sample's spectral reflectance values across wavelengths are summed up by these statistical indicators. The mean, standard deviation, skewness, kurtosis, and interquartile-based metrics like Q25, Q75, and IQR are the most often utilized features.

4.5.1.1 Mean (μ)

The **mean** represents the average spectral reflectance of a sample across all wavelengths. It provides a central tendency measure, giving an overall idea of the reflectance intensity.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad \dots\dots\dots (4)$$

Where:

- x_i is the reflectance value at wavelength i
- n is the number of spectral bands (e.g., 130 or 1949, depending on preprocessing),
- μ is the average reflectance of the sample.

The mean helps differentiate between brighter (higher reflectance) and darker (lower reflectance) samples, which can relate to compositional differences in food powders.

4.5.1.2. Standard Deviation (σ)

The standard deviation quantifies the spread or variability of reflectance values around the mean. A higher standard deviation indicates greater variability, which could be due to noise, complex material composition, or spectral mixing.

4.5.1.3. Skewness

Skewness measures the asymmetry of the reflectance distribution. A skewness of zero indicates a symmetric distribution.

- **Positive skew:** The right tail (higher values) is longer, indicating higher reflectance values.
- **Negative skew:** The left tail (lower values) is longer, indicating more low reflectance values.

Skewness can indicate irregular reflectance behavior due to contamination or varying particle size.

4.5.1.4. Kurtosis

Kurtosis quantifies the peakedness of the spectral distribution.

- A normal distribution has a kurtosis of 3 (mesokurtic).
- High kurtosis (>3) indicates a sharp peak (leptokurtic),
- Low kurtosis (<3) indicates a flatter distribution (platykurtic).

Kurtosis helps identify whether reflectance values are concentrated near the mean or have heavy tails, which may relate to the homogeneity of the sample.

4.5.1.5. Quartiles and Interquartile Range

Q25 (First Quartile): The value below which 25% of the spectral data falls.

Q75 (Third Quartile): The value below which 75% of the spectral data falls.

IQR is calculated as:

$$\text{IQR} = \text{Q75} - \text{Q25}$$

IQR is a robust measure of dispersion, less sensitive to outliers than standard deviation. It effectively captures the range within which the middle 50% of spectral data lies. In adulteration detection, a narrow IQR might indicate spectral uniformity, whereas a wide IQR may suggest mixture variability. These statistical measures serve as global descriptors of spectral behavior and are frequently used to convert high-dimensional spectral signatures into a compact set of numerical features for classification models. Their advantage lies in being computationally efficient and interpretable, making them ideal for integrating with machine learning algorithms in adulteration detection tasks.

4.5.2. Signal-Based Features

In addition to statistical descriptors, signal-based features play a crucial role in capturing the intrinsic structure and dynamics of hyperspectral reflectance curves. These features are extracted by analyzing the shape, complexity, and variation in the spectral signal across wavelengths, rather than summarizing values based on distribution alone. Signal-based features help preserve important patterns that statistical features may overlook, especially in the context of food adulteration, where subtle changes in spectral behavior often indicate the presence of contaminants.

4.5.2.1 Signal Energy

Signal energy is a measure of the **overall strength or magnitude** of the spectral reflectance signal. It is computed as the sum of the squares of the reflectance values. The most relevant signal-based features include:

$$\text{Energy} = \sum_{i=1}^n x^2_i \quad \dots\dots\dots (5)$$

A higher energy value may correspond to materials with intense spectral responses, while a lower energy value may suggest the presence of low-reflectance adulterants or diluents.

4.5.2.2. Entropy

Entropy quantifies the complexity or unpredictability in the spectral signal. It reflects how evenly distributed or random the reflectance values are across the spectrum:

$$\text{Entropy} = \sum_{i=1}^n p_i \log (p_i)^2 \dots\dots\dots (6)$$

Where p_i is the probability of the reflectance value at wavelength i . A highly ordered (pure) sample will have low entropy, while a mixed or adulterated sample often shows higher entropy due to spectral irregularity.

4.5.2.3. Spectral Slope and First Derivative Features

The first derivative of the spectrum captures the rate of change in reflectance with respect to wavelength. Signal slopes highlight transition zones and absorption edges. Analyzing slope patterns helps in distinguishing materials with similar reflectance levels but different absorption behaviors.

4.5.2.4. Area Under the Curve (AUC)

The area under the spectral curve represents the integrated reflectance over a given spectral range. It can serve as a cumulative index of reflectance behavior:

$$\text{AUC} = \int_{\lambda_1}^{\lambda_2} x(\lambda) d\lambda \quad \dots\dots\dots (7)$$

This is useful in quantifying broad absorption bands that may correspond to moisture, colorants, or specific chemical groups. Signal-based features are particularly valuable when dealing with subtle or non-linear adulteration patterns. While statistical features capture general trends, signal-based descriptors provide insights into how the spectral signal changes offering improved sensitivity to adulterants that affect specific absorption regions or alter the signal structure without significantly shifting the mean or variance.

By combining signal-based and statistical features, a more comprehensive and discriminative feature set can be constructed, enhancing the accuracy of machine learning models in food adulteration detection.

4.5.3. Peak and Valley Metrics

Peak and valley measures are important signal-based features that are taken from hyperspectral reflectance data to show where and how strong key absorption and reflection features are in the spectral curve. These measurements are very sensitive to the chemical make-up and physical structure of the substance. They work especially well for finding adulteration in food powders because adulterants can add or remove certain spectral features.

4.5.3.1. Definition and Importance

- A peak is a local maximum in the spectral reflectance curve. These are usually linked to reflectance properties like material brightness or surface scattering.
- A valley is a local minimum that is usually caused by absorption bands in the sample, which are caused by certain chemical bonds or components.

These features reflect material-specific interactions with light and are useful for:

- Find the diagnostic bands that are linked to known contaminants or natural ingredients.
- To get an idea of how strong the spectral absorption is, measure the absorption depth (the distance between the peak and the next valley).

- Looking at peak width and symmetry can show how pure or complicated the material is.

4.5.3.2. Common Metrics Extracted

- Peak point (wavelength where reflectance is highest)
- Valley location (wavelength where reflectance is lowest)
- Peak height or valley width
- The full width at half maximum (FWHM) is the width of the hill or valley when it is half as high or deep as it is at its widest point.
- How much a hill or valley stands out from the rest of the landscape.

In the setting of finding food that has been tampered with: Impurities may be present if the clarity of the peaks decreases, or the position of the valleys moves. Synthetic dyes and filler powders are examples of adulterants that can change the natural peaks or slopes by changing the way they are absorbed. Tracking changes in peak-valley characteristics between known adulterated and pure data makes it possible to classify and find patterns of adulteration. Most of the time, peak recognition algorithms or spectral derivative analysis are used to get these features. They are then used as reliable inputs for machine learning models.

4.5.4. PCA for Feature Extraction

Objective: Reduce dimensionality while retaining variance in the spectral features.

Principal Component Analysis (PCA):

Projects high-dimensional data into a lower-dimensional subspace

Components are orthogonal (uncorrelated) directions of maximum variance

Hyperspectral data typically consist of hundreds to thousands of closely spaced spectral bands, resulting in high dimensionality, data redundancy, and increased computational complexity. To address these challenges, Principal Component Analysis (PCA) was employed as a feature extraction technique. PCA transforms the original high-dimensional dataset into a new set of orthogonal (uncorrelated) axes known as principal components (PCs). These components are linear combinations of the original spectral variables, ordered by the amount of variance they capture from the data:

- First principal component (PC1) captures the maximum variance possible.

- Subsequent components (PC2, PC3, ...) capture the remaining variance in decreasing order, subject to orthogonality constraints.

This transformation projects the data into a lower-dimensional subspace, reducing the number of features while retaining the underlying structure and class-separating information. The number of retained principal components is typically determined based on a threshold of cumulative explained variance. PCA also helps in removing multicollinearity between spectral bands, enhancing model training speed, and improving classification performance. It is widely used in hyperspectral analysis for both visualization and dimensionality reduction tasks [11,12].

We use feature extraction using Principal Component Analysis (PCA) to diminish the dimensionality of hyperspectral data while retaining the most essential information for classification. Using PCA, we convert high-dimensional spectral information into a lower-dimensional space, encapsulating the most substantial variances in the data. This method improves the model's capacity to differentiate among various brands of tea, turmeric, and chili by recognizing essential spectral features. The retrieved characteristics, including statistical qualities such as mean, standard deviation, skewness, kurtosis, energy, and area under the curve (AUC), enhance classification accuracy. The PCA findings reveal clear clustering patterns, underscoring the efficacy of this technology in food authenticity and contamination identification.

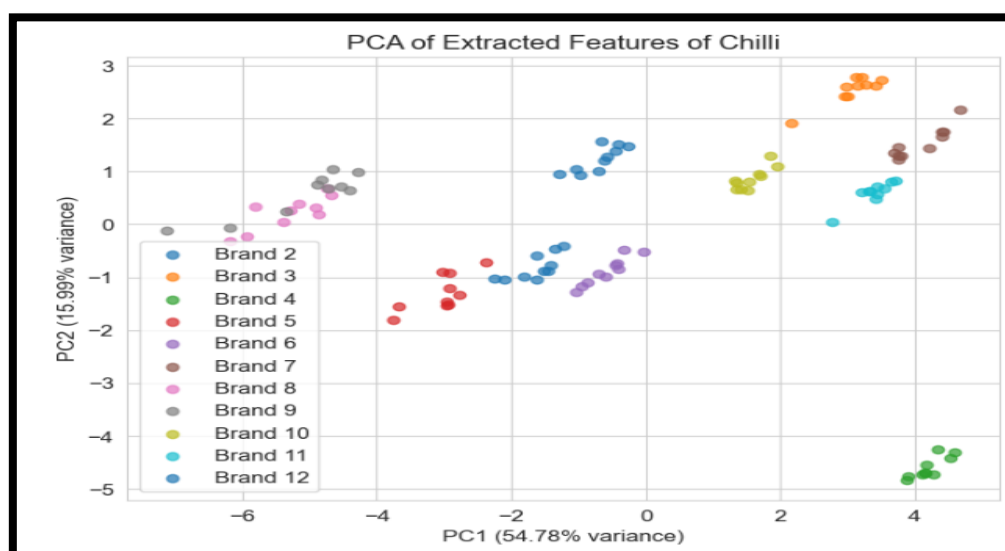


Figure 4.15: Principal Component Analysis (PCA) of Extracted Features for Chili

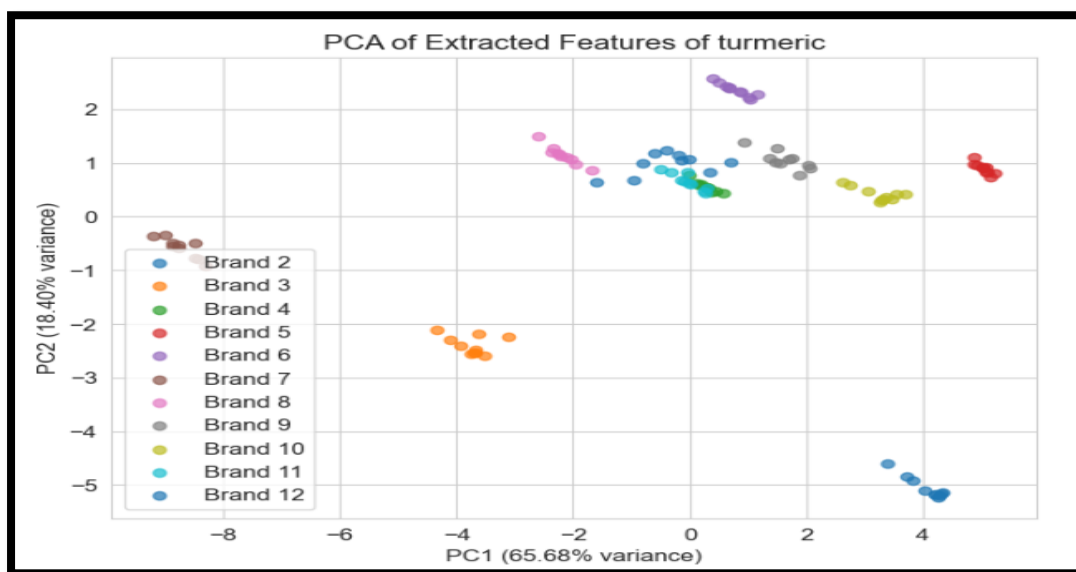


Figure 4.16: Principal Component Analysis (PCA) of Extracted Features for Turmeric

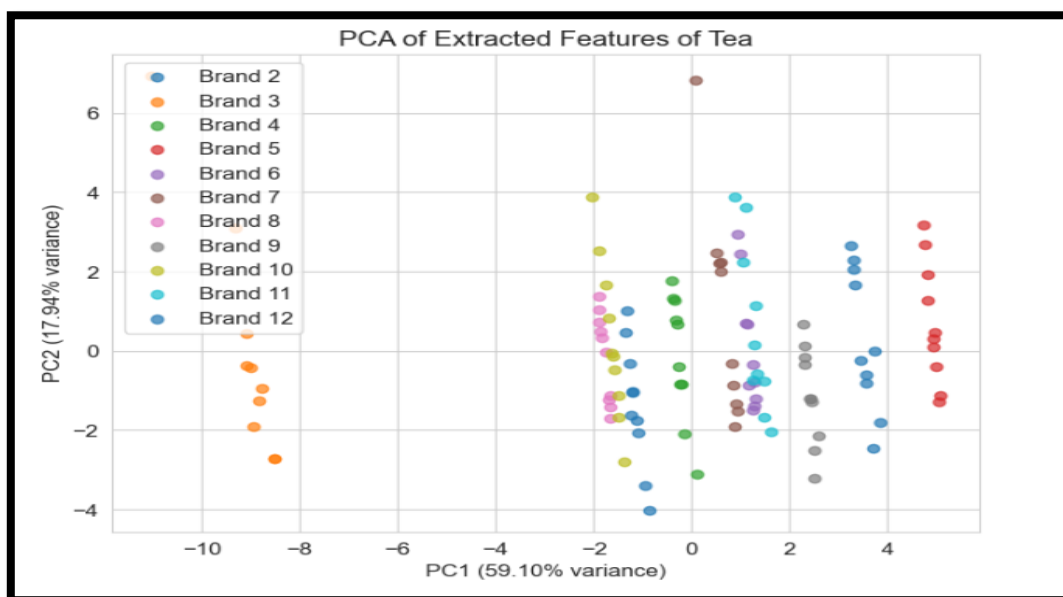


Figure 4.17: Principal Component Analysis (PCA) of Extracted Features for Tea

The PCA analysis of retrieved features for chili, turmeric, and tea elucidates significant insights into the spectral variances across various brands. The PCA scatter plot (figure17) for chili demonstrates a distinct separation across brands, indicating considerable changes in spectral features. PC1 elucidates 54.78% of the variation, whilst PC2 represents 15.99%, together including a substantial fraction of the data variability. The distinct clusters indicate that the derived features successfully discriminate chili samples according to their spectral characteristics. The PCA plot

(figure18) for turmeric indicates a robust variance explanation, with PC1 accounting for 65.68% and PC2 providing 18.40%. This signifies that the spectral characteristics exhibit significant discriminative capability, allowing distinct differentiation among various turmeric brands. The well delineated clusters indicate significant spectrum disparities, perhaps affected by variances in composition, purity, or contamination levels. The pronounced distinction in PCA space underscores the efficacy of the derived features in classification tasks for turmeric powder. The PCA scatter plot (figure19) for tea has a more scattered clustering pattern, indicating a greater degree of spectral overlap across brands. PC1 elucidates 59.10% of the variation, while PC2 represents 17.94%, including a significant percentage of the spectrum diversity. Nevertheless, several brands have overlapping clusters, indicating that more preprocessing or sophisticated classification methods may be required to improve brand distinction in tea samples.

The data demonstrate that chili and turmeric have superior separability in PCA space relative to tea, indicating that their spectral properties are more unique across various brands. Turmeric has the greatest variance explanation in PC1, indicating that a limited number of principal components adequately account for most of the spectral variability. The overlap seen in tea samples indicates that more refining of feature extraction approaches may be required for better discrimination. The derived features, including statistical qualities like mean, skewness, kurtosis, energy, AUC, and peak values, are essential for PCA-based separation, illustrating the efficacy of the suggested feature extraction method.

4.6. Classification

Following PCA) on the obtained spectral data, the samples are then classified using several machine learning techniques. Trained and assessed on the reduced feature set produced by PCA were four classifiers including Random Forest (RF), Support Vector Machine (SVM), k-Nearest Neighbours (k-NN) and XGBoost. Confusion matrices provided information on model performance; the classification models were evaluated using accuracy, precision, recall, and F1-score. With little misclassification, Random Forest and SVM showed the best performance among the tested models, deftly separating clean from contaminated data. The results confirm that PCA-transformed features, when combined with strong machine learning models, may significantly

enhance the identification and classification of food samples; therefore, this technique is suitable for practical food authenticity and quality control uses.

Figure 4.17 illustrates the training and testing accuracy of four machine learning models for chili categorizations, emphasizing the better efficacy of ensemble models. The Random Forest model achieved perfect classification, with both training and testing accuracy at 1.0000, making it the most dependable model. XGBoost exhibited exceptional performance, achieving a training accuracy of 1.0000 and a test accuracy of 0.9697, indicating robust generalization and little overfitting. KNN (k=5) had a training accuracy of 0.9091 and a test accuracy of 0.8485, indicating a modest level of overfitting. The SVM (RBF Kernel) achieved the lowest test accuracy of 0.8182, making it less appropriate for this dataset. The findings indicate that ensemble models like as XGBoost and RF surpass conventional models, providing enhanced generalization and accuracy for chili categorizations.

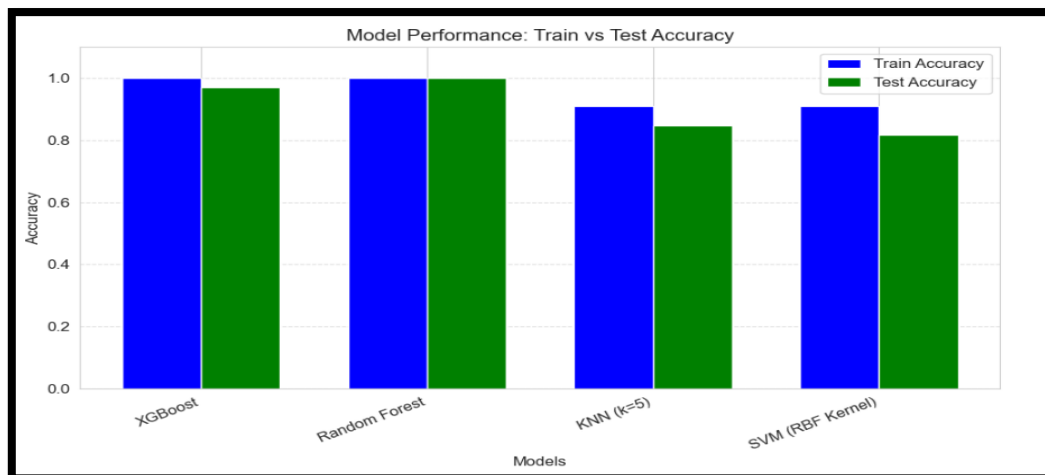


Figure 4.18: Training and Testing Accuracy of Four Machine Learning Models for Chilli Classification

The comparative comparison of performed models Table 4.1 for chili categorizations, evaluated on precision, recall, F1-score, and accuracy after PCA-based feature extraction, underscores the advantages of ensemble models. The RF model surpassed all others, attaining 100% accuracy, so establishing itself as the most efficient classifier. XGBoost exhibited remarkable performance with an accuracy of 97%, while it had somewhat worse recall for certain classes. Conversely, KNN (85%) and SVM (82%) showed reduced accuracy, suggesting that distance-based and kernel-based approaches may not generalize as successfully for this dataset. The results indicate that

RF and XGBoost are the most appropriate models for chili classification, providing superior accuracy, robustness, and generalization compared to KNN and SVM.

Table 4.1: Comparative Performance of Machine Learning Models for Chili Classification

Model	Precision	Recall	F1-Score	Accuracy
XGBoost	0.98	0.97	0.97	97%
Random Forest	1.00	1.00	1.00	100%
KNN (k=5)	0.86	0.85	0.85	85%
SVM (RBF Kernel)	0.83	0.82	0.82	82%

The figure 4.18 confusion matrices evaluate the efficacy of XGBoost, RF, KNN (k=5), and SVM (RBF Kernel) in chili classification. The Random Forest model achieved flawless classification with 100% accuracy, making it the superior model. XGBoost exhibited outstanding performance with little misclassifications. KNN (k=5) and SVM had a higher error rate, with SVM proving to be the least successful owing to many misclassifications. In summary, ensemble models (XGBoost and RF) surpassed KNN and SVM, demonstrating superior reliability for chili classification.

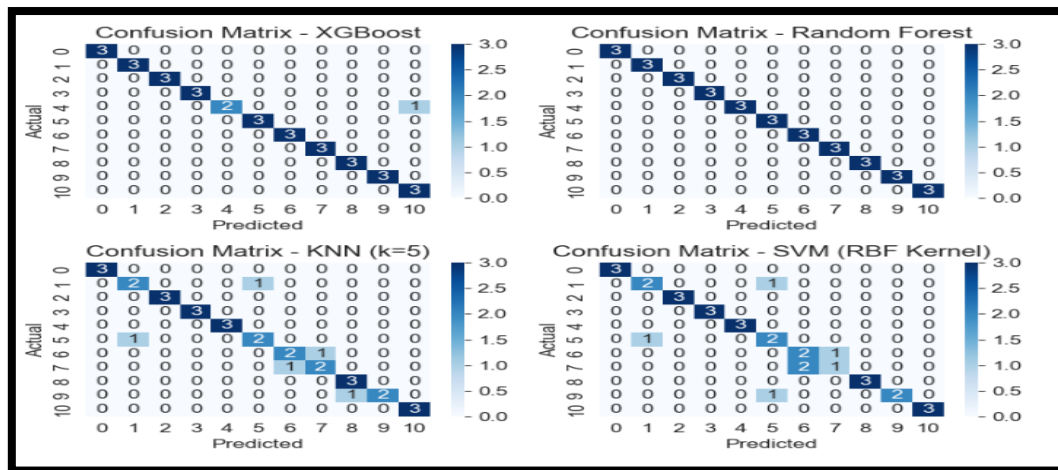


Figure 4.19: Confusion Matrices of Machine Learning Models for Chili Classification

The figure 4.19 illustrates the training and testing accuracy of four machine learning models for turmeric classification, highlighting the superior performance of ensemble models. The Random Forest model achieved perfect classification, with both training and testing accuracy at 1.0000, making it the most reliable model. XGBoost

also demonstrated excellent performance, attaining a training accuracy of 1.0000 and a test accuracy of 0.9697, showcasing strong generalization with minimal overfitting. KNN ($k=5$) recorded a training accuracy of 0.9091 and a test accuracy of 0.8485, indicating some degree of overfitting. The SVM (RBF Kernel) had the lowest test accuracy of 0.8182, suggesting that it may not be the best choice for turmeric classification. Overall, the results indicate that ensemble models like XGBoost and Random Forest outperform traditional models, offering better generalization and accuracy for turmeric classification.



Figure 4.20: Training and Testing Accuracy of Four Machine Learning Models for Turmeric Classification

The comparative analysis of machine learning models for turmeric classification, evaluated using precision, recall, F1-score, and accuracy after PCA-based feature extraction, highlights the superior performance of ensemble models Table 4.2. The Random Forest model outperformed all others, achieving 100% accuracy, making it the most effective classifier. XGBoost demonstrated strong performance with a test accuracy of 93.94%, though it had slightly lower recall for certain classes. In contrast, KNN (90.91%) and SVM (84.85%) showed lower accuracy, suggesting that distance-based and kernel-based methods may not generalize as effectively for this dataset. These findings indicate that Random Forest and XGBoost are the most suitable models for turmeric classification, offering higher accuracy, robustness, and better generalization compared to KNN and SVM.

Table 4.2: Comparative Performance of Different Machine Learning Models for Turmeric Classification

Model	Precision	Recall	F1-Score	Accuracy
XGBoost	0.96	0.94	0.93	93.94%
Random Forest	1.00	1.00	1.00	100.00%
KNN (k=5)	0.94	0.91	0.91	90.91%
SVM (RBF Kernel)	0.89	0.85	0.85	84.85%

The figure 4.20 presents confusion matrices assessing the performance of XGBoost, RF, KNN (k=5), and SVM (RBF Kernel) for turmeric classification. The RF model achieved perfect classification with 100% accuracy, confirming its superiority. XGBoost performed exceptionally well with only minor misclassifications. In contrast, KNN (k=5) and SVM exhibited higher misclassification rates, with SVM being the least effective due to multiple incorrect predictions. Overall, ensemble models (XGBoost and RF) outperformed KNN and SVM, proving to be the most reliable for turmeric classification.

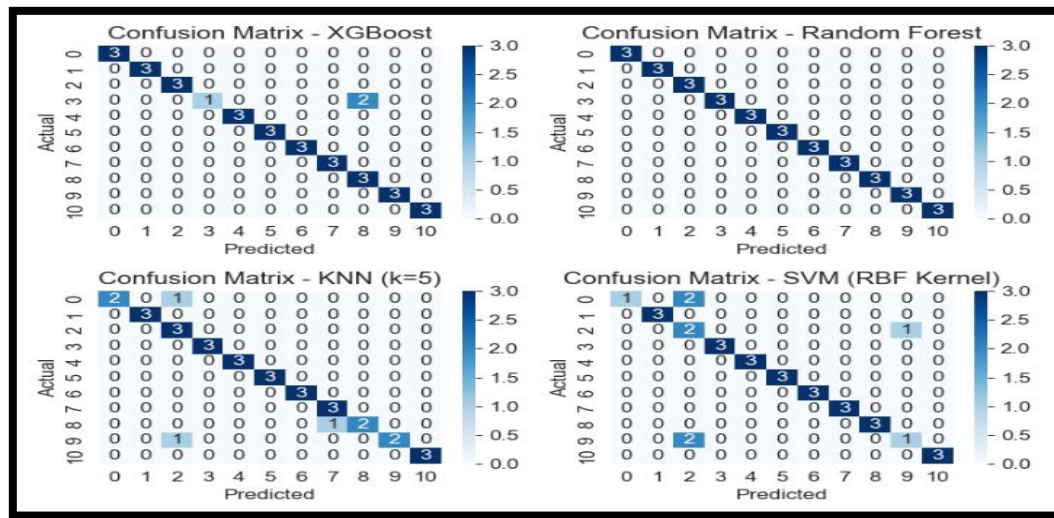


Figure 4.21: Confusion Matrices of Machine Learning Models for Turmeric Classification

The figure 4.21 illustrates the training and testing accuracy of four machine learning models for tea classifications, emphasizing the superior performance of ensemble models. The Random Forest model achieved perfect classification, with both training and testing accuracy at 1.0000, making it the most reliable. XGBoost also performed exceptionally well, with a training accuracy of 1.0000 and a test accuracy of

0.9697, demonstrating strong generalization and minimal overfitting. KNN (k=5) showed moderate overfitting, achieving a training accuracy of 0.9091 and a test accuracy of 0.8485. The SVM (RBF Kernel) recorded the lowest test accuracy (0.8182), indicating that it is less suitable for this dataset. These results confirm that ensemble models (XGBoost and RF) outperform traditional models, providing superior accuracy and generalization for tea classification.

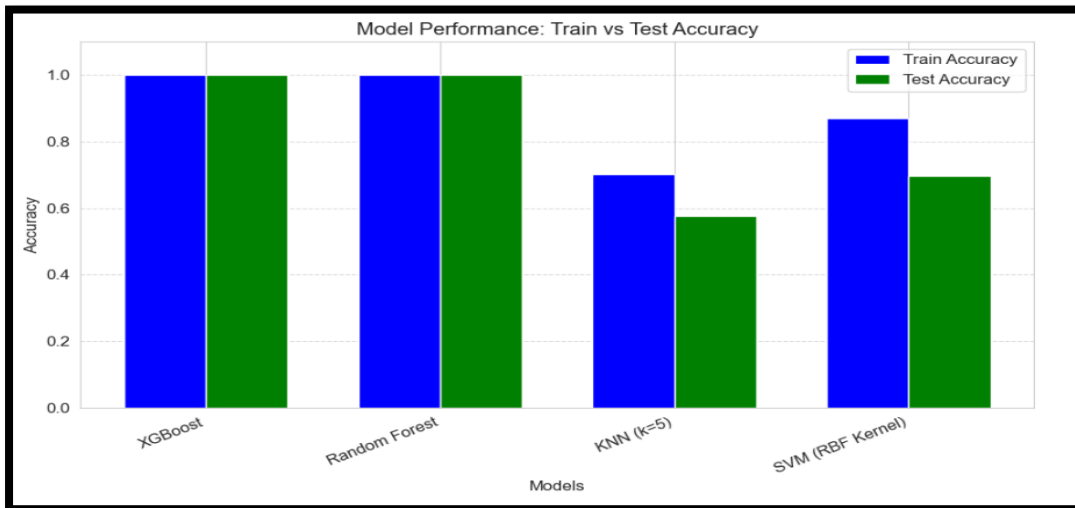


Figure 4.22: Training and Testing Accuracy of Four Machine Learning Models for Tea Classification

The comparative analysis of machine learning models for tea classification, evaluated using precision, recall, F1-score, and accuracy after PCA-based feature extraction, highlights the superior performance of ensemble models (Table 4.3). The Random Forest model outperformed all others, achieving 100% accuracy, making it the most effective classifier. XGBoost also exhibited exceptional performance, achieving 100% test accuracy, indicating its reliability in handling complex patterns. In contrast, KNN (k=5) (57.58%) and SVM (69.70%) showed lower accuracy, suggesting that distance-based (KNN) & kernel-based (SVM) methods may not generalize as effectively for this dataset. KNN struggled with misclassifications, particularly for certain classes, leading to a lower recall & F1-score. SVM demonstrated moderate performance, but its accuracy was significantly lower than ensemble models. These findings indicate that RF & XGBoost are the most suitable models for tea classification, offering higher accuracy, robustness, & better generalization compared to KNN & SVM.

Table 4.3: Performance Metrics of Machine Learning Models for Tea Classification

Model	Precision	Recall	F1-Score	Accuracy
XGBoost	0.96	0.94	0.93	93.94%
Random Forest	1.00	1.00	1.00	100.00%
KNN (k=5)	0.94	0.91	0.91	90.91%
SVM (RBF Kernel)	0.89	0.85	0.85	84.85%

The figure 4.22 presents confusion matrices assessing the performance of XGBoost, Random Forest (RF), KNN (k=5), and SVM (RBF Kernel) for tea classification. The RF model achieved perfect classification with 100% accuracy, establishing it as the most reliable model. XGBoost also performed exceptionally well, with only a few minor misclassifications. In contrast, KNN (k=5) and SVM exhibited higher misclassification rates, with SVM being the least effective due to multiple incorrect predictions. Overall, the results confirm that ensemble models (XGBoost and RF) outperform KNN and SVM, making them the most suitable models for tea classification.

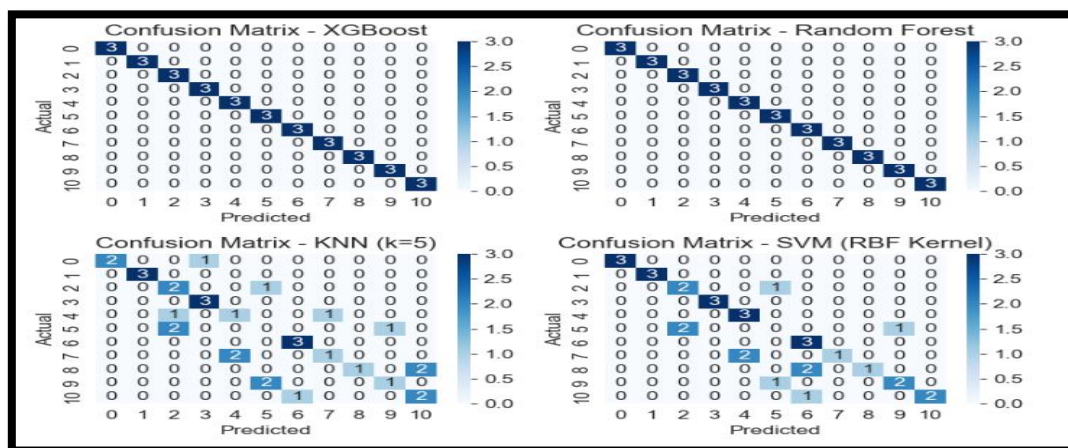


Figure 4.23: Confusion Matrices of Machine Learning Models for Tea Classification

The comparative analysis of XGBoost, RF, KNN (k=5), and SVM (RBF Kernel) for food adulteration detection highlights the superiority of ensemble models. The Random Forest model consistently achieved perfect classification (100% accuracy), making it the most robust and reliable model across different datasets, including turmeric and tea. XGBoost also demonstrated high accuracy with minimal misclassifications, confirming its strong generalization capabilities. In contrast, KNN

and SVM performed relatively poorly, with higher misclassification rates, indicating that distance-based and kernel-based approaches may not generalize well for food adulteration datasets. The confusion matrices further confirm that ensemble models (RF and XGBoost) offer better precision, recall, and overall classification accuracy, making them the best-suited approaches for detecting food adulteration.

4. 7 Food Adulteration Detection System

The Graphical User Interface (GUI) for the Food Adulteration Detection System is crafted to provide an intuitive and user-centric experience for the analysis of food samples. The interface enables users to enter client and sample information, including the client name, client ID, and sample name. It accommodates three categories of food items: Tea, Turmeric, and Chilly, each with a specific button for choosing the respective ASD file. Upon file selection, the system analyses the data using machine learning models and delivers the adulteration outcome. the system indicates if the chosen food item is contaminated, accompanied by a categorization result. The interface prominently displays an intuitive arrangement with distinctly named sections, color-coded buttons for various food categories, and an escape button for user convenience. This graphical user interface successfully incorporates machine learning models (Random Forest, XGBoost, KNN, and SVM) into a functional application, enabling fast detection of food adulteration.

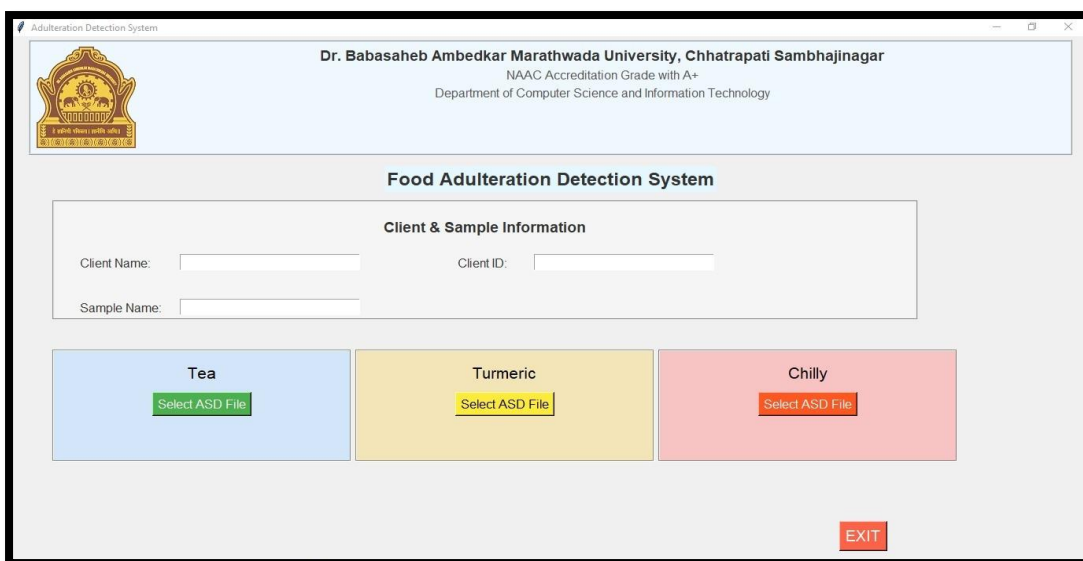


Figure 4.24: Graphical User Interface of the Food Adulteration Detection System

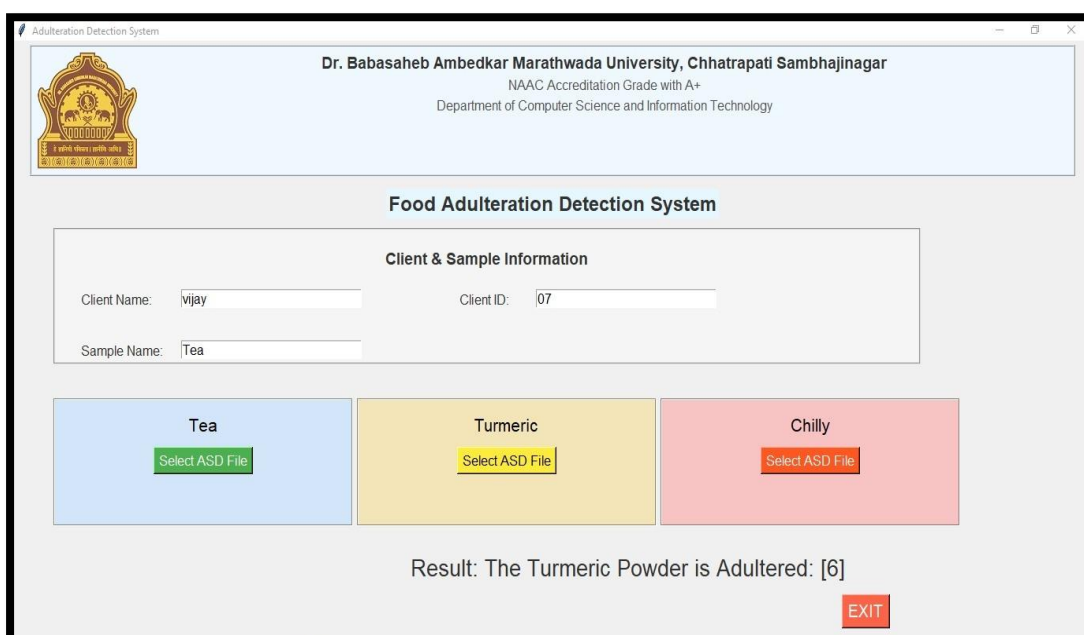


Figure 4.25: GUI with Sample Analysis Result

4.8 Creation of Spectral Library

Spectral libraries are an essential method in spectral analysis, offering a reference dataset for the categorization and thorough investigation of diverse dietary samples. This study examines the methodology for developing spectral libraries for chili powder, turmeric powder, and tea powder with ENVI software.

4.8.1. Process of Spectral Library Creation

1. Software Readiness
 - a. Launch ENVI Software
 - b. Access: Toolbox → Spectral → Spectral Libraries → Spectral Library

2. Input File Selection

There are three main input file types available:

- a. Data file
- b. ASCII file
- c. Spectrum file

In this case, the ASCII file format was selected for spectral data input.

3. Procedures for Data Processing

- a. Import the chosen input file using the import dialogue box.
- b. Compute and graph the average spectral signature of the sample.

- c. Export the spectral database.
- d. Save the result file with a period.SLI extension

We created a spectrum library to improve the precision of food adulteration detection in our study. This spectrum library serves as an extensive reference, documenting the distinct spectral characteristics of food samples.

The spectral libraries for chili (Figure 4.23), turmeric powder (Figure 4.24), and tea powder (Figure 4.25) exhibit unique reflectance patterns, highlighting their varied chemical compositions, colors, and structural features. Turmeric powder generally shows higher reflectance within the near-infrared spectrum (1000–2500 nm), primarily attributed to the existence of curcuminoids, starch, and cellulose. In contrast, chilli powder exhibits relatively lower reflectance in certain regions, showing unique absorption traits due to its rich pigment makeup, which includes carotenoids like capsanthin and capsorubin, as well as capsaicinoids that add to its heat. Additionally, turmeric exhibits a more consistent and stable spectral profile, while chili powder demonstrates greater spectral variability, likely due to differences in moisture levels, pigment concentration, and processing methods. Turmeric shows strong reflectance from 1000 to 1400 nm and significant absorption from 1500 to 2000 nm, whereas chili exhibits distinct spectral traits in the 500-1000 nm range due to its red pigments. Similarly, the spectral library of tea powder (Figure 4.25) provides a comprehensive representation of its reflectance characteristics over various wavelengths. The reflectance spectrum of tea powder exhibits distinct absorption and reflection properties, primarily influenced by its biological components, such as polyphenols, catechins, tannins, and pigments like chlorophyll and theaflavins. The spectral profile shows a marked increase in reflectance in the near-infrared range (1000–1400 nm), due to the existence of organic compounds and structural carbohydrates. Noticeable absorption bands are observed in the range of 1500 to 2000 nm, likely due to moisture, protein, and lignin content. The differences in the spectral curves reveal variations in tea processing, origin, and composition.

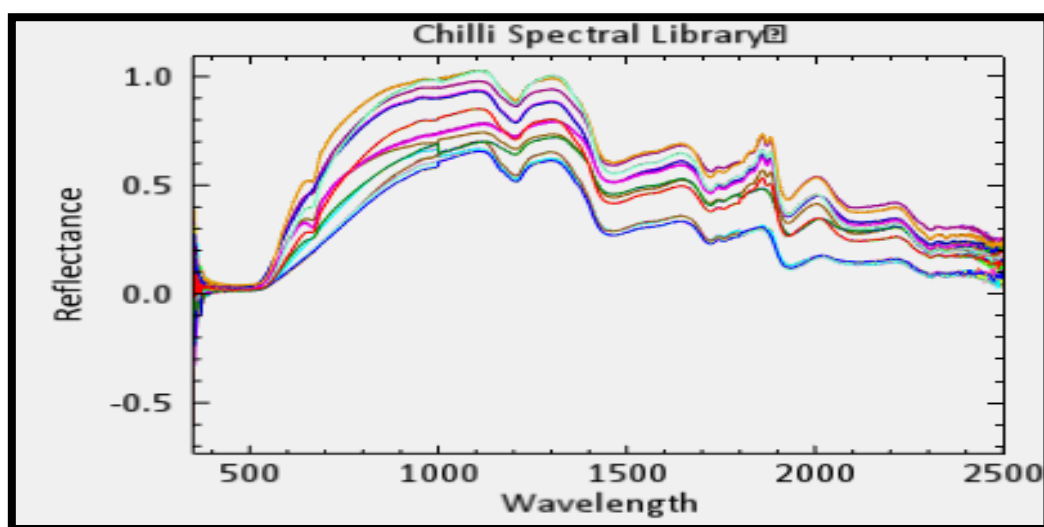


Figure 4.26: Spectral Library of Chili Powder

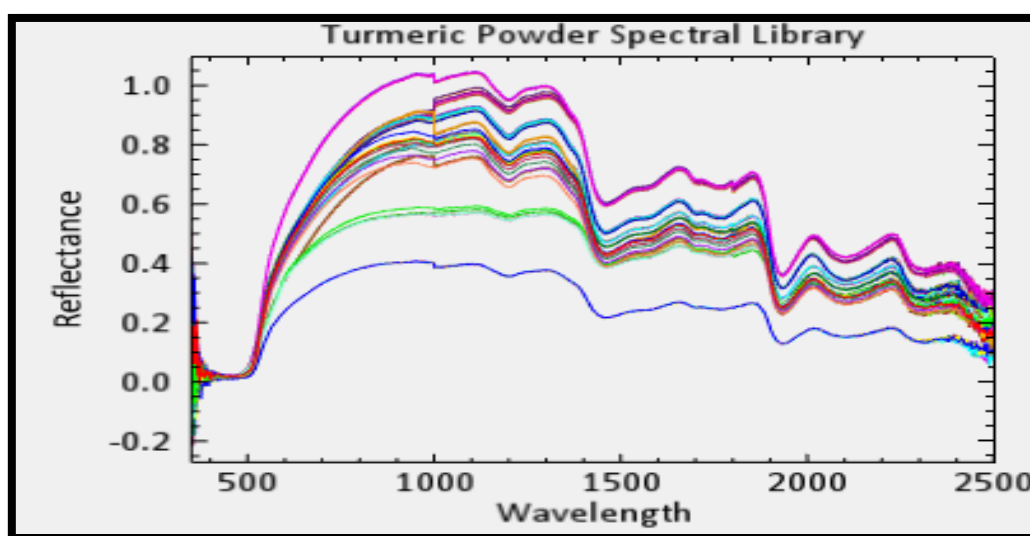


Figure 4.27: Spectral Library of Turmeric Powder

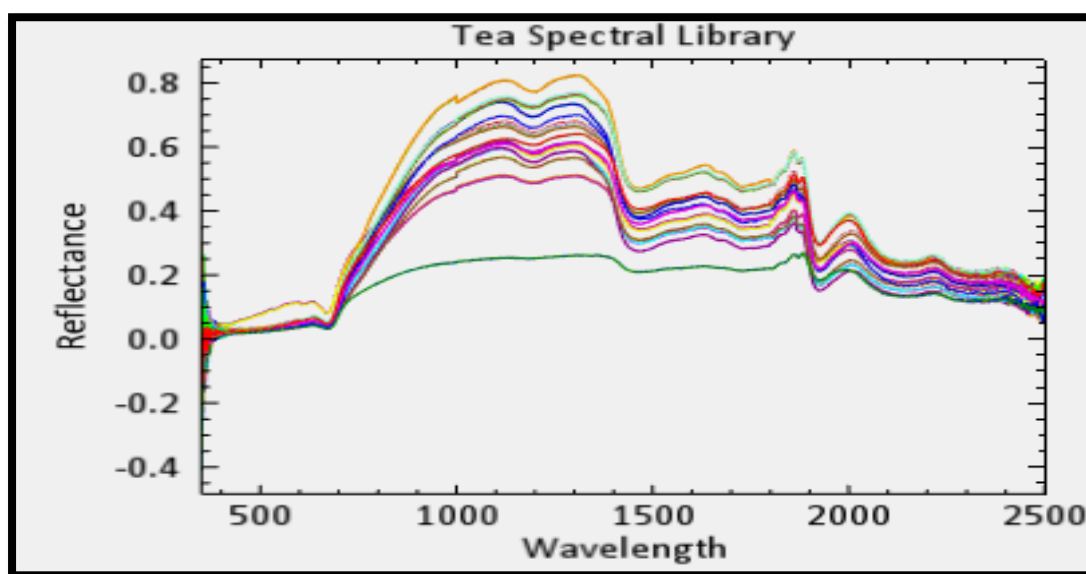


Figure 4.28: Spectral Library of Tea Powder

Turmeric, chili, and tea possess unique spectral signatures that reflect their varied chemical compositions. Turmeric shows enhanced reflectance in the infrared range due to curcuminoids, starch, and cellulose, while chilli exhibits greater absorption traits in the visible spectrum, resulting from its rich pigment components like carotenoids and capsaicinoids. Similarly, tea displays unique patterns of absorption and reflection determined by its content of pigments and polyphenols. Differences highlight the effectiveness of spectrum analysis as a powerful tool for distinguishing different food products and detecting potential adulterants, thereby ensuring quality and authenticity in food assessments.

4.9 Summary

chapter 4 presented the findings of our study on detecting adulteration in tea, turmeric, and chili powder using both traditional methods and advanced hyperspectral non-imaging technology combined with machine learning. The discussion began with simple household-level techniques such as visual inspection, water tests, and color spread checks, which proved useful for preliminary detection of common adulterants like brick powder, chalk, starch, and artificial dyes. Although these traditional approaches are cost-effective and easy to perform, they are limited in terms of accuracy and standardization. This highlighted the necessity of moving towards more robust and data-driven methods.

To address this gap, we collected hyperspectral reflectance data using the ASD FieldSpec 4 spectroradiometer and then applied a comprehensive preprocessing pipeline. The preprocessing steps included noise removal, Savitzky-Golay smoothing, second derivative transformation, normalization, and additional filtering to improve the signal-to-noise ratio. These steps significantly enhanced the spectral clarity, minimized baseline drift, and amplified subtle absorption features, making the data more reliable for classification. The before-and-after spectral comparisons clearly showed that preprocessing brought out distinct differences between pure and adulterated samples across chili, turmeric, and tea powders.

Feature extraction played a crucial role in simplifying the high-dimensional spectral data. We used statistical measures (mean, standard deviation, skewness, kurtosis, and quartiles), signal-based descriptors (energy, entropy, slopes, and area under the curve), and peak–valley metrics to capture meaningful spectral information. Principal Component Analysis (PCA) was then applied to reduce dimensionality while retaining the majority of variance. The PCA plots confirmed clear clustering for chili and turmeric powders, while tea showed more overlap, suggesting the need for more refined methods in some cases. The extracted features were further classified using different machine learning algorithms, namely Random Forest (RF), XGBoost, k-Nearest Neighbors (kNN), and Support Vector Machines (SVM). Among these, ensemble models Random Forest and XGBoost consistently outperformed the others, achieving near perfect-to-perfect classification accuracies across all three food powders. SVM and kNN showed lower performance, particularly struggling with misclassifications. These results strongly demonstrate that ensemble models, when combined with PCA-based feature extraction, are the most reliable for adulteration detection.

In addition to model training, we also developed a user-friendly Graphical User Interface (GUI) to make this research applicable in real-world scenarios. The GUI allows users to upload spectral files, run analyses, and instantly determine whether a food sample is pure or adulterated. This system integrates the trained machine learning models into a practical application, bridging the gap between laboratory research and consumer-level usage.

Finally, spectral libraries for chili, turmeric, and tea powders were created using ENVI software. These libraries documented the unique reflectance signatures of each powder, reflecting their chemical compositions and absorption patterns. Turmeric displayed strong reflectance in the near-infrared range, chili exhibited unique features in the visible spectrum due to pigments, and tea showed distinctive patterns influenced by polyphenols and tannins. These spectral libraries provide a valuable reference for future research and food quality assurance.

In summary, this chapter demonstrated that while traditional methods remain relevant for quick, initial checks, hyperspectral non-imaging technology combined with machine learning provides a far more accurate and standardized approach for adulteration detection. The study successfully validated Random Forest and XGBoost as the most effective classifiers and established spectral libraries that can serve as benchmarks for future work in food authenticity.

4.10 References

- [1].Cen, H., & He, Y. (2007). Theory and application of near-infrared reflectance spectroscopy in the determination of food quality. *Trends in Food Science & Technology*, 18(2), 72-83.
- [2].Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends in food science & technology*, 18(12), 590-598.
- [3].Dhangar, V. (2024). Exploring the Spectral Database of Food Samples Using ASD Field Spec 4 Spectroradiometer. Available at SSRN 4900729.
- [4].Vignesh, T., Selvakumar, D., & Jayavel, R. (2025). Detecting ferric oxide adulteration in chilli Powder: A Multimodal analytical approach for enhanced food safety. *Microchemical Journal*, 208, 112332.
- [5].Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends in food science & technology*, 18(12), 590-598.
- [6].Cen, H., & He, Y. (2007). Theory and application of near-infrared reflectance spectroscopy in the determination of food quality. *Trends in Food Science & Technology*, 18(2), 72-83.
- [7].Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends in food science & technology*, 18(12), 590-598.
- [8].Savitzky, A., & Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8), 1627-1639.
- [9].Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3), 37-52.
- [10].Al-Awadhi, M.A. and Deshmukh, R.R., 2021, October. Honey Classification using Hyperspectral Imaging and Machine Learning. In 2021 Smart Technologies, Communication and Robotics (STCR) (pp. 1-5). IEEE

- [11]. Jolliffe, I. T. (2002). Principal component analysis for special types of data (pp. 338-372). Springer New York. Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. Chemometrics and intelligent laboratory systems, 2(1-3), 37-52.

CONCLUSION AND FUTURE WORK

5.1. Conclusion

In this research, we have demonstrated the effectiveness of hyperspectral non-imaging spectroscopy combined with machine learning techniques for detecting adulteration in powdered food items such as spices and tea. Traditional methods of adulteration detection, although reliable, are often time-consuming, labor-intensive, and require chemical handling, making them less suitable for large-scale monitoring. In contrast, hyperspectral spectroscopy offers a non-destructive, rapid, and accurate alternative by capturing the unique spectral signatures of both pure and adulterated samples.

To improve classification performance, we applied advanced preprocessing techniques like Savitzky-Golay filtering and Principal Component Analysis (PCA). Various machine learning algorithms including XGBoost, Random Forest, SVM, and Decision Tree were evaluated, with Random Forest and XGBoost yielding the highest accuracy, even achieving 100% accuracy in certain cases. These results validate the potential of hyperspectral data in reliably distinguishing between pure and adulterated food powders. Additionally, we developed a user-friendly Graphical User Interface (GUI) to facilitate real-time analysis and simplify data interpretation. This system holds great promise for application in food quality control labs and by regulatory agencies, offering a faster, safer, and more efficient approach to adulteration detection.

Overall, the study establishes hyperspectral non-imaging spectroscopy as a viable and effective approach to food adulteration testing, highlighting its role in advancing food safety monitoring and ensuring transparency across the food supply chain.

5.2. Future work

Future research will extend the application of hyperspectral non-imaging techniques to a wider variety of food products, including milk, honey, flour, edible oils, and legumes. To enable more precise adulteration detection, the integration of hyperspectral imaging (HSI) will be explored, allowing spatial-spectral analysis to identify localized contaminants. Another important direction will involve the development of portable, real-time detection systems based on embedded hardware

platforms such as Raspberry Pi and FPGA, thereby facilitating on-site monitoring. In addition, advanced deep learning architectures, including Convolutional Neural Networks (CNNs) and Transformer-based models, will be implemented to further enhance classification accuracy and robustness. The fusion of hyperspectral data with complementary sensing modalities, such as Near-Infrared (NIR) and Raman spectroscopy, will also be pursued to improve detection reliability through multimodal analysis. Efforts will be directed toward establishing a comprehensive spectral library encompassing a broad range of adulterant profiles, as well as designing an automated, user-friendly graphical interface with cloud-based connectivity.

Finally, collaborations with regulatory authorities and industry stakeholders will be prioritized to ensure rigorous validation, facilitate standardization, and enable effective translation of the system into practical real-world applications.