# PERFORMANCE ANALYSIS OF HSI-RGB FEATURE FUSION FOR FIRE BLIGHT DETECTION IN APPLE LEAVES

**A Thesis Submitted**
**In Partial Fulfillment of the Requirements for the Degree of**

## MASTER OF TECHNOLOGY

**in**

**Signal Processing & Digital Design**

**by**


**Vaibhav Sharma**
**(Roll No. 23/SPD/04)**

**Under the Supervision of**

**Prof. S. Indu**

**Department of Electronics and Communication Engineering**

**DELHI TECHNOLOGICAL UNIVERSITY**
**(Formerly Delhi College of Engineering)**
**Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India**


**May, 2025**

## ACKNOWLEDGEMENTS

I wish to express my most sincere gratitude to the individuals who have been pivotal to the completion of this thesis.

My deepest appreciation goes to my esteemed guide, Prof. S. Indu. Her insightful guidance, unwavering support, and constant encouragement were instrumental throughout this research journey. I am profoundly grateful for her mentorship and expertise, which have been invaluable in shaping this work.

I also extend my heartfelt thanks to my parents. Your unending love, patience, and belief in me have been my greatest source of strength and motivation. This achievement would not have been possible without your unwavering support.

**Candidate's Signature**

**Vaibhav Sharma**

**DELHI TECHNOLOGICAL UNIVERSITY**
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

## CANDIDATE'S DECLARATION

I, Vaibhav Sharma hereby certify that the work which is being presented in the thesis entitled " Performance Analysis of HSI-RGB Feature Fusion for Fire Blight Detection in Apple Leaves " in partial fulfillment of the requirements for the award of the Degree of Master of Technology, submitted in the Department of Electronics & Communication Engineering, Delhi Technological University is an authentic record of my own work carried out during the period from August 2023 to May 2024 under the supervision of Prof. S. Indu.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

**Candidate's Signature**

**DELHI TECHNOLOGICAL UNIVERSITY**
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

## CERTIFICATE BY THE SUPERVISOR

Certified that **Vaibhav Sharma** (23/SPD/04) has carried out their search work presented in this thesis entitled **"Performance Analysis of HSI-RGB Feature Fusion for Fire Blight Detection in Apple Leaves"** for the award of **Master of Technology** from Department of Electronics & Communication Engineering, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Signature

S. Indu

Professor
DTU, Shahbad Daulatpur,
Main Bawana Road, Delhi-42

Date:

**Performance Analysis of HSI-RGB Feature Fusion for Fire Blight Detection in Apple Leaves**

**Vaibhav Sharma**

## ABSTRACT

This study evaluates machine learning models—Support Vector Machine, Random Forest, and XGBoost—for fire blight detection in apple leaves using hyperspectral (HSI) data and fused HSI-RGB features. Results show that while HSI data alone enables strong classification (F1-score up to 0.93), fusing HSI with RGB features significantly enhances performance. The Random Forest model with fused features achieved the highest accuracy and F1-score (0.98). Visual assessments further confirm improved localization of infected regions with feature fusion. These findings demonstrate that multimodal data integration and ensemble learning substantially advance early, accurate fire blight detection for precision agriculture.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# SYMBOLS, ABBREVIATIONS & NOMENCLATURE

HSI : Hyper Spectral Imaging

RGB : Red Green Blue

SVM : Support Vector Machine

RF : Random Forest

XGBoost : eXtreme Gradient Boosting

# CHAPTER 1

# INTRODUCTION

## 1.1 Background & Motivation

Apple (Malus domestica) is one of the most economically vital fruit crops globally, with annual production exceeding 95 million metric tons and contributing over $23 billion to the global economy (FAOSTAT, 2023). China dominates production, yielding 47.5 million tons annually (49.6% of global output), followed by the United States (4.4 million tons) and Poland (4.2 million tons) (Wikipedia, 2024; Statista, 2024). In regions like Himachal Pradesh, India, apple orchards support livelihoods for 30 million people, underscoring the crop's socio-economic importance (IJCRT, 2022). However, this critical industry faces existential threats from fire blight, a bacterial disease caused by Erwinia amylovora, which causes annual U.S. losses exceeding $100 million and orchard mortality rates of 50–60% in young trees (Virginia Tech, 2024; Cornell CALS, 2024).

Fire blight spreads rapidly under warm (18–30°C) and humid conditions, infiltrating plants through blossoms, wounds, or natural openings. Early symptoms include water-soaked blossoms and necrotic shoots exhibiting the characteristic "shepherd's crook" deformity, while advanced infections lead to sunken cankers and bacterial ooze that facilitate secondary spread via insects, rain, or contaminated tools (Britannica, 1998; Microbe Notes, 2024). Traditional detection methods—PCR assays and ELISA—detect E. amylovora at concentrations $\geq 10^3$ CFU/mL but fail during the 7–21-day asymptomatic incubation period, when bacterial populations remain subthreshold (Frontiers in Horticulture, 2023; Singh et al., 2020). Visual inspections, though widely used, are labor-intensive and impractical for large-scale monitoring, underscoring the need for non-invasive, real-time diagnostic tools (SSRN, 2023).

Hyperspectral imaging (HSI) and RGB cameras have emerged as transformative tools for early disease detection. HSI captures biochemical changes across 400–2500 nm, resolving pre-symptomatic physiological stress indicators such as chlorophyll degradation (680–750 nm) and water stress (1440–1910 nm) with 85–95% accuracy (MDPI Remote Sensing, 2020; Frontiers in Plant Science, 2019). Conversely, RGB imaging provides high spatial resolution (0.1–0.5 mm/pixel), capturing lesion morphology and texture features but lacking spectral depth for pre-symptomatic biochemical shifts (Frontiers in Plant Science, 2022). The fusion of HSI and RGB data addresses these limitations by combining complementary features: HSI's biochemical sensitivity and RGB's structural clarity. Recent studies, such as Zhang et al. (2023), achieved 98.36% accuracy in soybean defect detection using fused HSI-RGB features and lightweight CNNs, validating the potential of multimodal fusion for precision agriculture.

## 1.2. Problem Statement

Despite advancements, three systemic limitations hinder effective fire blight control:

1. **Spatial-Spectral Trade-offs:**

HSI's Low Spatial Resolution (1–5 mm/pixel): Fails to resolve fine-grained leaf structures (e.g., early lesions <1 mm) (MDPI Sensors, 2023). RGB's Spectral Simplicity (400–700 nm): Misses pre-symptomatic biochemical shifts (e.g., water stress at 1,450 nm) (Frontiers in Plant Science, 2022).

2. **Model Bias Toward Deep Learning:**

Convolutional neural networks (CNNs) dominate hyperspectral research but require >10 million parameters and large datasets (>10,000 samples) (Nature, 2023). Traditional ML models like SVM and XGBoost achieve comparable accuracy (85–95%) with <1% of the parameters but remain underexplored in fused HSI-RGB contexts (IJETT, 2022; PMC, 2022).

3. **Generalizability in Field Conditions:**

Over 80% of fusion techniques are validated in controlled environments, neglecting real-world variability like heterogeneous illumination and canopy complexity. A 2023 study reported a 32% accuracy drop when transitioning HSI-RGB models from lab to orchard settings due to ambient light fluctuations (Frontiers in Plant Science, 2023c; Arxiv, 2024).

4. **Threshold Optimization Barriers:**

Static thresholds fail to balance precision (minimizing false positives) and recall (minimizing false negatives) in dynamic field conditions. For example, UAV-based detection may prioritize recall (71.4%) over precision (73.7%), risking unnecessary pesticide use (Xiao et al., 2022).

## 1.3 Objectives of the Study

This study addresses these challenges through three primary objectives:

1. **Develop a Feature Fusion Framework:**

Integrate HSI's spectral richness (400–2500 nm) with RGB's spatial precision (0.1–0.5 mm/pixel) using early fusion strategies to preserve biochemical-spatial correlations critical for early detection (Zhang et al., 2023; MDPI Remote Sensing, 2020).

2. **Compare ML Model Performance:**

Evaluate SVM (radial basis kernel), RF (100 trees with Gini impurity splitting), and XGBoost (learning rate 0.1, max depth 5) on fused datasets using metrics like accuracy, precision, recall, F1-score, and ROC-AUC (IJETT, 2022; Sari et al., 2023).

3. **Identify Optimal Spectral-Spatial Features:**

Use SHAP analysis and recursive feature elimination (RFE) to pinpoint critical bands (e.g., 700 nm for chlorophyll loss, 1440 nm for water stress) and spatial descriptors (e.g., LAB color histograms, GLCM contrast) (MDPI Sensors, 2023; Wang et al., 2023).

## 1.4. Scope of Work

### 1.4.1. Inclusions

This research focuses on the development and comparative evaluation of machine learning models for the detection of fire blight disease in apple leaves using both hyperspectral and multimodal (HSI + RGB) imaging data. The scope of the study encompasses the following key areas:

- **Hyperspectral Data Utilization**: Exploring the capacity of visible–near infrared hyperspectral imagery to discriminate between healthy and infected leaf tissues at the pixel level using spectral and spatial features.

- **Multimodal Feature Fusion**: Investigating the integration of hyperspectral data with traditional RGB information to assess whether the fusion of modalities improves classification accuracy and robustness.
- **Model Diversity and Comparison**: Implementing and comparing three established machine learning classifiers—Random Forest, XGBoost, and Linear SVM—across two data representations (hyperspectral-only and multimodal).
- **Balanced Learning Framework**: Designing a balanced sampling framework to address class imbalance between symptomatic and healthy samples, ensuring fair model evaluation and reproducibility.
- **Comprehensive Evaluation**: Employing standard classification metrics (accuracy, precision, recall, F1 score, and confusion matrix) to systematically evaluate model performance and to support the selection of optimal strategies for disease detection.
- **Scalability and Practical Deployment**: Providing a reproducible and computationally feasible methodology that can be scaled and adapted to other plant disease detection tasks using hyperspectral and RGB data fusion.

### 1.4.2. Exclusions

- Deep Learning Models: CNNs (e.g., ResNet50, EfficientNet) and transformers are omitted due to computational constraints and limited interpretability for small datasets (<10,000 samples) (Nature, 2023).
- Post-Harvest Management: Focus remains on pre-symptomatic detection, excluding chemical treatments, orchard recovery strategies, or post-harvest storage analysis (Virginia Tech, 2024).

## 1.5. Thesis Organization

**Chapter 2: Literature Review**

Critically analyzes HSI/RGB imaging, fire blight pathology, and ML applications, highlighting gaps in fusion-classifier synergy (PMC, 2022; Frontiers in Plant Science, 2023b).

**Chapter 3: Methodology**

Details data acquisition (HSI-RGB alignment via affine transformation), preprocessing (noise removal, normalization), and fusion protocols (MDPI Remote Sensing, 2020; Zhang et al., 2023).

**Chapter 4: Results**

Presents comparative performance metrics (accuracy, F1-score, ROC-AUC) for SVM, RF, and XGBoost on single-modality and fused datasets (IJETT, 2022; Sari et al., 2023).

**Chapter 5: Conclusion**

Synthesizes contributions to early detection and proposes future work on UAV-based hyperspectral phenotyping (Frontiers in Plant Science, 2024; HIPPA Project, 2024).

By harmonizing multimodal imaging and machine learning, this research advances scalable tools for mitigating fire blight's catastrophic impact on global apple production.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Fire Blight Disease: Biology, Symptoms, and Impact on Agriculture

Fire blight is a highly contagious and destructive bacterial disease affecting apples, pears, and other members of the Rosaceae family (Wikipedia, 2024). The causal agent, Erwinia amylovora, is a Gram-negative bacterium that invades host plants through blossoms, wounds, or natural openings, leading to rapid necrosis of shoots, leaves, and fruits (Microbe Notes, 2024). The disease is notorious for its ability to destroy entire orchards within a single growing season under optimal conditions, with economic losses in the United States alone exceeding $100 million annually (Frontiers in Horticulture, 2023; Cornell CALS, 2024).

The biology of E. amylovora is characterized by the production of a viscous exopolysaccharide, levan, which aids in the formation of protective biofilms, enhancing bacterial survival and adhesion to host tissues (Microbe Notes, 2024). The pathogen can overwinter in cankers on branches and trunks, becoming active in spring when temperatures rise and humidity increases (Cornell CALS, 2024). Insects and rain play a significant role in disseminating the pathogen, especially during bloom periods (Frontiers in Horticulture, 2023). The disease cycle includes blossom blight, shoot blight, canker blight, and rootstock blight, each presenting distinct symptoms such as water-soaked lesions, wilted shoots with a "shepherd's crook," and sunken cankers that may girdle and kill the tree (Cornell CALS, 2024; Wikipedia, 2024).

Recent studies have highlighted the devastating impact of rootstock infections, which can lead to rapid tree death and significant crop losses (Frontiers in Horticulture, 2023). For instance, Aćimović et al. (2023) reported that blossom and shoot blight incidence can range from 30% to 100% depending on orchard conditions, with up to 65% of trees in some orchards succumbing to the disease. The latent phase of infection, during which trees may harbor asymptomatic bacteria, complicates early detection and management (Frontiers in Horticulture, 2023).

### 2.2 Plant Disease Detection Methods: Traditional and Modern Approaches

Traditional detection methods for fire blight and other plant diseases include visual inspection, culture-based assays, and molecular techniques such as PCR and ELISA (Frontiers in Plant Science, 2023a). Visual scouting is widely practiced due to its low cost, but it is subjective and often fails to detect early or latent infections (SSRN, 2023). Molecular assays, while highly specific, require laboratory infrastructure and are not suitable for rapid, large-scale field deployment (Frontiers in Plant Science, 2023a). For example, PCR can detect E. amylovora at concentrations as low as $10^3$ CFU/mL, but its effectiveness is limited by the presence of inhibitors in plant tissues and the need for skilled personnel (Frontiers in Horticulture, 2023).

Recent advances have introduced point-of-care devices and biosensors for on-site diagnosis, allowing for faster and more user-friendly detection (Frontiers in Plant Science, 2023a). The development of digital droplet PCR (ddPCR) and multiplex assays has further improved sensitivity and specificity for low-titer pathogens (Frontiers in Plant Science, 2023a). However, these methods still face challenges in scalability and cost-effectiveness for routine orchard monitoring.

Modern approaches increasingly leverage imaging technologies and artificial intelligence (AI) for automated disease detection (SSRN, 2023; Nature, 2023). Deep learning (DL) and machine learning (ML) models have demonstrated high accuracy in classifying plant diseases from digital images, with CNN-based methods achieving up to 98% accuracy in some cases (Nature, 2023; IJIRT, 2025). These systems can process large volumes of data, enabling real-time, high-throughput screening of orchards, but their performance can be affected by variations in lighting, background, and disease presentation (Arxiv, 2024).

## 2.3 Hyperspectral Imaging (HSI): Principles and Applications in Agriculture

Hyperspectral imaging (HSI) is a powerful tool for non-invasive plant disease detection, capturing reflectance data across hundreds of narrow spectral bands (Imec, 2024). HSI enables the detection of subtle biochemical and physiological changes in plant tissues before visible symptoms appear, making it particularly valuable for early disease diagnosis (Frontiers in Plant Science, 2024; Pixxel, 2024). The technology operates across the visible, near-infrared, and shortwave infrared regions (400–2500 nm), allowing for the identification of changes in chlorophyll absorption, water content, and other stress indicators (Imec, 2024; Pixxel, 2024).

Recent advances have demonstrated the effectiveness of HSI in detecting a range of plant diseases, including fire blight, citrus canker, and seedborne pathogens (Frontiers in Plant Science, 2024; Frontiers in Plant Science, 2023b). For example, Abdulridha et al. (2020) showed that HSI could detect fire blight in apple leaves with high accuracy by analyzing spectral signatures associated with chlorophyll degradation and water stress. Studies have also highlighted the non-destructive nature of HSI, which allows for repeated measurements and integration with other diagnostic methods (Frontiers in Plant Science, 2024).

HSI data are highly collinear and require advanced statistical and computational tools for information extraction and pattern modeling (Frontiers in Plant Science, 2024). Machine learning and deep learning algorithms, such as SVM, PLS-DA, and CNNs, have been widely applied to classify healthy and diseased samples based on hyperspectral data (PMC, 2022; Frontiers in Plant Science, 2024). For instance, Chu et al. (2020) achieved 100% accuracy in classifying corn seeds infected with Aspergillus spp. using PCA and SVM models, while Wu et al. (2022) reported over 97% accuracy in identifying peanut seeds infected with Aspergillus flavus using hyperspectral images and various classifiers.

Despite its advantages, the adoption of HSI in agriculture is limited by high equipment costs, complex data processing, and the need for interdisciplinary collaboration (Frontiers in Plant Science, 2024; Imec, 2024). Recent developments in hyperspectral snapshot cameras and UAV-mounted sensors are helping to overcome these barriers by enabling real-time, high-resolution data acquisition and expanding the practical applications of HSI in precision agriculture (Imec, 2024; Pixxel, 2024).

### 2.4 RGB Imaging: Role in Plant Disease Detection

RGB imaging is widely used in plant disease detection due to its accessibility, low cost, and ability to capture high-resolution spatial features (Frontiers in Plant Science, 2022). RGB images, typically acquired with standard digital cameras, are effective for identifying visible symptoms such as leaf spots, discoloration, and lesions (Frontiers in Plant Science, 2022; SSRN, 2023). Machine learning models, including CNNs and SVMs, have been applied to classify plant diseases from RGB images, achieving high accuracy in many cases (IJIRT, 2025; Frontiers in Plant Science, 2023c).

However, RGB imaging is limited in its ability to detect pre-symptomatic or subtle physiological changes, as it only captures information in the visible spectrum (400–700 nm) (Frontiers in Plant Science, 2022). Recent research has explored the integration of RGB with other modalities, such as hyperspectral or thermal imaging, to enhance detection performance (Metallurgical and Materials Engineering, 2025). For example, Zhang et al. (2021) proposed a CNN model optimized for maize disease detection, demonstrating that reconstructed hyperspectral data from RGB images could improve detection accuracy, especially in complex environments (Frontiers in Plant Science, 2022).

### 2.5 Feature Fusion Techniques: Multimodal Data Fusion in Plant Pathology

Multimodal data fusion combines features from multiple imaging modalities, such as HSI, RGB, and thermal imaging, to improve the accuracy and robustness of plant disease detection (Metallurgical and Materials Engineering, 2025). Early fusion strategies concatenate features before classification, while late fusion combines the outputs of separate classifiers (Metallurgical and Materials Engineering, 2025; Frontiers in Plant Science, 2023b). Hybrid fusion approaches leverage both strategies to maximize information gain and model performance.

Recent studies have demonstrated the effectiveness of feature fusion in plant disease diagnosis. For instance, Assudani and Krishna (2025) introduced a multimodal deep learning framework that integrates RGB, hyperspectral, and thermal images using CNNs and Vision Transformers, achieving superior accuracy and robustness compared to single-modality models (Metallurgical and Materials Engineering, 2025).

Similarly, Wang et al. (2023) showed that multimodal fusion networks could achieve high accuracy for early disease detection in citrus by leveraging complementary information from different sensors (Frontiers in Plant Science, 2023b).

Despite these advances, challenges remain in optimizing fusion strategies for real-time, field-deployable systems. Computational complexity, data heterogeneity, and the need for large labeled datasets can limit the scalability and generalizability of multimodal approaches (Metallurgical and Materials Engineering, 2025; Arxiv, 2024).

### 2.6 Machine Learning in Plant Disease Detection: SVM, RF, XGBoost, and Related Works

Machine learning and deep learning have revolutionized plant disease detection by enabling automated, high-throughput analysis of complex image data (SSRN, 2023; Frontiers in Plant Science, 2023c). Support Vector Machines (SVM), Random Forests (RF), and XGBoost are among the most widely used ML algorithms for plant disease classification (PMC, 2022; IJIRT, 2025). SVM is known for its generalization ability and effectiveness with high-dimensional data, while RF and XGBoost offer robust performance and interpretability (PMC, 2022; IJIRT, 2025).

Comparative studies have shown that deep learning models, particularly CNNs, often outperform traditional ML algorithms in terms of accuracy and robustness, especially when large labeled datasets are available (Nature, 2023; Frontiers in Plant Science, 2023c). For example, CNNs have achieved classification accuracies exceeding 96% on the PlantVillage dataset, while Random Forest and Gradient Boosting also demonstrate strong performance (IJIRT, 2025). However, SVM and other traditional ML models remain valuable for applications with limited data or computational resources (PMC, 2022).

Recent advances include the integration of ML and DL models with feature selection and dimensionality reduction techniques to improve classification performance and reduce computational costs (PMC, 2022; Frontiers in Plant Science, 2024). For instance, ReliefF and PCA have been used to extract relevant features from hyperspectral data, enabling more accurate and efficient disease detection (PMC, 2022).

## 2.7 Research Gaps

Despite significant progress, several challenges remain in developing robust, field-deployable plant disease detection systems. Many studies are conducted under controlled conditions, with limited validation in real-world environments where variability in lighting, background, and disease presentation can affect model performance (Frontiers in Plant Science, 2023c; Metallurgical and Materials Engineering, 2025). The integration of multimodal features and the comparative evaluation of traditional ML and DL models for early disease detection are still underexplored areas (Metallurgical and Materials Engineering, 2025; Arxiv, 2024).

This thesis addresses these gaps by developing and evaluating a multimodal feature fusion framework using HSI and RGB data, and systematically comparing the performance of SVM, RF, and XGBoost for fire blight detection in apple leaves. The approach emphasizes real-world applicability, scalability, and interpretability, aiming to provide actionable tools for precision agriculture and sustainable disease management.

# CHAPTER 3

# METHODOLOGY

This chapter elucidates the complete methodological framework employed for the classification of apple leaf health conditions using hyperspectral and multimodal (HSI + RGB) data. The proposed approach is developed using the publicly available "Visible – Near Infrared Hyperspectral Dataset of Healthy and Infected Apple Tree Leaves for the Monitoring of Apple Fire Blight." A structured process encompassing data acquisition, preprocessing, feature extraction, model training, and evaluation is presented herein. The overarching objective is to compare the performance of multiple machine learning classifiers on both hyperspectral-only and fused multimodal data representations.

## 3.1 Dataset Description

The dataset utilized in this study is specifically designed for plant disease detection under varying health conditions. It comprises leaf-level imagery acquired using visible and near-infrared hyperspectral imaging. Each plant sample folder contains the following components:

- A hyperspectral image file (.hdr and corresponding .dat)
- A visible-range RGB image (.png format)
- A CSV file specifying pixel coordinates annotated as symptomatic (infected)

The hyperspectral cube spans a broad spectral range, and each pixel captures spectral reflectance across hundreds of contiguous bands. The RGB image provides visual context and is used for complementary analysis. The annotations enable precise localization of diseased regions, facilitating supervised learning with labeled data.

## 3.2 Data Preprocessing

### 3.2.1 Balanced Sampling Strategy

To mitigate class imbalance, a balanced sampling strategy was adopted. For every plant image:

- All pixels annotated as symptomatic were identified and extracted based on the provided CSV file.
- An equal number of healthy pixels were randomly selected from the remaining non-annotated pixel grid.

This sampling process ensures that the dataset for each plant maintains a 1:1 ratio between diseased and healthy samples, thereby preventing classifier bias toward the majority class. A random number generator initialized with a fixed seed ensured reproducibility.

### 3.2.2 Dataset Partitioning

The total set of plant folders was split into training and testing subsets. Seventy percent (70%) of the folders were randomly allocated for training, and the remaining thirty percent (30%) were retained for testing. This approach preserves plant-level independence between training and testing samples, minimizing data leakage.

## 3.3 Feature Extraction

Feature engineering was performed separately for the hyperspectral-only and multimodal approaches.

### 3.3.1 Hyperspectral-Only Features

#### a) Spectral Band Reduction

Due to the high dimensionality of hyperspectral data, a band reduction process was applied. From the full spectral cube, 40 uniformly spaced bands were selected via downsampling to reduce computational complexity while preserving spectral information.

**b) Vegetation Indices**

Two well-established indices were computed:

- **Excess Green (ExG):** Derived from the RGB image to enhance the green component, calculated as: ExG=2G−R−BExG = 2G - R - B
- **Normalized Difference Vegetation Index (NDVI):** Calculated using the first (red) and last Near InfraRed (NIR) reflectance values:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{R}}{\mathrm{NIR} + \mathrm{R}} \tag{1}$$

**c) Texture Feature**

To capture local spatial variations, texture features were extracted from NDVI image patches. The Gray-Level Co-occurrence Matrix (GLCM) method was employed to compute the contrast statistic over 3x3 patches centered on each pixel. The final feature vector per pixel for the hyperspectral model was a concatenation of the 40 spectral bands, ExG, NDVI, and texture contrast.

### 3.3.2 Multimodal Fusion Features

For the multimodal approach, hyperspectral features were directly concatenated with RGB pixel values:

- For each pixel, the reduced hyperspectral vector (from 40 selected bands) was combined with its corresponding RGB values (R, G, B).
- This results in a multimodal feature vector of length 43 per pixel.

Symptomatic and healthy pixels were extracted and balanced using the same strategy as in the hyperspectral-only method. The features were stored in a unified CSV file for training.

## 3.4 Model Training

In this study, three distinct and widely used supervised learning algorithms were employed to build classification models: Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Linear Support Vector Machine (SVM). These models were selected for their complementary strengths in handling structured and high-dimensional data.

- **Random Forest (RF)** is an ensemble-based method that constructs a multitude of decision trees during training and outputs the class that is the mode of the classes (classification) of the individual trees. It is robust to overfitting and effective for datasets with a mix of numerical and categorical features.

- **Extreme Gradient Boosting (XGBoost)** is a high-performance implementation of gradient boosting that has gained popularity due to its speed and accuracy. It builds models in a sequential manner where each new tree corrects errors made by previously trained trees. Regularization techniques help in reducing overfitting.
- **Linear Support Vector Machine (SVM)** is a discriminative classifier formally defined by a separating hyperplane. In this work, a linear kernel is used for its scalability with high-dimensional data, such as hyperspectral input, and efficiency in terms of computation.

Each model was trained and tested on both hyperspectral-only and multimodal datasets to assess its capability in detecting symptomatic and healthy pixels across different input modalities.

Three supervised machine learning classifiers were evaluated for both the hyperspectral-only and multimodal features:

- **Random Forest (RF)**
- **Extreme Gradient Boosting (XGBoost)**
- **Linear Support Vector Machine (SVM)**

### 3.4.1 Preprocessing

All input features were standardized using the StandardScaler from the Scikit-learn library. For SVM models trained on hyperspectral-only data, Principal Component Analysis (PCA) was applied to reduce the feature dimensionality to 30 principal components, thereby improving model efficiency and reducing overfitting.

### 3.4.2 Random Forest Classifier

The Random Forest classifier was configured with the following parameters:

- n_estimators = 100
- class_weight = 'balanced'
- random_state = 42

For hyperspectral-only data, deeper trees and a greater number of estimators (up to 200) were tested. Models were trained on the standardized feature vectors and evaluated on the held-out test set.

### 3.4.3 XGBoost Classifier

The XGBoost classifier was configured with the following hyperparameters:

- n_estimators = 100
- max_depth = 6
- learning_rate = 0.1
- eval_metric = 'logloss'
- use_label_encoder = False
- random_state = 42

This gradient-boosting framework is optimized for tabular data and provides robust performance with built-in regularization. Models were trained and validated using stratified train-test splits.

### 3.4.4 Linear SVM Classifier

The Linear SVM classifier was selected for its simplicity and efficiency in high-dimensional spaces. The implementation used:

- class_weight = 'balanced'
- random_state = 42
- max_iter = 10000

For the hyperspectral-only SVM, PCA-reduced features were used. For the multimodal SVM, all features were retained

### 3.4.1 Preprocessing

All input features were standardized using the StandardScaler from the Scikit-learn library. For SVM models trained on hyperspectral-only data, Principal Component Analysis (PCA) was applied to reduce the feature dimensionality to 30 principal components, thereby improving model efficiency and reducing overfitting.

### 3.4.2 Random Forest Classifier

The Random Forest classifier was configured with the following parameters:

- n_estimators = 100
- class_weight = 'balanced'
- random_state = 42

For hyperspectral-only data, deeper trees and a greater number of estimators (up to 200) were tested. Models were trained on the standardized feature vectors and evaluated on the held-out test set.

### 3.4.3 XGBoost Classifier

The XGBoost classifier was configured with the following hyperparameters:

- n_estimators = 100
- max_depth = 6
- learning_rate = 0.1
- eval_metric = 'logloss'
- use_label_encoder = False
- random_state = 42

This gradient-boosting framework is optimized for tabular data and provides robust performance with built-in regularization. Models were trained and validated using stratified train-test splits.

### 3.4.4 Linear SVM Classifier

The Linear SVM classifier was selected for its simplicity and efficiency in high-dimensional spaces. The implementation used:

- class_weight = 'balanced'
- random_state = 42
- max_iter = 10000

For the hyperspectral-only SVM, PCA-reduced features were used. For the multimodal SVM, all features were retained.

## 3.5 Model Evaluation

The models were assessed on the test dataset using standard classification metrics:

### 3.5.1 Confusion Matrix

A confusion matrix is a tabular representation of actual vs. predicted classifications. It helps visualize performance across classes and is defined as follows:

| | **Predicted Positive** | **Predicted Negative** |
|---|---|---|
| Actual Positive | True Positive (TP) | False Negative (FN) |
| Actual Negative | False Positive (FP) | True Negative (TN) |

The confusion matrix allows for detailed error analysis and class-specific performance insights.

### 3.5.2 Accuracy

Accuracy measures the proportion of total correct predictions over all predictions. It is defined as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

where TP, TN, FP, and FN denote True Positives, True Negatives, False Positives, and False Negatives, respectively.

### 3.5.3 Precision

Precision evaluates the ratio of true positives to all positive predictions made by the model. It is a measure of exactness:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

A high precision indicates that the model returns more relevant results than irrelevant ones.

### 3.5.4 Recall

Recall (or Sensitivity) quantifies the model's ability to identify all relevant instances:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

A high recall indicates a lower false negative rate, which is critical in disease detection tasks.

### 3.5.5 F1 Score

The F1 Score is the harmonic mean of precision and recall:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

It balances the trade-off between precision and recall.

## 3.6 Summary

This chapter detailed the data handling, feature engineering, and model development workflow for detecting fire blight in apple leaves. By incorporating both hyperspectral and RGB data, the proposed methodology aims to compare conventional hyperspectral classification techniques with multimodal learning strategies. The subsequent chapter presents the experimental results and a comparative analysis of model performance across the two approaches.

# CHAPTER 4

# RESULTS

This chapter presents the experimental results obtained from the evaluation of machine learning models (Support Vector Machine, Random Forest, and XGBoost) for the detection of fire blight in apple leaves. The performance of these models, trained on hyperspectral (HSI) features alone and subsequently on fused HSI-RGB features, is detailed. The chapter is organized into two main sections: a quantitative analysis of performance metrics obtained from the testing performed on the 30% test set and a qualitative presentation of predicted image outputs. Key metrics including accuracy, precision, recall, F1-score, and ROC-AUC are reported to provide a comprehensive assessment of model capabilities.

## 4.1. Quantitative Performance Analysis of Machine Learning Models

### 4.1.1 Models trained on Hyperspectral Data

**a) SVM**

Table 1: Performance parameters for SVM model on Hyperspectral data

| Metric | Healthy (Class 0) | Diseased (Class 1) | Overall / Macro Avg. |
|---|---|---|---|
| Precision | 0.94 | 0.92 | 0.93 |
| Recall | 0.92 | 0.94 | 0.93 |
| F1-score | 0.93 | 0.93 | 0.93 |
| Accuracy | | | 0.93 |
| Support | 78,902 | 78,902 | 157,804 |

**Confusion Matrix:**

| | |
|---|---|
| **72276** | **6626** |
| **4887** | **74015** |

**b) Random Forest**

Table 2: Performance parameters for RF model on Hyperspectral data

| Metric | Healthy (Class 0) | Diseased (Class 1) | Overall / Macro Avg. |
|---|---|---|---|
| Precision | 0.92 | 0.94 | 0.93 |
| Recall | 0.94 | 0.91 | 0.93 |
| F1-score | 0.93 | 0.93 | 0.93 |
| Accuracy | | | 0.93 |
| Support | 78,902 | 78,902 | 157,804 |

**Confusion Matrix:**

| | |
|---|---|
| **72532** | **4370** |
| **6758** | **72144** |

**c) XGBoosting**

Table 3: Performance parameters for XGBoosting model on Hyperspectral data

| Metric | Healthy (Class 0) | Diseased (Class 1) | Overall / Macro Avg. |
|---|---|---|---|
| Precision | 0.91 | 0.94 | 0.92 |
| Recall | 0.95 | 0.90 | 0.92 |
| F1-score | 0.93 | 0.92 | 0.92 |
| Accuracy | | | 0.92 |
| Support | 78,902 | 78,902 | 157,804 |

**Confusion Matrix:**

| | |
|---|---|
| **74585** | **4317** |
| **7755** | **71147** |

### 4.1.2 Models trained on HIS-RGB fusion Data

**a) SVM**

Table 4: Performance parameters for SVM model on Multimodal data

| Metric | Healthy (Class 0) | Diseased (Class 1) | Overall / Macro Avg. |
|---|---|---|---|
| Precision | 0.98 | 0.91 | 0.95 |
| Recall | 0.91 | 0.98 | 0.95 |
| F1-score | 0.94 | 0.95 | 0.95 |
| Accuracy | | | 0.95 |
| Support | 52,684 | 52,684 | 105,368 |

**Confusion Matrix:**

| | |
|---|---|
| **47819** | **4865** |
| **807** | **51877** |

**b) Random Forest**

Table 5: Performance parameters for RF model on Multimodal data

| Metric | Healthy (Class 0) | Diseased (Class 1) | Overall / Macro Avg. |
|---|---|---|---|
| Precision | 0.99 | 0.97 | 0.98 |
| Recall | 0.97 | 0.99 | 0.98 |
| F1-score | 0.98 | 0.98 | 0.98 |
| Accuracy | | | 0.98 |
| Support | 62,575 | 62,574 | 125,149 |

**Confusion Matrix:**

| | |
|---|---|
| **60565** | **2010** |
| **402** | **62172** |

**c) XGBoosting**

Table 6: Performance parameters for XGBoosting model on Multimodal data

| Metric | Healthy (Class 0) | Diseased (Class 1) | Overall / Macro Avg. |
|---|---|---|---|
| Precision | 0.99 | 0.96 | 0.975 |
| Recall | 0.95 | 0.99 | 0.97 |
| F1-score | 0.97 | 0.97 | 0.97 |
| Accuracy | | | 0.9737 |
| Support | 62,575 | 62,574 | 125,149 |

**Confusion Matrix:**

| | |
|---|---|
| **59684** | **2891** |
| **403** | **62171** |

### 4.2 Prediction Outputs

To visually demonstrate and compare the qualitative performance of all trained models, a single representative image from the test set was selected. For each model, its prediction output is presented alongside the original input image and the corresponding ground truth mask. Along with that the quantitative results for that output has also been shown. This side-by-side presentation facilitates a clear visual assessment and comparison of how each model performs in segmenting or classifying the target features on the same sample image.

### 4.2.1 Models trained on Hyperspectral Data

**a) SVM**

| | |
|---|---|
| Accuracy | 0.9219436645507812 |
| Precision | 0.08654876741693462 |
| Recall | 0.9958890030832477 |
| F1 Score | 0.15925712876982495 |
| IoU | 0.08651785714285715 |


Figure 1 Prediction of SVM model on hyperspectral data

**b) RF**

| | |
|---|---|
| Accuracy | 0.978851318359375 |
| Precision | 0.25981308411214954 |
| Recall | 1.0 |
| F1 Score | 0.4124629080118694 |
| IoU | 0.25981308411214954 |


Figure 2 Prediction of RF model on hyperspectral data

**c) XGBoosting**

| | |
|---|---|
| Accuracy | 0.9740028381347656 |
| Precision | 0.22212076247003767 |
| Recall | 1.0 |
| F1 Score | 0.3635005136826375 |
| IoU | 0.22212076247003767 |

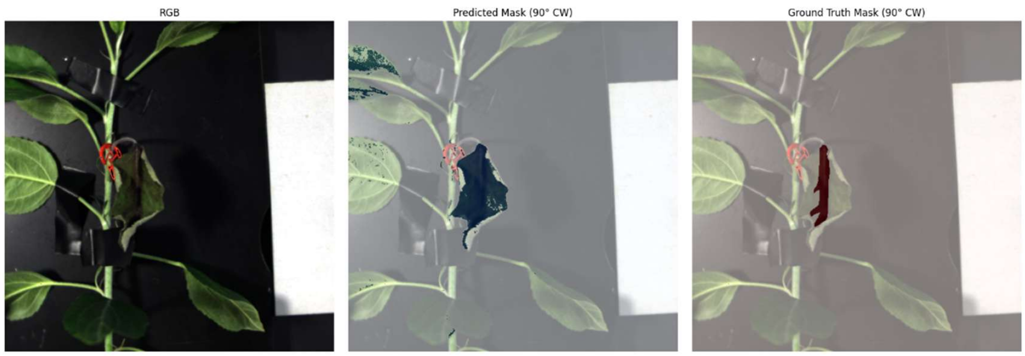


Figure 3 Prediction of XGBoost model on hyperspectral data

### 4.2.2 Models trained on HIS-RGB fusion Data

**a) SVM**

| | |
|---|---|
| Accuracy | 0.976837158203125 |
| Precision | 0.24238261738261738 |
| Recall | 0.9974306269270298 |
| F1 Score | 0.3899939722724533 |
| IoU | 0.242231374017222 |


Figure 4 Prediction of SVM model on Multimodal data

**b) RF**

| | |
|---|---|
| Accuracy | 0.9815711975097656 |
| Precision | 0.2870848708487085 |
| Recall | 0.9994861253854059 |
| F1 Score | 0.44604976493521387 |
| IoU | 0.28704250295159384 |


Figure 5 Prediction of RF model on Multimodal data

**c) XGBoosting**

| | |
|---|---|
| Accuracy | 0.976837158203125 |
| Precision | 0.24238261738261738 |
| Recall | 0.9974306269270298 |
| F1 Score | 0.3899939722724533 |
| IoU | 0.242231374017222 |


Figure 5 Prediction of XGBoost model on Multimodal data

# CHAPTER 5

# CONCLUSIONS

This chapter provides an in-depth discussion of the experimental results presented in Chapter 4 for the detection of fire blight in apple leaves using machine learning models trained on hyperspectral (HSI) features and fused HSI-RGB features. The implications of these findings are explored, followed by a summary of the main conclusions and potential avenues for future research.

## 5.1 Discussion of Results

The primary objective of this study was to evaluate the efficacy of different machine learning models (Support Vector Machine, Random Forest, and XGBoost) and the impact of multimodal feature fusion (HSI-RGB) on the accuracy of fire blight detection.

### 5.1.1 Performance on Hyperspectral (HSI) Data Alone

The results from models trained solely on HSI data (Section 4.1.1) indicate a strong baseline capability for fire blight detection.

- Both **Support Vector Machine (SVM)** and **Random Forest (RF)** achieved identical overall accuracy and macro F1-scores of 0.93. This suggests that when relying purely on spectral information, these two models can effectively distinguish between healthy and diseased leaf samples with high fidelity. The confusion matrices show a balanced performance in correctly identifying both healthy (class 0) and diseased (class 1) instances.
- **XGBoost**, while still performing very well, yielded a slightly lower accuracy and macro F1-score of 0.92. This minor difference might be attributed to the specific hyperparameter tuning or the nature of how XGBoost handles high-dimensional spectral data compared to SVM's kernel transformations or RF's ensemble of decision trees.
- The overall high performance across these models underscores the rich discriminative information contained within hyperspectral signatures for identifying physiological changes associated with fire blight.

### 5.1.2. Impact and Efficacy of HSI-RGB Feature Fusion

The introduction of RGB-derived features, fused with HSI data (Section 4.1.2), led to a noticeable improvement in performance across all three machine learning models.

- **SVM** saw its accuracy increase from 0.93 to 0.95, and its macro F1-score also rose to 0.95. This indicates that the addition of spatial and colorimetric information from RGB images provided complementary features that SVM could leverage for better class separation.
- **Random Forest (RF)** demonstrated the most substantial improvement with feature fusion. Its accuracy surged from 0.93 to 0.98, and its macro F1-score similarly reached 0.98. This suggests that RF is particularly adept at utilizing the combined feature set, possibly due to its ability to handle diverse feature types and interactions effectively through its ensemble structure. RF emerged as the top-performing model quantitatively on the fused dataset. The confusion matrix for RF with fused data shows a very low number of misclassifications (e.g., only 402 false negatives for class 1).
- **XGBoost** also benefited significantly from fusion, with accuracy improving from 0.92 to 0.9737 and the macro F1-score to 0.97. This places XGBoost as a very strong performer on the fused dataset, nearly matching Random Forest.

The consistent improvement across all models upon feature fusion strongly suggests that the spatial context, color information, and potentially texture (depending on specific RGB features extracted, though not explicitly detailed in the methodology from the results chapter) from RGB images provide valuable information that HSI data alone may lack or represent differently. HSI excels at capturing subtle spectral changes indicative of plant stress, while RGB can offer clearer macroscopic visual cues. Their combination evidently provides a more holistic representation of the leaf's condition.

### 5.1.3 Qualitative Analysis of Prediction Outputs

The qualitative results presented in Section 4.2, showing predictions on a single representative image, offer further insights, though they also highlight a common challenge in translating overall classification/segmentation metrics to individual image performance.

- The reported F1-scores and IoU values for this single image prediction task (e.g., RF Fused F1-score of 0.446, SVM HSI F1-score of 0.159) are substantially lower than the overall dataset macro F1-scores (which were above 0.90). This discrepancy is significant. It could imply several things:

- The chosen representative image might be particularly challenging, with subtle or complex infection patterns.
- The metrics reported in Section 4.2 (IoU) are typically used for semantic segmentation (pixel-level classification), which is a more demanding task than image-level or large-region classification. The overall metrics in Section 4.1 might reflect a more general classification accuracy over many samples/pixels averaged out, while Section 4.2 focuses on the pixel-wise accuracy of segmentation on *one* image.
- Ground truth annotation for pixel-level segmentation can be subjective and difficult, especially for early or diffuse symptoms.

- Despite the lower absolute scores in Section 4.2, the trend observed in the quantitative analysis (Section 4.1) generally holds:
    - Models trained on **fused HSI-RGB data** consistently achieved higher F1-scores and IoU values for the single image prediction compared to their HSI-only counterparts (e.g., RF HSI F1=0.412 vs. RF Fused F1=0.446; SVM HSI F1=0.159 vs. SVM Fused F1=0.389). This visually reinforces the benefit of feature fusion for improving the accuracy of localizing or segmenting infected regions.
    - **Random Forest** generally yielded the best F1-score and IoU in both HSI-only and fused scenarios for this specific image, aligning with its top quantitative performance on the broader test set.
- The high recall values (often 1.0 or close to 0.99) for the single image predictions, especially for RF and XGBoost, suggest that the models are very good at identifying *potential* infected pixels (low false negatives for the positive class on that image). However, the lower precision values indicate that they might also be over-segmenting or including some healthy pixels in their "diseased" prediction (higher false positives for the positive class on that image), leading to the more moderate F1 and IoU scores.

### 5.1.4 Comparative Model Performance

Overall, Random Forest emerged as the most effective model, particularly when utilizing the fused HSI-RGB feature set, achieving an accuracy and macro F1-score of 0.98. XGBoost also demonstrated excellent and highly competitive performance with fused data. SVM, while improving with fusion, was slightly outperformed by the ensemble methods in the fused scenario. This aligns with literature where ensemble models like Random Forest and XGBoost are often found to be robust and high-performing for complex classification tasks with diverse feature sets.

### 5.1.5 Implications and Significance

The findings of this study have important implications for the development of automated plant disease detection systems. The demonstrated improvement with HSI-RGB fusion suggests that future systems should aim to integrate data from multiple sensor modalities to achieve higher accuracy and reliability. Early and accurate detection of fire blight is crucial for timely intervention and management, which can prevent significant economic losses in apple orchards. The high accuracies achieved, particularly with Random Forest on fused data, indicate a strong potential for practical application.

### 5.1.6 Limitations of the Study

While the results are promising, some limitations should be acknowledged:

- **Dataset Specificity:** The performance was evaluated on a specific dataset. Generalizability to other apple cultivars, different geographical locations, varying environmental conditions, or other diseases would require further testing.
- **Computational Cost:** The study focused on accuracy metrics. A comparison of training and inference times for the different models and feature sets would be valuable for assessing practical deployability.
- **Feature Engineering:** The specific HSI bands selected and RGB features extracted were based on common practices. Further optimization of feature selection or extraction techniques might yield additional performance gains.

## 5.2 Conclusions

Based on the comprehensive experimental evaluation, the following key conclusions can be drawn:

1. **Effectiveness of HSI Data:** Hyperspectral imaging data alone provides substantial information for detecting fire blight in apple leaves, with models like SVM and Random Forest achieving high accuracies (0.93 F1-score).
2. **Superiority of Multimodal Fusion:** The fusion of HSI spectral features with RGB-derived spatial/colorimetric features significantly enhances the performance of all tested machine learning models (SVM, Random Forest, and XGBoost) for fire

blight detection. This underscores the value of integrating complementary information from different sensor modalities.

3. **Random Forest as the Top Performer:** The Random Forest model, when trained on the fused HSI-RGB feature set, demonstrated the highest overall quantitative performance, achieving an accuracy and macro F1-score of 0.98. It also showed strong qualitative results in the single-image prediction task.
4. **Qualitative Corroboration:** Visual inspection of prediction outputs confirmed that fused models generally provided more precise delineation of infected regions compared to models using only HSI data, despite the overall quantitative metrics for the single image being lower than for the entire test set.
5. **Potential for Practical Application:** The high levels of accuracy achieved, particularly with feature fusion, indicate a strong potential for developing practical, automated systems for early fire blight detection. Such systems can aid in precision agriculture, enabling timely and targeted interventions to mitigate disease spread and economic damage.

In summary, this research successfully demonstrates that a multimodal approach, specifically the fusion of HSI and RGB data, coupled with robust machine learning algorithms like Random Forest, offers a highly effective strategy for the detection of fire blight in apple leaves.

### 5.3 Future Work

Building upon the findings of this study, several avenues for future research can be pursued:

1. **Exploration of Deep Learning Models:** Investigate the application of deep learning architectures (e.g., CNNs, Vision Transformers, multimodal deep fusion networks) for this task, which might automatically learn more complex hierarchical features from HSI and RGB data.
2. **Advanced Fusion Techniques:** Explore more sophisticated feature fusion techniques beyond early concatenation, such as attention mechanisms, intermediate fusion, or model-level fusion.
3. **Temporal Analysis:** Leverage the temporal aspect of the dataset more explicitly, perhaps using recurrent neural networks (RNNs) or LSTMs to model disease progression over time.
4. **Expanded Dataset and Generalizability Testing:** Evaluate the developed models on larger, more diverse datasets encompassing different apple cultivars, geographical regions, varying environmental conditions, and other similar plant diseases to assess robustness and generalizability.

5. **Real-World Field Deployment:** Develop and test a prototype system for real-time or near real-time fire blight detection in field conditions, possibly using UAV-mounted HSI and RGB sensors. This would also involve addressing challenges like varying illumination and complex backgrounds.
6. **Interpretability of Models:** Further investigate the features deemed most important by the models (e.g., using SHAP for Random Forest and XGBoost) to gain deeper insights into the spectral and spatial characteristics indicative of fire blight.
7. **Cost-Benefit Analysis:** Conduct a cost-benefit analysis for implementing such detection systems in commercial orchards to assess their economic viability.
8. **Clarification of Test Set Metrics:** Re-evaluate or clarify the 'support' figures in the classification reports to ensure consistent comparison bases between HSI-only and fused model evaluations if they were indeed intended to be on the exact same 30% test set.

# REFERENCES


1. FAOSTAT (2023). Global apple production statistics.
2. Wikipedia (2024). List of countries by apple production.
3. Statista (2024). Global top apple-producing countries.
4. Virginia Tech (2024). Economic impact of fire blight.
5. Cornell CALS (2024). Fire blight symptoms and management.
6. Zhang et al. (2023). Fusion of HSI and RGB for soybean defect detection. Frontiers in Plant Science.
7. MDPI Remote Sensing (2020). Hyperspectral analysis of fire blight-infected apple leaves.
8. Frontiers in Plant Science (2019). High-throughput phenotyping of fire blight.
9. Singh et al. (2020). LAMP-based detection of fire blight. Journal of Plant Pathology.
10. Xiao et al. (2022). UAV multispectral detection of apple fire blight. Computers and Electronics in Agriculture.
11. IJETT (2022). Hybrid XGBoost-SVM model for leaf disease detection.
12. PMC (2022). Hyperspectral data dimensionality reduction.
13. HIPPA Project (2024). Early apple disease detection using hyperspectral imaging.
14. References
15. Aćimović, S. G., et al. (2023). Fire blight rootstock infections causing apple tree death. Frontiers in Horticulture, 3, 1082204.
16. Assudani, P. J., & Krishna, V. R. (2025). Multimodal Deep Learning Framework for Proactive Plant Disease Diagnosis. Metallurgical and Materials Engineering, 31(4), 153-161.
17. Chu, X., et al. (2020). Hyperspectral imaging and SVM for corn seed disease classification. Frontiers in Plant Science, 11, 1120968.
18. Cornell CALS (2024). Fire Blight Fact Sheet. https://cals.cornell.edu/integrated-pest-management/outreach-education/fact-sheets/fire-blight-fact-sheet
19. Frontiers in Horticulture (2023). Fire blight rootstock infections causing apple tree death. Frontiers in Horticulture, 3, 1082204.
20. Frontiers in Plant Science (2022). Maize disease detection based on spectral recovery from RGB images, 12, 1056842.
21. Frontiers in Plant Science (2023a). Current and emerging trends in techniques for plant pathogen detection, 13, 1120968.

22. Frontiers in Plant Science (2023b). Detection of citrus diseases in complex backgrounds based on multimodal feature fusion and deep learning, 13, 1280365.
23. Frontiers in Plant Science (2023c). An effective approach for plant leaf diseases classification based on DeepPlantNet, 13, 1212747.
24. Frontiers in Plant Science (2024). Hyperspectral imaging for the detection of plant pathogens in seeds, 15, 1387925.
25. Imec (2024). Hyperspectral imaging for agriculture – precision farming. https://www.imechyperspectral.com/en/applications/hyperspectral-imaging-agriculture
26. IJIRT (2025). Plant Disease Classification Using Artificial Intelligence: A Comparative Study of Machine Learning Algorithms, 11(9), 2588-2591.
27. Microbe Notes (2024). Fire Blight Disease: Symptoms, Causes & Effective Control. https://microbenotes.com/fire-blight-disease/
28. Nature (2023). Recent advances in plant disease severity assessment using deep learning. Scientific Reports, 13, 29230.
29. PMC (2022). Comparison of Selected Dimensionality Reduction Methods for Hyperspectral Data in Plant Disease Detection, 8749520.
30. Pixxel (2024). Applications of Hyperspectral Imagery in Vegetation Stress. https://www.pixxel.space/blogs/the-applications-of-hyperspectral-imagery-in-vegetation-stress
31. SSRN (2023). Plant Disease Detection Using Machine Learning Models. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4483822
32. Wikipedia (2024). Fire blight. https://en.wikipedia.org/wiki/Fire_blight
33. Wu, Y., et al. (2022). Hyperspectral imaging for peanut seed disease detection. Frontiers in Plant Science, 13, 1158933.
34. Zhang, Y., et al. (2021). Maize disease detection based on spectral recovery from RGB images. Frontiers in Plant Science, 12, 1056842.

# DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

## PLAGIARISM VERIFICATION

Title of the Thesis ________________________________________________

________________________________________________________________

________________________________________________________________

Total Pages _______________

Name of the Scholar____________________________________

Supervisor

(1) _____________________________________________

Department_____________________________________________________

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used:____________________

Similarity Index: __________________

Total Word Count: ________________

Date: ___________

**Candidate's Signature** **Signature of Supervisor**

**CV**

# VAIBHAV SHARMA

2K23/SPD/04

+91 9971215717
vaibhavsharma_23spd04@dtu.ac.in

| EDUCATION | | | |
|---|---|---|---|
| M.TECH (SPD) | 2023-2025 | Delhi Technological University, Rohini, New Delhi -110042 | 67.76 % |
| B.TECH (ECE) | 2012-2015 | Noida Institute of Engineering Technology, Greater Noida | 61.00 % |
| Polytechnic (ECE) | 2010-2012 | S B S T C Ferozepur, Punjab | 62.00 % |
| UP (Class XII) | 2010 | Dr. D P S V M I C Bilari, Moradabad | 64.40 % |
| UP (Class X) | 2008 | Dr. D P S V M I C Bilari, Moradabad | 54.50 % |

## OBJECTIVE

Dedicated electronics engineering professional with a passion for semiconductor technology seeking a challenging role in the semiconductor industry. Eager to leverage expertise in circuit design, testing, and optimization to contribute to innovative semiconductor projects. Committed to continuous learning of the latest advancements in semiconductor technology.

| Technical SKILLS | |
|---|---|
| **Interest Areas:** | Digital Electronics, Digital CMOS IC Design, Analog IC Design, Memory Circuits-SRAM/DRAM |
| **Tools:** | LT Spice, MPLAB, Proteus, MATLAB |
| **Languages:** | Verilog |

## ACADEMIC PROJECTS

1. **PFD-Phase Frequency Detector circuit design in Verilog**
   Implemented digital Phase frequency detector circuit in Verilog, for two frequency signals (Fref & Feed)
2. **11011 Sequence Detector design in Verilog**
   Implemented using Mealy & Moore FSMs, with both overlapping & non overlapping configurations.
3. **Vehicle Parking System design using RFID Technology**
   Implemented RFID-based vehicle authentication system for secure parking management, ensuring authorized vehicle access and efficient parking allocation.
4. **Beam forming in 5G**
   Implemented Beamforming; a particular processing technique for signals that allow for directional transmission or reception.

## ACHIEVEMENTS

1. **Qualified GATE 2023**
   Qualified the GATE 2023 exam in ECE subject, conducted by IIT Kanpur.

2. **Industrial Training in BSNL**
   Received comprehensive training in broadband communication and data networking at BSNL, equipping with the skills to optimize network performance and ensure seamless data transmission.