Early Detection and Prediction of Cracking in Grapes using Hyperspectral Imaging and Machine Learning

By Yovel Gani

Supervised by:
Dr. Alon Sela, Prof. Victor Alchanatis, Prof. Irad Ben-Gal

August 2026

A thesis submitted toward the degree of Master of Science (M.Sc.)

Abstract

Grape cracking is a severe physiological disorder that leads to significant economic losses in the viticulture industry by reducing marketable yield and increasing susceptibility to secondary infections like sour rot. Early and non-destructive detection of cracking potential is critical for implementing preventative measures. This research explores the feasibility of using hyperspectral imaging (HSI) in the 400-1000 nm range, coupled with machine learning, for the identification and prediction of cracks in table grapes.

The study followed a multi-stage approach: initial identification, prediction modeling, and comparative analysis of spectral signatures. For the identification task, a pixel-level dataset was meticulously created by manually selecting and labeling pixels from healthy and cracked berries. An XGBoost classifier trained on these spectral signatures achieved outstanding performance, with an accuracy of approximately 97% and an AUC of 0.99, demonstrating that a distinct spectral difference exists between intact and cracked grape tissue.

However, applying this pixel-level model to whole-image segmentation proved challenging. The model produced noisy results, often misclassifying background elements such as plastic ties or dry twigs, which exhibited spectral signatures similar to those of decay associated with cracks.

To overcome this limitation and move towards prediction, the research shifted to analyzing the temporal evolution of mean spectral signatures of entire grape clusters under different irrigation regimes. This approach revealed significant and consistent trends over time, particularly in the Near-Infrared (NIR) region (700-900 nm). These findings suggest that changes in the average spectral signature can serve as a robust indicator of physiological stress and cracking susceptibility, potentially offering a reliable method for early warning. The results demonstrate that while pixel-level identification is highly accurate under controlled conditions, a macro-level analysis of mean spectral signatures provides a more practical and promising avenue for predicting grape cracking in complex field environments.

Contents

1 Introduction **9**

1.1 The Problem of Fruit Cracking . . . 9

1.2 Economic Impact . . . 9

1.3 Research Motivation . . . 9

1.3.1 Main Research Question . . . 10

1.4 Research Objectives and Hypotheses . . . 10

1.4.1 Research Objectives . . . 10

2 Literature Review **11**

2.1 Grape Cracking as an Agricultural and Physiological Problem . . . 11

2.1.1 Definition and visual manifestation of fruit cracking . . . 11

2.1.2 Phenological timing (veraison, late ripening stages) . . . 11

2.1.3 Environmental and agronomic triggers . . . 12

2.2 Economic Impact and Practical Consequences . . . 12

2.2.1 Yield losses and reduction in marketable produce . . . 12

2.2.2 Quality degradation and shelf-life reduction . . . 12

2.2.3 Limitations of reactive management strategies . . . 13

2.3 Physiological and Biomechanical Mechanisms of Grape Cracking . . . 13

2.3.1 Berry water relations and turgor pressure dynamics . . . 14

2.3.2 Cuticle structure, micro-cracks, and mechanical failure . . . 14

2.3.3 Hormonal influences and agronomic interventions . . . 15

2.3.4 Cultivar-dependent susceptibility and variability . . . 15

2.3.5 Skin pathology as an indicator of structural instability . . . 15

2.3.6 Summary . . . 16

2.4 Precision Agriculture and Non-Destructive Sensing . . . 16

2.4.1 Principles of precision agriculture and data-driven crop management 16

2.4.2 Role of proximal and remote sensing technologies . . . 16

2.4.3 Limitations of conventional physiological measurements . . . 17

2.4.4 Computer Vision and AI in Agriculture . . . 17

2.5 Hyperspectral Imaging for Agricultural Applications . . . 18

2.5.1 Fundamentals of Reflectance and Spectral Signatures 18
2.5.2 Spectral Imaging Modalities and the Electromagnetic Spectrum . 19
2.5.3 Advantages of Hyperspectral Imaging over RGB Imaging 20
2.5.4 VIS–NIR Sensitivity to Physiological and Structural Changes . . 21
2.5.5 Applications in Fruit Quality Assessment and Defect Detection . . 21
2.6 Machine Learning for Hyperspectral Data Analysis 22
2.6.1 Deep learning and non-linear models 23
2.6.2 Pixel-level versus image-level modeling paradigms 24
2.6.3 Challenges: class imbalance, background clutter, and generalization 25
2.7 Related Work on Fruit Defect Detection and Prediction 25
2.7.1 HSI-based detection of visible fruit defects 26
2.7.2 Temporal and Longitudinal Spectral Analysis 27
2.7.3 Attempts Toward Early and Pre-Symptomatic Defect Prediction . 27
2.7.4 Limitations of Existing Approaches under Field Conditions 28
2.8 Research Gap and Study Positioning 29
2.8.1 Summary of Established Knowledge 29
2.8.2 Identified Methodological and Application Gaps 30
2.8.3 Lack of Early, Non-Destructive Crack Prediction Frameworks . . . 30
2.8.4 Positioning of the Present Study within the Literature 31

3 Research Objectives and Hypotheses 32
3.1 Research Objectives 32
3.2 Research Hypotheses 32

4 Materials and Methods 35
4.1 Data Acquisition 35
4.1.1 Study Area and Experimental Design 35
4.1.2 Imaging Systems 37
4.1.3 Data Collection Protocol 40
4.2 Preprocessing and Dataset Preparation 42
4.2.1 Data Management and Pre-processing 43
4.2.2 Pixel-Level Dataset Preparation 43
4.2.3 Expansion to Multi-Class Dataset 45
4.2.4 Whole-Image Dataset Construction and Experimental Design . . 46
4.2.5 Automated Signature Extraction for Temporal Analysis 47
4.3 Modeling and Evaluation Methodology 48
4.3.1 Global Spectral Preprocessing 49
4.3.2 Spectral Discriminability and Informative Wavelength Analysis . . 50
4.3.3 Pixel-Level Classification Framework 51

4.3.4 Anomaly Detection Framework 53
4.3.5 Dimensionality Reduction Strategy 54
4.3.6 Whole-Image Detection and Aggregation Pipeline 55
4.3.7 Evaluation Protocols and Performance Metrics 57

5 Experimental Results **59**
5.1 Overview of Experimental Results 59
5.2 Spectral Signature Characterization 59
5.2.1 Mean Spectral Signatures of Healthy and Cracked Tissue 59
5.2.2 Spectral Separability Across Wavelength Regions 61
5.3 Pixel-Level Classification Performance 63
5.3.1 Experimental Setup and Evaluation Protocol 63
5.3.2 Binary Pixel-Level Classification Results 63
5.3.3 Pixel-Level Classification Results Under Different Labeling and Balancing Regimes (10 classes) 65
5.3.4 Autoencoder-Based Pixel-Level Anomaly Detection 68
5.3.5 Full-Image Inference: Early vs. Late Season Experiments 72
5.3.6 Wavelength selection (BFS) setup 75
5.3.7 Full-image wavelength-selection results 76
5.3.8 BFS Wavelength-Selection Stability Analysis 76

6 Discussion **84**
6.1 Summary of Key Findings 84
6.2 Spectral and Physiological Interpretation 85
6.2.1 VIS Region: Pigments and Skin Integrity 85
6.2.2 NIR and Red-Edge Sensitivity to Water Stress 86
6.3 Why Pixel-Level Detection Works in Controlled Settings but Fails in Field Conditions 87
6.4 Temporal Aggregation as a Robust Strategy for Early Crack Prediction . 88
6.5 Feature Selection and Spectral Stability 90
6.5.1 Dimensionality Reduction via BFS 90
6.5.2 Stability Across Random Seeds and Sensor Design Implications . 90
6.6 Generalization Gaps Between Calibration and Test Sets 91
6.7 Supervised Classification versus Anomaly Detection 93
6.8 Practical Implications for Precision Viticulture 94
6.9 Limitations of the Study 95
6.10 Future Research Directions 96
6.11 Concluding Perspective 98

List of Figures

2.1 The formula for reflectance Blackburn, 2007. 19
2.2 Hyperspectral signature of barley grains with varying moisture content T. Sun et al., 2020. 19
2.3 Segments of the electromagnetic spectrum commonly used in hyperspectral imaging (Middleton Spectral Vision, n.d.). 20

4.1 Map of the experimental vineyard showing the five randomized blocks (A–E) and the four irrigation regimes applied within each block. 36
4.2 Hyperspectral imaging setup and data acquisition using the Specim IQ system under vineyard conditions. 37
4.3 Thermal imaging system and example thermal data acquired from grape clusters under vineyard conditions. 39
4.4 Graphical user interface of the HSI – RGB annotation tool. (1) Hyperspectral image visualized as an RGB composite with a SAM2-generated segmentation overlay, (2) wavelength-selectable grayscale HSI band with the same SAM2 segmentation mask, and (3) RGB image of the same grape cluster captured on the same day using a Canon EOS R5. (4) Spectral signature extracted from a single pixel within the SAM2-segmented region, and (5) autoencoder reconstruction of the same spectral signature, discussed later in this work. 44
4.5 Schematic overview of the whole-image detection and spatial aggregation pipeline. 56

5.1 Mean spectral signatures (± standard deviation) for healthy and cracked tissue across the VIS–NIR range. 60
5.2 Wavelength-wise spectral separability metrics between healthy and cracked tissue, including absolute mean difference, Cohen's d, and Fisher score. . 61
5.3 Regional visual summary of mean reflectance (±1 STD) and absolute mean difference between healthy and cracked tissue in the VIS and NIR ranges. This figure provides an intuitive visualization of the regional trends quantified in Table 5.2. 62

5.4 BFS results: CRACK F1 and CRACK PR–AUC as a function of the number of selected wavelengths. 75
5.5 BFS performance curves with markers indicating the maximum score and the first wavelength counts where the score drops beyond 0.5% and 1.0% from the maximum. 75
5.6 Distribution and popularity of the top-30 selected wavelengths across five varying-split seeds. 78
5.7 Wavelength popularity at different bin sizes (1, 5, 10, 20 nm) for the top-30 subsets. 79
5.8 Wavelength stability heatmap (5 nm bins) for the top-30 subsets. Each cell shows how many of the five seeds selected a wavelength within that bin. 80
5.9 Wavelength selection frequency along the spectrum (2 nm sliding window) for the top-30 subsets. 81
5.10 Distribution and popularity of the top-11 selected wavelengths across five varying-split seeds. 81
5.11 Wavelength popularity at different bin sizes (1, 5, 10, 20 nm) for the top-11 subsets. 82
5.12 Wavelength stability heatmap (5 nm bins) for the top-11 subsets. Each cell shows how many of the five seeds selected a wavelength within that bin. 83
5.13 Wavelength selection frequency along the spectrum (2 nm sliding window) for the top-11 subsets. 83

List of Tables

1.1 Estimated financial losses from fruit cracking in major crops.(Alchanatis et al., 2022). 10

4.1 Overview of datasets used for crack detection and prediction experiments. 42
4.2 Evaluation metrics for pixel-level and anomaly detection experiments . . 58
4.3 Evaluation metrics for whole-image and patch-based detection 58

5.1 Summary statistics of reflectance values across spectral regions for healthy and cracked tissue. 60
5.2 Regional spectral separability metrics between healthy and cracked tissue. Mean values are averaged across wavelengths within each range, while peak values indicate the maximum separability and the corresponding wavelength. 62
5.3 Pixel-level classification metrics (unbalanced training set). Values are reported as mean ± standard deviation across folds where available. Precision and recall are reported for the *Cracked* class; F1 is the weighted F1-score. 64
5.4 Pixel-level classification metrics (balanced training set). Values are reported as mean ± standard deviation across folds where available. Precision and recall are reported for the *Cracked* class; F1 is the weighted F1-score. 64
5.5 Pixel-level classification metrics (**unbalanced** training set). Values are reported as mean ± standard deviation across folds. 66
5.6 Pixel-level classification metrics (**balanced** training set). Values are reported as mean ± standard deviation across folds. 67
5.7 Pixel-level anomaly detection performance using autoencoders trained on non-cracked samples (unbalanced sampling). Values are reported as mean ± standard deviation across folds. Metrics correspond to the *Cracked* class. 69
5.8 Pixel-level anomaly detection performance using autoencoders trained on non-cracked samples (balanced sampling). Values are reported as mean ± standard deviation across folds. Metrics correspond to the *Cracked* class. 69

5.9 Pixel-level anomaly detection performance using autoencoders trained exclusively on *Cracked* pixels and evaluated in a multi-class setting (unbalanced sampling). Values are reported as mean across LOGO folds. Metrics correspond to the *Cracked* class. 70
5.10 Pixel-level anomaly detection performance using autoencoders trained exclusively on *Cracked* pixels and evaluated in a multi-class setting (balanced sampling). Values are reported as mean ± standard deviation across LOGO folds. Metrics correspond to the *Cracked* class. 70
5.11 Pixel-level anomaly detection performance using autoencoders trained on all semantic classes except *Cracked* and evaluated in a multi-class setting (unbalanced sampling). Values are reported for the held-out test split. Metrics correspond to the *Cracked* class. 71
5.12 Pixel-level anomaly detection performance using autoencoders trained on all semantic classes except *Cracked* and evaluated in a multi-class setting (balanced sampling). Values are reported for the held-out test split. Metrics correspond to the *Cracked* class. 71
5.13 Dataset composition for full-image crack-detection experiments under early- and late-season conditions. Row 1 was used for calibration, while Row 2 served as an independent test set. 72
5.14 Full-image crack detection performance for early- and late-season experiments. Metrics are reported separately for calibration and test splits. Confusion-matrix entries correspond to the number of grape clusters classified as cracked or healthy. All derived statistics refer to the *Cracked* class. 73
5.15 Full-image crack detection using autoencoder-based anomaly detection (128–64–32, trained on *Cracked* pixels). Reconstruction errors (MSE) were converted to pseudo-probabilities via sigmoid transformation; the same post-processing pipeline (thresholding, morphological filtering, patch-based aggregation) was applied. Metrics are reported for calibration and test splits under early- and late-season conditions. All derived statistics refer to the *Cracked* class. 74
5.16 Summary of BFS wavelength-count thresholds for CRACK metrics. . . . 75
5.17 Full-image crack-detection performance for different wavelength subsets under early- and late-season conditions. 76
5.18 Model Generalization Analysis: Comparison between Calibration and Test MCC. 76
5.19 BFS stability: per-seed results under varying train/test splits. 77
5.20 Pairwise Jaccard similarity of top-30 wavelengths across seeds (varying split). Mean Jaccard = 0.493. 77

5.21 Pairwise Jaccard similarity of top-11 wavelengths across seeds (varying split). Mean Jaccard = 0.398. 77

Chapter 1

Introduction

1.1 The Problem of Fruit Cracking

Fruit cracking is a significant physiological disorder affecting a wide range of commercially important crops, including sweet cherries, tomatoes, pomegranates, and particularly table grapes (Vitis vinifera L.). The phenomenon is characterized by the rupture of the fruit's skin (cuticle and epidermis), exposing the internal flesh to the environment. These cracks render the fruit unmarketable and create entry points for pathogens, leading to rapid decay from secondary infections such as sour rot. Cracking typically occurs during the final stages of fruit ripening, a period when the fruit is most valuable. It is primarily caused by an imbalance between the turgor pressure within the fruit and the mechanical strength of its skin, a condition often exacerbated by environmental factors like rain, high humidity, and irrigation practices.

1.2 Economic Impact

The economic consequences of grape cracking are substantial. Direct losses can amount to up to 30% of the marketable yield. For table grapes, this translates to an estimated financial loss of €625 per ton, culminating in a total industry loss of approximately €175 million annually, based on European market data (Table 1.1). Beyond direct yield loss, cracking negatively impacts fruit appearance and quality, leading to price reductions of 20-50% and a shortened shelf-life, further diminishing profitability for growers.

1.3 Research Motivation

Current methods for managing cracking risk are reactive and often insufficient. There is a critical need for a non-destructive, reliable method to detect early signs of cracking susceptibility before visible fissures appear. Hyperspectral imaging (HSI) has emerged

Table 1.1: Estimated financial losses from fruit cracking in major crops.(Alchanatis et al., 2022).

Fruit	**Prod. (K t)**	**Coverage (K ha)**	**Price (€/kg)**	**Loss (%)**	**Fin. Loss (€/t)**	**Total loss (M€)**
Table grapes	280	7.0	2.5	25	625	175.0
Citrus	2300	58.5	1.1	20	220	506.0
Pomegranate	43	1.0	3.3	25	825	35.5
Sweet cherry	560	37.3	2.0	20	400	224.0

as a powerful tool in precision agriculture for assessing plant health and fruit quality due to its ability to capture detailed spectral information beyond the visible range. This research aims to investigate the potential of HSI and machine learning to address the challenge of grape cracking.

1.3.1 Main Research Question

Can hyperspectral imaging and machine learning models be used to identify existing grape cracks and predict their future occurrence?

1.4 Research Objectives and Hypotheses

1.4.1 Research Objectives

1. **Identification:** To develop a high-accuracy classification model to differentiate between healthy and cracked grape tissue at the pixel level using their hyperspectral signatures.

2. **Prediction:** To investigate methods for predicting cracking potential by analyzing temporal changes in the spectral characteristics of grapes over the ripening period.

3. **Comparative Analysis:** To compare the mean spectral signatures of grape clusters under different agricultural treatments (e.g., irrigation regimes) to identify spectral indicators of stress related to cracking.

Chapter 2

Literature Review

2.1 Grape Cracking as an Agricultural and Physiological Problem

2.1.1 Definition and visual manifestation of fruit cracking

Fruit cracking in table grapes (*Vitis vinifera* L.) is defined as a physiological disorder in which the berry skin ruptures, compromising fruit quality and marketability (Guirao et al., 2025). The visible manifestation typically appears as surface splits or fissures on the berry, reflecting the mechanical failure of the cuticle–skin system under stress (La Spada et al., 2024). Recent physiological evidence further links cracking susceptibility to cultivar-dependent differences in water uptake and internal tissue expansion during critical developmental phases (Yu et al., 2025).

2.1.2 Phenological timing (veraison, late ripening stages)

The phenological development of grape berries is characterized by well-defined stages that culminate in ripening, with veraison representing a critical transition point from berry growth to maturation. Veraison marks the onset of ripening and is associated with coordinated physiological changes, including berry softening, shifts in water and solute transport, and the initiation of biochemical processes that define fruit quality (Guirao et al., 2025). The timing and duration of veraison and subsequent late ripening stages are strongly influenced by thermal accumulation and environmental conditions, leading to variability in ripening dynamics across seasons and cultivars. Recent phenological studies have demonstrated that growing degree days (GDD) provide a robust framework for describing the progression from veraison to harvest, with differences in heat accumulation directly affecting the length of late ripening stages and final harvest timing (Monteiro et al., 2025).

2.1.3 Environmental and agronomic triggers

Environmental and agronomic factors play a central role in modulating grape berry development and triggering physiological disorders during ripening, particularly under conditions of altered water availability and tissue hydration. Variations in water uptake driven by aquaporin activity have been shown to directly influence internal berry pressure and cracking susceptibility (Yu et al., 2025), while changes in cuticle structure and integrity during ripening modulate the berry's mechanical resistance to environmental stress (La Spada et al., 2024). In parallel, agronomic practices affecting calcium transport and availability after veraison further influence fruit firmness and skin stability, thereby impacting the berry's response to environmental stressors during late ripening stages (Guirao et al., 2025).

2.2 Economic Impact and Practical Consequences

2.2.1 Yield losses and reduction in marketable produce

Fruit cracking leads to substantial yield losses in table grape production by directly reducing the proportion of berries suitable for fresh market commercialization. Even limited cracking of the berry skin is sufficient to render fruit unmarketable, as it compromises visual quality and accelerates postharvest deterioration processes such as water loss, tissue softening, and increased susceptibility to decay (C. Zhang et al., 2025). As a result, a significant discrepancy often emerges between biological yield and effective marketable yield, particularly in commercial vineyards exposed to unfavorable conditions during ripening. Based on large-scale production and economic estimates, cracking-related losses in table grapes can reach approximately €175 million, assuming a 25% yield loss rate, an average production volume of 280 kt, and a mean market price of 2.5 €/kg (Alchanatis et al., 2022). At a broader scale, fruit cracking is recognized as a major physiological disorder that contributes to yield instability and economic loss across multiple fruit crops, including grapes, by increasing preharvest rejection rates and postharvest waste (Santos et al., 2023).

2.2.2 Quality degradation and shelf-life reduction

Reactive or purely reactive management strategies for fruit cracking have shown limited effectiveness in table grape production, where berry susceptibility is strongly governed by the interaction between environmental conditions during ripening and intrinsic skin and tissue properties. Reviews addressing fruit cracking across fleshy fruit crops emphasize that abrupt changes in water availability, humidity, and temperature frequently exceed the buffering capacity of short-term corrective interventions once berries reach vulnera-

ble developmental stages, rendering post-hoc management largely insufficient (Seo et al., 2022; Xu et al., 2025). In grapes, even minor cracking events are sufficient to compromise marketability, which further amplifies the economic consequences of delayed or reactive responses. Evidence from field-based mitigation studies in other thin-skinned fruits indicates that targeted preharvest treatments, such as salicylic acid sprays, can reduce cracking incidence but do not eliminate it, suggesting that similar reactive approaches in grapes are inherently constrained and highly dependent on timing and environmental context (Karantzi et al., 2021). Collectively, these findings highlight the limitations of reactive management strategies in table grapes and underscore the need for preventive, monitoring-driven, and predictive approaches to cracking risk mitigation.

2.2.3 Limitations of reactive management strategies

Reactive or purely reactive management strategies for berry cracking in table grapes remain inherently limited, largely because cracking is triggered by rapid shifts in water relations and microclimatic conditions during ripening, interacting with cultivar-specific skin and cuticle properties. Evidence from table-grape studies indicates that cracking susceptibility is tightly linked to berry structural integrity and water uptake dynamics, with critical windows during ripening in which short-term corrective actions may be insufficient to prevent splitting once mechanical resistance declines (Chang & Keller, 2021; Chang et al., 2019; Yu et al., 2025). Accordingly, while preharvest interventions aimed at strengthening the cuticle or improving tissue integrity (e.g., calcium-based or combined treatments) can contribute to improved berry skin traits and potentially reduce susceptibility, their effectiveness is typically partial and context-dependent rather than fully preventive (La Spada et al., 2024). Taken together, these findings highlight that reactive management alone is unlikely to provide robust control of cracking in commercial table-grape systems and motivate preventive, monitoring-driven, and predictive strategies that enable earlier risk identification and timely intervention (Alchanatis et al., 2022).

2.3 Physiological and Biomechanical Mechanisms of Grape Cracking

Reactive or purely reactive management strategies for fruit cracking have consistently demonstrated limited efficacy across diverse horticultural crops. Comprehensive reviews indicate that cracking is driven by complex interactions between environmental triggers—such as rainfall, humidity fluctuations, and abrupt temperature changes—and intrinsic physiological and structural fruit properties, which often overwhelm short-term or corrective interventions applied after stress exposure (Seo et al., 2022; Xu et al., 2025). As emphasized in these studies, no single management practice is capable of fully

preventing cracking once susceptible developmental stages are reached, underscoring the inherent limitations of post-hoc approaches and the need for preventive, integrated strategies. Field-based intervention trials further support this conclusion: although targeted preharvest treatments such as foliar salicylic acid applications can significantly reduce cracking incidence and improve marketable yield, their effectiveness remains partial and crop- and condition-dependent, highlighting fundamental constraints of reactive measures when applied in isolation (Karantzi et al., 2021).

2.3.1 Berry water relations and turgor pressure dynamics

Beyond softening, water uptake plays a central role. Even minor fluctuations in soil moisture or surface wetness can cause rapid increases in berry volume. When mesocarp expansion outpaces the extensibility of the skin, tensile stress accumulates in the cuticle and epidermal cell walls. As reviewed by Khadivi-Khub Khadivi-Khub, 2014, this imbalance—driven by osmotic gradients, turgor dynamics, and altered permeability—forms a generalized mechanism for cracking across thin-skinned fruit species. However, grapes are particularly sensitive because cuticular deposition slows near maturity, while mesocarp growth continues. This mismatch in developmental timing reduces the elastic reserve of the skin and predisposes berries to micro-fissuring and eventual rupture.

2.3.2 Cuticle structure, micro-cracks, and mechanical failure

The grape berry cuticle functions as both a diffusion barrier and mechanical shell. Its composition—cutin, polysaccharides, and epicuticular waxes—determines the skin's tensile strength, elasticity, and hydration behavior. Studies in sweet cherry and tomato have demonstrated that microcracks in the cuticle dramatically increase water permeability and serve as initiation sites for macroscopic cracking Emmons and Scott, 1998; Peschel and Knoche, 2005. Recent work now shows that grapes exhibit similar patterns. Chang and Keller Chang and Keller, 2021 revealed that the grape cuticle and epidermal cell walls play both overlapping and distinct roles in resisting cracking. Their microscopy studies indicate that microcracks can form days before visible splitting and propagate through weakened regions of the cuticle under mechanical stress or hydration events.

Environmental conditions exacerbate these mechanical vulnerabilities. High humidity, rain, or dew hydration increases cuticular permeability and reduces its fracture toughness. Conversely, heat and drought can cause structural fatigue through repeated cycles of dehydration and rehydration. Because these microcracks do not manifest visually, early detection requires sensitive optical or spectral tools—highlighting the relevance of hyperspectral imaging (HSI) approaches.

2.3.3 Hormonal influences and agronomic interventions

Plant growth regulators, particularly gibberellins, are widely applied in table grape production to increase berry size, reduce compactness, and improve market quality. However, these treatments also modify skin biomechanics. Lichter et al. Lichter et al., 2014 investigated the effects of GA_3 on the seeded table grape 'Zainy' and found that while berry enlargement was achieved, the treatment altered skin structure and significantly affected cracking susceptibility. Their findings highlight the trade-off between agronomic optimization and physiological stability, emphasizing the need for cultivar-specific GA_3 protocols.

Other hormones—most notably abscisic acid (ABA)—also shape berry softening and water transport. Castellarin et al. Castellarin and other, 2016 characterized the interplay between ABA metabolism, turgor loss, and cell wall modification during ripening. These hormonal signals regulate both mesocarp softness and cuticle extensibility, indirectly influencing cracking risk. Together, these works suggest that cracking is not only mechanically driven but is also hormonally modulated.

2.3.4 Cultivar-dependent susceptibility and variability

Cracking susceptibility varies widely among grape cultivars, reflecting differences in cuticle thickness, cell wall architecture, berry firmness, water relations, and patterns of sugar accumulation. Chang and Keller Chang and Keller, 2021 demonstrated cultivar-dependent variation in skin biomechanical properties, showing that small differences in cuticle thickness or epidermal morphology can translate into large differences in cracking rates. Thin-cuticle cultivars or those with rapid late-season sugar accumulation tend to be more vulnerable.

For 'Scarlotta Seedless', the cultivar investigated in the present study, several postharvest evaluations have documented sensitivity to temperature fluctuations and hydration-induced damage A. Author et al., 2022; B. Author et al., 2024. While these works focus primarily on postharvest quality, they provide important evidence that the skin of Scarlotta berries is structurally fragile compared to other table grape varieties. Given the cultivar's widespread commercial use, understanding and detecting early signs of cuticular stress is a significant applied challenge.

2.3.5 Skin pathology as an indicator of structural instability

Although not a cracking disorder per se, grape sour rot provides insight into skin structural failure. Hall et al. Hall et al., 2018 demonstrated that sour rot development is strongly associated with breaches in berry skin integrity, allowing entry of yeasts, acetic acid bacteria, and insects. These findings reinforce the broader concept that once the

cuticle–epidermis barrier is compromised—whether by microcracks, overripening, or mechanical abrasion—the fruit becomes highly vulnerable to physiological and pathological deterioration. Thus, the onset of cracking represents a pivotal point of failure in grape berry biology.

2.3.6 Summary

Overall, grape cracking emerges from the interaction of mechanical constraints, hormonal regulation, developmental processes, and cultivar-specific structural traits. Although these mechanisms have been extensively investigated, most studies focus on visible or post-symptomatic cracking stages. Consequently, early pre-symptomatic structural and physiological changes in the berry skin remain poorly characterized, highlighting the need for non-destructive optical approaches such as hyperspectral imaging.

2.4 Precision Agriculture and Non-Destructive Sensing

2.4.1 Principles of precision agriculture and data-driven crop management

Precision agriculture (PA) has emerged as a data-driven management paradigm that seeks to optimize agricultural operations through the integration of sensing technologies, data analytics, and site-specific interventions. At its core, PA relies on systematic acquisition and analysis of large heterogeneous datasets to improve efficiency, productivity, and long-term sustainability in crop production Mgendi, 2024. Recent advances in the Internet of Things (IoT), Big Data infrastructures, and artificial intelligence (AI) have further accelerated the adoption of data-centric farming practices, enabling continuous environmental monitoring and facilitating sustainable resource management at multiple spatial and temporal scales Ahmed and Shakoor, 2025.

2.4.2 Role of proximal and remote sensing technologies

The proliferation of smart sensor technologies—including proximal sensing devices, wireless sensor networks, and real-time monitoring platforms—has substantially expanded the types of agronomic variables that can be measured in situ. Such sensors now capture environmental, physiological, and operational parameters with unprecedented resolution, which supports precise and resource-efficient crop management Soussi et al., 2024.

2.4.3 Limitations of conventional physiological measurements

Managing these heterogeneous datasets—which often vary in granularity, modality, and update frequency—requires robust data engineering pipelines capable of processing information across various temporal and spatial dimensions. This complexity is further amplified by the need to deliver decision-ready outputs in formats that support real-time or near-real-time analysis Krisnawijaya et al., 2022. To address these challenges, modern PA increasingly incorporates advanced computational methods, including big data platforms, predictive modeling, and AI-driven analytics. These tools play a key role in transforming raw sensor data into predictive insights, allowing early detection of anomalies, improved planning, and automated decision-support workflows Ahmed and Shakoor, 2025; Mgendi, 2024.

2.4.4 Computer Vision and AI in Agriculture

Computer vision (CV) has become a cornerstone technology in modern agriculture, driven by the rapid digital transformation of the sector and the growing need for automated, scalable, and data-driven crop management solutions. As global food demand increases and available farmland diminishes, CV systems—particularly those based on deep learning—provide the computational foundation required for precise, efficient, and reliable agricultural operations Patrício and Rieder, 2018; H. Tian et al., 2020. These systems enable machines to acquire, interpret, and analyze visual information, supporting tasks ranging from seed quality assessment and soil analysis to plant health monitoring, weed detection, livestock management, and yield estimation Cao et al., 2025; Dhanya et al., 2022. Over the last decade, convolutional neural networks (CNNs) have emerged as the dominant architecture for agricultural CV applications due to their strong feature extraction capabilities, scalability, and adaptability to diverse imaging conditions Attri et al., 2023a; Benos et al., 2021. Recent advances such as generative adversarial networks (GANs), transfer learning, and vision transformers (ViT) further extend the potential of CV systems, enabling improved robustness under field conditions characterized by variable lighting, occlusion, and environmental noise Dhanya et al., 2022; Ghazal et al., 2024.

Despite these technological advancements, the effectiveness of CV models in agriculture is fundamentally dependent on the quality of the data used to train them. High-quality datasets must accurately represent the natural variability found in agricultural environments—including crop phenology, environmental dynamics, sensor noise, and geographical differences—to ensure generalization and prevent model bias. Poorly curated datasets, imbalanced class distributions, low-resolution imagery, or inconsistent labeling practices can lead to significant reductions in model performance, especially under real-world conditions where visual ambiguity is common Patrício and Rieder, 2018; H. Tian

et al., 2020. Several studies highlight that annotation quality is often a major limiting factor in agricultural CV, as labeling requires domain expertise and is sensitive to human subjectivity. For example, plant disease severity, fruit defects, weed species boundaries, and growth-stage differentiation may be interpreted differently across annotators, leading to inconsistencies that degrade model accuracy and reliability Benos et al., 2021; Dhanya et al., 2022. This challenge is amplified in pixel-level tasks such as segmentation, where precise delineation of plant organs or symptomatic tissue is crucial for downstream analytical accuracy.

The reliance on high-quality annotations is particularly evident in deep learning models, which are strongly data-driven and require large, well-balanced labeled datasets to achieve stable convergence. In many agricultural applications, assembling such datasets is difficult due to seasonal variability, labor-intensive labeling, and the scarcity of annotated images for rare conditions such as early-stage infections or latent physiological defects Attri et al., 2023a; Ghazal et al., 2024. To mitigate these constraints, recent research has explored synthetic data generation using GANs, semi-supervised learning, active learning, and weakly supervised labeling strategies. These approaches aim to reduce annotation costs while improving model robustness and are increasingly adopted in advanced CV pipelines for agriculture Dhanya et al., 2022. Additionally, improvements in sensor technologies—including hyperspectral, multispectral, and thermal imaging—offer richer feature representations that help compensate for noise or ambiguity in labeled datasets, strengthening both classification and segmentation performance H. Tian et al., 2020.

Overall, the literature demonstrates that while deep-learning-based CV systems have achieved remarkable progress across a wide range of agricultural applications, their true effectiveness in real-world environments hinges on the availability of high-quality, well-annotated datasets, robust model architectures, and reliable deployment pipelines. Continued innovation in data acquisition, labeling methodologies, and domain adaptation techniques will therefore be essential for advancing next-generation CV solutions capable of supporting smart, scalable, and fully automated agricultural systems.

2.5 Hyperspectral Imaging for Agricultural Applications

2.5.1 Fundamentals of Reflectance and Spectral Signatures

A central concept in spectral analysis is reflectance, defined as the ratio between the radiation reflected by a surface and the radiation received by that surface. This ratio is influenced by the material's chemical properties, surface micro-topography, and the angle

of the light source Blackburn, 2007. Reflectance is the primary value captured at each pixel for every wavelength in a hyperspectral image, providing a normalized measure that is independent of the illumination intensity.

$$Reflectance = \phi out \,/\, \phi in$$
$$\text{where } \phi = \text{radiation}$$

Figure 2.1: The formula for reflectance Blackburn, 2007.

When a material is exposed to a light source, it reflects specific portions of electromagnetic energy according to its molecular vibrations and chemical composition. This unique reaction is known as the spectral signature of a material Shaw and Manolakis, 2002. Because these signatures are specific to chemical structures, hyperspectral imaging enables reliable discrimination between visually similar targets that would otherwise appear identical under conventional RGB or multispectral sensors.

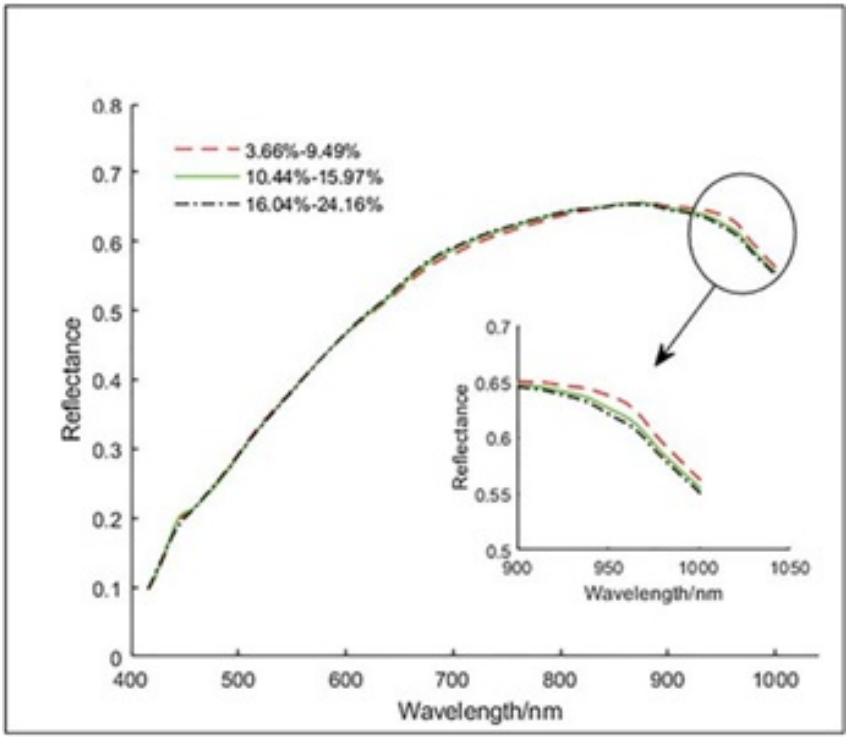


Figure 2.2: Hyperspectral signature of barley grains with varying moisture content T. Sun et al., 2020.

2.5.2 Spectral Imaging Modalities and the Electromagnetic Spectrum

Imaging systems differ substantially in the density and breadth of the spectral information they capture. Traditional RGB cameras record reflected light using only three broad

bands of visible radiation: blue, green, and red Benelli et al., 2020. Multispectral imaging (MSI) typically expands this capability by capturing between 3 to 15 distinct bands and has a spectral resolution greater than 10 nm. In contrast, hyperspectral imaging (HSI) acquires reflectance data across hundreds of narrow, contiguous spectral bands, typically with a spectral resolution of less than 5 nm Mehta, 2018.

These systems capture broad segments of the electromagnetic spectrum, ranging from the ultraviolet (UV) around 200 nm, through the visible (VIS) light range (380 nm to 800 nm), and extending into the near-infrared (NIR) up to 1700 nm Ball, 2007. Additionally, the range often includes the short-wave infrared (SWIR) from 970 nm to 2500 nm Ball, 2007; Zwinkels, 2015. This high spectral resolution results in a three-dimensional hyperspectral data cube where each spatial pixel is associated with a detailed spectral vector, enabling the detection of both surface and subsurface characteristics that are not visible through conventional imaging Hong et al., 2025.

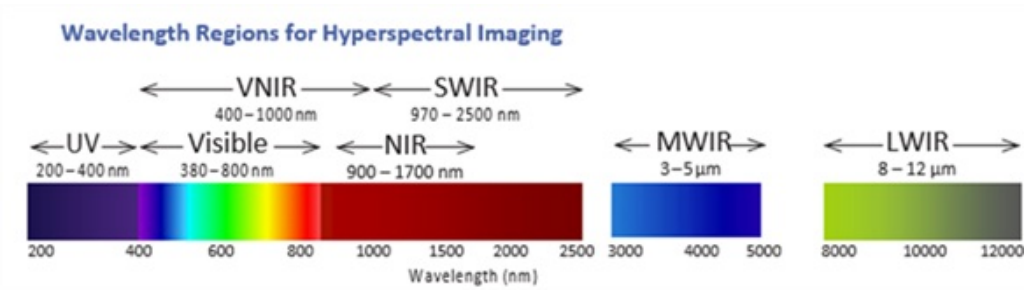


Figure 2.3: Segments of the electromagnetic spectrum commonly used in hyperspectral imaging (Middleton Spectral Vision, n.d.).

2.5.3 Advantages of Hyperspectral Imaging over RGB Imaging

Hyperspectral imaging offers substantial advantages over conventional RGB imaging by capturing detailed spectral information across a wide range of wavelengths. While RGB imaging provides only coarse color-based information, HSI records a dense spectral signature for each pixel, enabling precise characterization of material composition and condition Imaging, 2024. This enhanced spectral resolution allows HSI to detect subtle biochemical, structural, and physicochemical variations that are often indistinguishable in RGB images L. Chen et al., 2025; Liang et al., 2025.

In agricultural and food-related applications, HSI has demonstrated clear superiority over RGB imaging in tasks such as disease detection, quality assessment, and safety monitoring, owing to its ability to integrate spatial and spectral information in a non-destructive manner L. Chen et al., 2025; Liang et al., 2025. This capability facilitates the identification of pre-visual physiological stress, such as moisture imbalances or cell

wall degradation, which often occur before physical symptoms like grape cracking become visible to the naked eye Abade et al., 2025; Valme et al., 2025.

2.5.4 VIS–NIR Sensitivity to Physiological and Structural Changes

The visible and near-infrared (VIS–NIR) spectral range exhibits high sensitivity to physiological and structural changes in plant tissues. Variations in VIS reflectance (400–700 nm) are primarily associated with biochemical components such as photosynthetic pigments (e.g., chlorophyll), whereas NIR reflectance (700–1000 nm) is strongly influenced by internal tissue structure and water-related properties T. Wang and Zhao, 2024; Y. Zhang et al., 2022.

Recent studies have demonstrated that VIS–NIR hyperspectral imaging enables the early detection of plant physiological stress prior to the appearance of visible symptoms, highlighting its sensitivity to subtle functional changes at the tissue level Liu et al., 2023; Panda and Mishra, 2025. In fruit systems, VIS–NIR reflectance has been shown to correlate with internal quality attributes, including structural integrity, sugar content, and water status **postharvest2024VISNIR**; Q. Sun et al., 2024. These findings establish VIS–NIR HSI as a powerful tool for capturing microstructural and physiological variations that are not discernible using conventional imaging approaches.

2.5.5 Applications in Fruit Quality Assessment and Defect Detection

Hyperspectral imaging has been widely investigated as a non-destructive technique for fruit quality assessment and defect detection. Recent studies emphasize that combining HSI with advanced machine learning and deep learning methods enables accurate evaluation of fruit quality attributes and improves food safety monitoring Yang et al., 2025; X. Zhang, Li, et al., 2024. The rich spectral signatures captured by HSI facilitate the identification of internal composition and surface characteristics that are not discernible using conventional imaging.

Experimental studies demonstrate the effectiveness of HSI for detecting both external and internal fruit defects, such as bruises, scars, and disease-related decay Akter et al., 2025; Min et al., 2023a. In addition, HSI combined with feature selection techniques has shown strong potential for predicting key quality indicators such as soluble solid content (SSC), firmness, and aroma-related attributes in a variety of fruits Kanwal et al., 2025; P. Tian et al., 2023; Zhu et al., 2025. Overall, these studies establish HSI as a robust framework for postharvest management and industrial sorting, providing significant potential for industrial food safety assurance.

2.6 Machine Learning for Hyperspectral Data Analysis

Traditional machine learning approaches for hyperspectral imaging

Traditional machine learning (ML) approaches have been widely adopted for hyperspectral imaging (HSI) analysis in agricultural and food-quality applications, primarily due to their effectiveness in small- to medium-sized datasets and their reliance on well-established spectral preprocessing pipelines. Comprehensive review studies indicate that classical ML methods continue to play a central role in precision agriculture, supporting a broad range of tasks including crop and disease detection, stress monitoring, yield assessment, and produce quality evaluation García-Vera et al., 2024; Ram et al., 2024.

In typical traditional HSI workflows, each sample is represented as a spectral vector extracted at the pixel level, from a defined region of interest, or through spatial or spectral aggregation. Supervised learning is then applied using conventional algorithms such as partial least squares regression or discrimination (PLS/PLSR), support vector machines and regression (SVM/SVR), and other statistical or kernel-based models. These models are commonly combined with dedicated spectral preprocessing techniques aimed at mitigating sensor noise and compensating for scattering and illumination variability. Widely used preprocessing methods include Savitzky–Golay smoothing, standard normal variate (SNV) normalization, multiplicative scatter correction (MSC), and derivative-based spectral transformations.

Numerous studies have demonstrated the effectiveness of such classical pipelines for non-destructive assessment of fruit quality attributes. For instance, hyperspectral data combined with PLSR and SVR models have been successfully used to predict physicochemical properties of apples during storage following Savitzky–Golay filtering and SNV normalization Khaled et al., 2023. Similarly, PLS-based models employing tailored preprocessing strategies have achieved high predictive accuracy for parameters such as soluble solids content, acidity, and phenolic compounds in apples, underscoring the strong dependence of classical ML performance on appropriate spectral normalization Hasanzadeh et al., 2022.

A major advantage of traditional ML approaches lies in their interpretability and their ability to identify informative wavelengths within the high-dimensional hyperspectral space. Consequently, feature extraction and wavelength selection techniques are frequently integrated into classical pipelines to reduce redundancy and computational complexity while preserving predictive performance. Competitive adaptive reweighted sampling (CARS), for example, has been employed to extract effective wavelength subsets and to construct reduced-band models suitable for fast and potentially real-time applications

Khaled et al., 2023. Comparable results have been reported in pear quality assessment studies, where wavelength selection combined with LS-SVM and optimized preprocessing substantially improved model robustness and prediction accuracy Z. Zhang et al., 2024. These findings are consistent with broader reviews emphasizing the central role of preprocessing and feature selection in traditional HSI-based learning frameworks Ram et al., 2024.

Despite their proven effectiveness, traditional ML approaches are often evaluated under relatively controlled experimental conditions, such as laboratory or postharvest environments with standardized acquisition protocols. As a result, many existing studies focus on detection or quality estimation tasks based on already manifested spectral differences rather than on truly pre-symptomatic inference. Review studies further highlight persistent challenges in transferring HSI models from laboratory settings to real-world field deployment, including illumination variability, background complexity, sensor limitations, and domain shift effects García-Vera et al., 2024; Ram et al., 2024. Recent preharvest investigations demonstrate that strong predictive performance can be achieved through careful band reduction and model design; however, they also reveal that robustness and generalization remain key obstacles for operational field applications Fass et al., 2025. Overall, prior work establishes traditional ML as a powerful and interpretable foundation for hyperspectral data analysis, while motivating methodological choices that explicitly address variability and scalability when targeting early and field-relevant prediction tasks.

2.6.1 Deep learning and non-linear models

Computer vision (CV) has become a cornerstone technology in modern agriculture, driven by the rapid digital transformation of the sector and the growing need for automated, scalable, and data-driven crop management solutions. As global food demand increases and available farmland diminishes, CV systems—particularly those based on deep learning—provide the computational foundation required for precise, efficient, and reliable agricultural operations Patrício and Rieder, 2018; H. Tian et al., 2020. These systems enable machines to acquire, interpret, and analyze visual information, supporting tasks ranging from seed quality assessment and soil analysis to plant health monitoring, weed detection, livestock management, and yield estimation Cao et al., 2025; Dhanya et al., 2022. Over the last decade, convolutional neural networks (CNNs) have emerged as the dominant architecture for agricultural CV applications due to their strong feature extraction capabilities, scalability, and adaptability to diverse imaging conditions Attri et al., 2023a; Benos et al., 2021. Recent advances such as generative adversarial networks (GANs), transfer learning, and vision transformers (ViT) further extend the potential of CV systems, enabling improved robustness under field conditions characterized by variable lighting, occlusion, and environmental noise Dhanya et al., 2022; Ghazal et al.,

2024. Despite their substantial potential, the adoption of deep learning models in agricultural computer vision is constrained by practical data-related limitations. As highlighted by recent review studies, most deep learning applications in agriculture rely on supervised learning paradigms and task-specific datasets, which are often limited in size and expensive to acquire and annotate Attri et al., 2023b. The lack of large-scale, standardized annotated datasets restricts the availability of agriculture-specific pre-trained models and hampers knowledge transfer across crops, environments, and sensing conditions. As a result, deep learning models frequently need to be trained from scratch or extensively fine-tuned, requiring significant computational resources, long development cycles, and coordinated research efforts.

2.6.2 Pixel-level versus image-level modeling paradigms

In hyperspectral image analysis, the choice of representation unit plays a central role in determining both the sensitivity of predictive models to localized changes and their robustness under real-world conditions. Hyperspectral data are inherently acquired at the pixel level, where each pixel is represented as a high-dimensional spectral vector spanning contiguous wavelength channels. Pixel-level modeling directly exploits this representation and enables fine-grained discrimination of material properties and subtle biochemical variations, which can be critical for detecting early stress signals or micro-scale defects Moeini and Wang, 2018.

Despite these advantages, pixel-level approaches face substantial practical limitations. Pixel-wise classification or prediction typically requires dense and accurate pixel-level annotations, which are costly and time-consuming to obtain in agricultural contexts. Moreover, pixel-based predictions are often sensitive to noise, mixed pixels, and background variability, particularly under non-uniform illumination and complex field conditions. To address these challenges, many modern hyperspectral methods incorporate spectral–spatial strategies that aggregate information across local neighborhoods using patch-based, region-based, or superpixel representations. By integrating spatial context with spectral information, such approaches reduce sensitivity to pixel-level inconsistencies and often achieve improved robustness and generalization Z. Chen et al., 2024; Moeini and Wang, 2018.

By contrast, image-level modeling treats the entire hyperspectral cube or large spatial regions as holistic inputs to machine learning models, thereby simplifying annotation requirements and enabling global context to inform predictions. While image-level approaches may be less sensitive to highly localized early anomalies, they benefit from reduced labeling effort and increased stability under variable acquisition conditions. Recent empirical evaluations further highlight that the incorporation of spatial context is a key factor in achieving reliable and generalizable performance in deep learning-based

hyperspectral classification Z. Chen et al., 2024. Overall, the trade-offs between pixel-level, spectral–spatial, and image-level modeling paradigms reflect fundamental balances between spatial granularity, label efficiency, and operational robustness in hyperspectral applications.

2.6.3 Challenges: class imbalance, background clutter, and generalization

Despite these technological advancements, the effectiveness of CV models in agriculture is fundamentally dependent on the quality of the data used to train them. High-quality datasets must accurately represent the natural variability found in agricultural environments—including crop phenology, environmental dynamics, sensor noise, and geographical differences—to ensure generalization and prevent model bias. Poorly curated datasets, imbalanced class distributions, low-resolution imagery, or inconsistent labeling practices can lead to significant reductions in model performance, especially under real-world conditions where visual ambiguity is common Patrício and Rieder, 2018; H. Tian et al., 2020. Several studies highlight that annotation quality is often a major limiting factor in agricultural CV, as labeling requires domain expertise and is sensitive to human subjectivity. For example, plant disease severity, fruit defects, weed species boundaries, and growth-stage differentiation may be interpreted differently across annotators, leading to inconsistencies that degrade model accuracy and reliability Benos et al., 2021; Dhanya et al., 2022. This challenge is amplified in pixel-level tasks such as segmentation, where precise delineation of plant organs or symptomatic tissue is crucial for downstream analytical accuracy.

While the modeling strategies reviewed above span a wide range of algorithms and data representations, their application in fruit analysis literature can be broadly divided into two conceptual objectives: (i) the detection of existing, observable defects or quality attributes, and (ii) the prediction of latent or future physiological disorders prior to visible manifestation. The following section reviews prior work in both domains, with particular emphasis on hyperspectral approaches for fruit defect detection and early-stage prediction.

2.7 Related Work on Fruit Defect Detection and Prediction

Previous studies applying hyperspectral imaging to fruit quality analysis have addressed both defect detection and defect prediction, although these objectives are often treated implicitly rather than explicitly. In the context of this review, detection refers to the

identification of defects that are already visually or structurally manifested, whereas prediction denotes the identification of pre-symptomatic or latent conditions that precede visible damage.

2.7.1 HSI-based detection of visible fruit defects

Hyperspectral imaging (HSI) is a well-established non-destructive technique for detecting visible fruit defects, including skin blemishes, decay, and mechanical injuries, particularly in postharvest inspection and grading applications. By combining high spectral resolution with spatial information, HSI enables effective discrimination between healthy and defective tissues based on characteristic reflectance differences associated with structural damage, biochemical changes, or pathogen activity.

Numerous studies demonstrate the effectiveness of classical hyperspectral analysis pipelines for visible defect detection. Wang et al. B. Wang et al., 2023 employed hyperspectral imaging together with spectral preprocessing and wavelength selection methods (CARS and SPA) to classify natural skin defects in *Cerasus humilis* fruit. Their approach achieved high classification accuracy and enabled spatial localization of defective regions, highlighting the importance of integrating spectral discrimination with image-level analysis. These findings are consistent with broader reviews showing that Vis/NIR spectroscopy and hyperspectral techniques are robust tools for identifying surface degradation across diverse fruit types under controlled conditions C. Wang et al., 2025.

Beyond surface-level defects, hyperspectral imaging has also been applied to the detection of decay processes during postharvest storage. Min et al. Min et al., 2023b reviewed recent advances in early decay detection using hyperspectral imaging, demonstrating that HSI can capture spectral changes associated with fungal infection before severe tissue degradation becomes visually apparent. Nevertheless, most reported applications focus on monitoring ongoing decay during storage rather than on predicting latent physiological disorders prior to visible symptom development.

Comprehensive reviews further confirm the maturity of hyperspectral imaging as a quality assessment tool within the fruit and vegetable supply chain. Vignati et al. Vignati et al., 2023 summarized key concepts, data processing strategies, and applications of HSI for quality monitoring, emphasizing its effectiveness for defect detection and postharvest inspection, while also noting limitations related to environmental variability and model transferability.

Overall, existing literature establishes hyperspectral imaging as a reliable and powerful technology for detecting visible fruit defects and postharvest decay under controlled acquisition conditions. However, the majority of these studies address defects that are already externally or structurally expressed, thereby motivating further research toward predic-

tive, pre-symptomatic frameworks capable of identifying latent physiological changes prior to visible manifestation, particularly under realistic field conditions.

2.7.2 Temporal and Longitudinal Spectral Analysis

Temporal and longitudinal spectral analysis extends conventional hyperspectral approaches by explicitly modeling changes in spectral signatures over time, moving beyond static, single-time-point observations. In fruit quality assessment, such approaches aim to capture progressive physiological, biochemical, or structural shifts associated with ripening, environmental stress, or decay processes. By analyzing spectral trajectories across multiple acquisition points, hyperspectral imaging (HSI) enables the monitoring of dynamic quality evolution that may precede visible defect formation.

Several studies have demonstrated the utility of temporal spectral analysis in postharvest monitoring scenarios. Min et al., 2023b showed that HSI can detect early fruit decay by capturing subtle spectral variations associated with fungal infection prior to pronounced tissue degradation. These variations evolve systematically over time, providing reliable indicators of quality deterioration during storage. Similarly, C. Wang et al., 2025 emphasized that Vis/NIR-based systems are capable of tracking time-dependent changes in physicochemical properties, supporting applications such as shelf-life estimation and decay monitoring.

Despite these advances, most longitudinal hyperspectral studies remain confined to controlled postharvest environments, where acquisition conditions and sampling intervals are strictly regulated. Reviews by Vignati et al., 2023 highlight that temporal spectral analysis is predominantly used in a descriptive or confirmatory manner, focusing on monitoring ongoing quality degradation rather than enabling predictive inference. Consequently, temporal modeling is typically applied to identify or validate existing deterioration, rather than to infer latent physiological disorders at pre-symptomatic stages.

Overall, while temporal and longitudinal spectral analysis provides a powerful framework for characterizing dynamic fruit systems, its application remains largely limited to postharvest scenarios. This underscores a clear need for methodologies that leverage temporal spectral information to predict early-stage physiological disorders under realistic preharvest field conditions.

2.7.3 Attempts Toward Early and Pre-Symptomatic Defect Prediction

Several investigations have demonstrated that HSI can reveal latent internal disorders before external manifestation. Dong et al. Dong et al., 2025 showed that early-stage bitter pit in apples can be identified using VIS–NIR hyperspectral data combined with

spectral and texture features, achieving high classification accuracy prior to visible pit formation. This highlights the capacity of HSI to detect internal physiological imbalances before macroscopic damage occurs.

Pre-symptomatic detection of pathogen infection has also been reported across multiple crops. Haghbin et al. Haghbin et al., 2023 demonstrated that gray mold infection in kiwifruit can be identified using hyperspectral data and chemometric models, capturing infection-related spectral changes before visible decay appears. Similarly, hyperspectral imaging enabled early detection of anthracnose in pear leaves through multi-source feature fusion Y. Zhang et al., 2024, as well as the identification of bacterial leaf spot in tomato prior to symptom development X. Zhang, Vinatzer, and Li, 2024.

Recent advancements have further integrated HSI with deep learning to enhance sensitivity to subtle stress-related patterns. Ou et al. Ou et al., 2024 combined HSI with deep neural networks for the early detection of gray mold in strawberry leaves, identifying spectral signatures associated with initial pathogen stress.

Despite these results, most early-detection studies remain limited to controlled environments and localized leaf- or fruit-level analysis. Truly predictive frameworks that anticipate future defect emergence under realistic, field-scale vineyard conditions remain scarce. This underscores the need for methodologies that leverage longitudinal hyperspectral data for robust, pre-symptomatic prediction in real-world agricultural settings.

2.7.4 Limitations of Existing Approaches under Field Conditions

Despite the substantial progress demonstrated by hyperspectral imaging (HSI) in fruit quality assessment and early defect detection, the majority of existing approaches remain challenged by limited applicability under real-world field conditions. Most studies are conducted in controlled laboratory or postharvest environments, where illumination, background, viewing geometry, and acquisition timing are carefully regulated. While such settings facilitate methodological development and performance benchmarking, they do not adequately reflect the complexity and variability encountered in operational orchard or vineyard scenarios.

Several reviews emphasize that environmental variability represents a major obstacle for field deployment of HSI systems. Vignati et al. Vignati et al., 2023 note that changes in ambient lighting, surface moisture, temperature, and background composition can significantly distort spectral signatures, reducing model robustness and transferability. Similarly, Wang et al. C. Wang et al., 2025 highlight that Vis/NIR-based fruit quality models often exhibit degraded performance when transferred across seasons, cultivars, or acquisition setups, necessitating frequent recalibration.

Another key limitation lies in data annotation and scalability. Many HSI-based studies rely on dense pixel-level or region-level labeling, which is labor-intensive, subjective, and difficult to obtain consistently under field conditions. This constraint is particularly pronounced in early-detection studies, where pre-symptomatic changes are subtle and often lack clear visual reference points for annotation. As a result, most early-detection experiments remain restricted to localized fruit- or leaf-level samples acquired under controlled conditions Haghbin et al., 2023; Y. Zhang et al., 2024.

Temporal variability further complicates field deployment. Although longitudinal spectral analysis has shown promise for monitoring decay progression during storage, temporal sampling under field conditions is far less structured. Min et al. Min et al., 2023b demonstrated that early decay signatures evolve systematically in postharvest settings; however, extending such temporal modeling to preharvest environments introduces additional sources of uncertainty related to plant phenology, weather dynamics, and management practices.

Finally, while recent studies have explored the integration of deep learning with HSI to enhance sensitivity to subtle stress patterns, these approaches remain data-intensive and highly task-specific. Deep learning models typically require large, consistently labeled datasets that are rarely available in agricultural field contexts, limiting their generalizability and operational feasibility Ou et al., 2024. Consequently, most existing frameworks prioritize detection of already manifested defects or localized early symptoms, rather than robust, pre-symptomatic prediction at field scale.

Collectively, these limitations indicate that although HSI is a powerful sensing modality, current methodologies fall short of enabling reliable early prediction of fruit disorders under realistic field conditions. Addressing environmental variability, annotation constraints, temporal inconsistency, and model transferability remains essential for advancing HSI-based systems from experimental studies toward practical, field-deployable early warning tools.

2.8 Research Gap and Study Positioning

2.8.1 Summary of Established Knowledge

Hyperspectral imaging (HSI) has been extensively investigated as a non-destructive tool for agricultural monitoring, particularly for the detection of visible defects, postharvest decay, and quality-related attributes in fruits and vegetables. Numerous studies have demonstrated that spectral and spectral–spatial features can effectively discriminate between healthy and defective tissues, quantify physicochemical properties, and support automated grading and sorting processes. In addition, temporal hyperspectral analy-

sis has been successfully applied to monitor quality degradation and disease progression during postharvest storage under controlled conditions.

In parallel, the physiological and biomechanical mechanisms underlying fruit cracking have been widely studied using destructive measurements, biochemical assays, and controlled experiments. These studies have established the roles of cuticle integrity, cell wall composition, water relations, and hormonal regulation in crack development, providing valuable biological insight into cracking susceptibility.

Despite this extensive body of work, the integration of hyperspectral imaging with physiological knowledge of cracking remains limited. In particular, while HSI has proven effective for detecting existing defects, its use for predicting the future occurrence of a physiological disorder such as grape berry cracking—based on early spectral trends and temporal dynamics—remains largely unexplored.

2.8.2 Identified Methodological and Application Gaps

A central gap in the existing literature lies in the focus on reactive rather than predictive frameworks. Most hyperspectral studies in fruit analysis are designed to identify defects or diseases after they have already manifested visually or structurally. Even studies described as "early detection" typically operate at a stage where symptoms are already incipient or localized, rather than truly pre-symptomatic.

Methodologically, many existing approaches rely on high-dimensional spectral data without explicitly addressing spectral redundancy or computational efficiency. While feature selection and wavelength reduction techniques have been applied in quality assessment tasks, they are rarely integrated with temporal analysis to support early prediction. Moreover, most models are trained and evaluated under controlled laboratory or postharvest conditions, limiting their robustness and transferability to preharvest field environments characterized by variable illumination, background complexity, and plant heterogeneity.

From an application perspective, there is a notable lack of non-destructive tools capable of providing actionable early warnings for physiological disorders in fruit crops. In the context of table grapes, cracking is typically identified only after visible damage occurs, at which point economic loss is often unavoidable. Existing monitoring strategies therefore do not support timely intervention or informed decision-making at the vineyard level.

2.8.3 Lack of Early, Non-Destructive Crack Prediction Frameworks

Although grape cracking has been extensively studied from a physiological and mechanistic standpoint, few studies have attempted to detect or predict cracking at an early stage using optical sensing techniques. In particular, the detection of pre-symptomatic

microcracks or latent physiological stress—prior to visible skin rupture—remains largely unaddressed in the literature.

This gap reflects both technical and conceptual challenges. Pre-symptomatic cracking is characterized by subtle structural and biochemical changes that do not produce obvious visual cues, making traditional imaging approaches insufficient. Furthermore, the lack of dense, time-resolved hyperspectral datasets acquired under field conditions has limited the development of predictive models capable of learning early spectral signatures associated with future cracking events.

As a result, no established framework currently exists that combines non-destructive hyperspectral imaging, temporal spectral analysis, and machine learning to predict grape cracking before it becomes visually apparent.

2.8.4 Positioning of the Present Study within the Literature

The present study is positioned at the intersection of physiological research on grape cracking and applied hyperspectral data analysis. By leveraging VIS–NIR hyperspectral imaging, machine learning, and spectral feature selection, this work aims to move beyond defect detection toward early prediction of cracking susceptibility.

Specifically, the proposed approach integrates temporal spectral information with efficient wavelength selection to capture subtle changes in berry skin integrity and physiological state prior to visible cracking. Unlike prior studies that focus on postharvest or single-time-point analysis, this work emphasizes early-stage, preharvest monitoring under realistic conditions.

In doing so, this study bridges a critical gap between fundamental understanding of cracking mechanisms and practical, non-destructive vineyard monitoring. The resulting framework contributes a novel perspective to the literature by explicitly addressing the distinction between detection and prediction and by demonstrating the feasibility of hyperspectral imaging as an early warning tool for grape cracking.

Chapter 3

Research Objectives and Hypotheses

3.1 Research Objectives

The primary objective of this research is to develop and validate hyperspectral imaging (HSI) methods capable of detecting and predicting cracking in table grapes ('Scarlotta' cultivar) at both symptomatic and pre-symptomatic stages. This work aims to bridge the gap between spectroscopic research on cracking mechanisms and practical, non-destructive agricultural monitoring by leveraging advances in imaging, machine learning, and spectral analysis.

3.2 Research Hypotheses

The research is guided by the following hypotheses, structured hierarchically to reflect a progression from physical feasibility to algorithmic performance, robustness, early detection capability, and spectral efficiency.

Hypothesis 1: Spectral Distinguishability (Physical Basis)

This hypothesis examines whether a fundamental physical difference exists in the spectral reflectance properties of healthy and cracked grape tissues.

- H_0**:** There is no statistically significant difference in the spectral signatures of healthy and cracked tissues in 'Scarlotta' grapes.
- H_1**:** Statistically significant differences exist in the spectral signatures of healthy versus cracked grape skins at specific wavelengths, providing a physical basis for subsequent algorithmic discrimination.

Hypothesis 2: Pixel-Level Classification Feasibility

This hypothesis evaluates whether machine learning algorithms can effectively learn the spectral differences identified in Hypothesis 1 under controlled, pixel-level conditions.

- H_0**:** A machine learning model trained on hyperspectral data cannot achieve classification performance, as measured by standard metrics (e.g., F1-score or accuracy), significantly higher than a random or baseline classifier at the pixel level.
- H_1**:** The proposed machine learning framework can successfully classify individual pixels as *Cracked* or *Healthy* with high performance, effectively learning the underlying spectral decision boundary.

Hypothesis 3: Whole-Image Detection and Robustness

This hypothesis assesses the robustness of the trained models when applied to realistic whole-image hyperspectral scenes containing complex background elements not present in the clean pixel dataset.

- H_0**:** The model fails to distinguish between the spectral signatures of cracks and complex background clutter (e.g., leaves, woody tissue, or plastic), resulting in a high false positive rate when applied to full hyperspectral images.
- H_1**:** The spectral signatures of cracks are sufficiently distinct from background clutter to allow the model to effectively suppress noise and detect cracks within full-image scenes with high precision.

Hypothesis 4: Early Detection Capability

This hypothesis investigates the ability of the whole-image framework to detect cracking at its onset, prior to the appearance of visually observable symptoms.

- H_0**:** The spectral changes associated with the early onset of cracking are too subtle to be reliably detected, limiting the model's applicability to late-stage damage.
- H_1**:** The spectral model is sufficiently sensitive to detect early physiological changes and micro-cracks, enabling pre-symptomatic identification of clusters at risk of cracking.

Hypothesis 5: Spectral Efficiency and Operational Feasibility

This hypothesis examines whether the detection capabilities demonstrated in Hypotheses 3 and 4 can be preserved when using a reduced subset of informative spectral bands.

- H_0**:** Reducing the number of spectral bands results in a significant degradation in detection performance, either in whole-image scenarios or during early-stage detection, implying that the full hyperspectral spectrum is required.
- H_1**:** A carefully selected subset of informative wavelengths captures the key physiological markers of cracking, enabling robust detection in both whole-image and early-stage scenarios using cost-effective multispectral sensing configurations.

This thesis is organized into several interconnected chapters. The *Literature Review* chapter provides a comprehensive overview of grape cracking mechanisms and relevant precision agriculture techniques. The *Materials and Methods* chapter describes the experimental design, data acquisition procedures, hyperspectral imaging systems, and modeling approaches used in this study. The *Results* chapter presents the experimental findings, including pixel-level classification, whole-image crack detection, temporal spectral analysis, and wavelength selection experiments. The *Discussion* chapter interprets these results in the context of early crack detection, physiological relevance, and practical deployment considerations. Finally, the *Conclusions and Future Work* chapter summarizes the main contributions of the study and outlines directions for future research.

Chapter 4

Materials and Methods

4.1 Data Acquisition

4.1.1 Study Area and Experimental Design

The study was conducted in the Lachish region of Israel, an area characterized by a semi-arid climate with a hot, dry summer and a mild winter. These conditions make the region a central hub for grape cultivation (viticulture) in Israel. The experiment was conducted in a commercial table grape vineyard (*Vitis vinifera* L., cv. 'Scarlotta'), in collaboration with the "Tali Grapes" brand. The vineyard was grafted onto 1103 Paulsen rootstock and was planted in 2018.

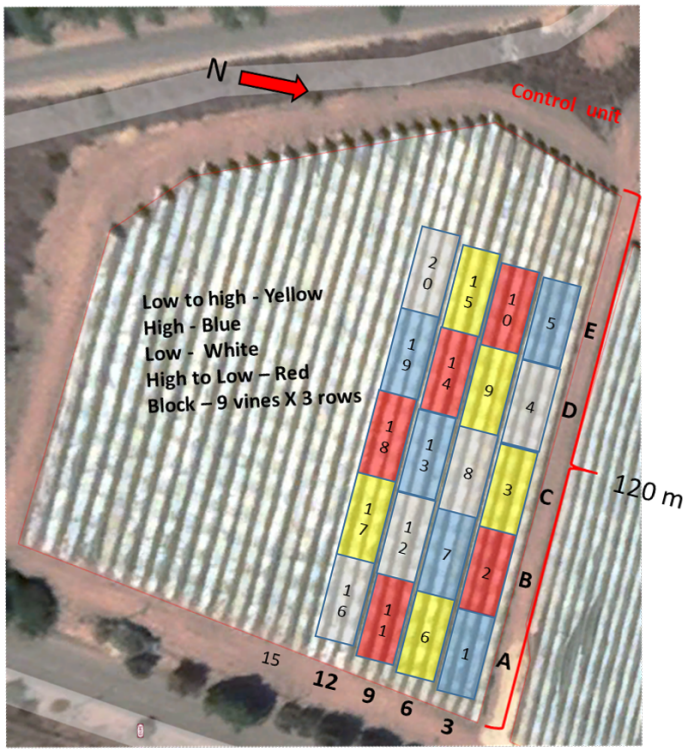


Figure 4.1: Map of the experimental vineyard showing the five randomized blocks (A–E) and the four irrigation regimes applied within each block.

The experimental design followed a Randomized Complete Block Design (RCBD), including five replicates (Blocks A–E). Within each block, four different irrigation regimes were tested, resulting in a total of 20 experimental plots (5 blocks × 4 treatments).

Each experimental plot consisted of three consecutive rows of nine vines each (27 vines per plot). To minimize edge effects and water spillover between irrigation regimes, only vines from the middle row of each plot were selected for measurement. From each middle row, 12 representative grape clusters were tagged, yielding a total sample size of 240 clusters (20 plots × 12 clusters).

The four irrigation regimes tested were:

- **High Irrigation (Blue):** 796 mm of water throughout the season.
- **Low Irrigation (White):** 376 mm of water throughout the season.
- **Reducing Irrigation (Red):** A transition from high to low irrigation (total 532 mm), beginning at the post-veraison stage.
- **Increasing Irrigation (Yellow):** A transition from low to high irrigation (total 645 mm), beginning at the post-veraison stage.

Irrigation commenced on April 29, 2024, and was applied daily or every other day via a drip irrigation system.

The experiment continued until the end of September 2024. Physiological and yield measurements—including stem water potential, stomatal conductance, fruit properties, and percentage of healthy clusters—were performed weekly. Physiological data were collected using a Scholander Pressure Bomb and a Porometer, and soil moisture was measured using Soil Moisture Sensors.

4.1.2 Imaging Systems

To study grape cracking in a comprehensive way, a multi-modal imaging setup was used. The system combined three imaging sensors placed at the same location, each capturing different types of information from a different part of the electromagnetic spectrum. Together, these sensors provided complementary spectral, spatial, and thermal data, allowing a more complete analysis of grape berry condition and cracking development. All images were collected from a fixed distance of approximately 1 m from the grape clusters to ensure consistent scale and alignment across all measurements.

Hyperspectral Imaging System (HSI)

Hyperspectral data were collected using the **Specim IQ** (Specim, Spectral Imaging Ltd., Oulu, Finland), a portable snapshot hyperspectral camera designed for field applications. The camera operates in the Visible and Near-Infrared (VNIR) spectral range (400–1000 nm) and produces 3D datacubes (spatial × spatial × spectral) for each snapshot.

(a) Field deployment of the Specim IQ system in the vineyard.

(b) Hyperspectral image acquisition of grape clusters using the Specim IQ.

Figure 4.2: Hyperspectral imaging setup and data acquisition using the Specim IQ system under vineyard conditions.

- **Spectral Resolution:** The system captures 204 contiguous spectral bands with a Full Width at Half Maximum (FWHM) of 7 nm and a spectral sampling rate of

approximately 3 nm. The high spectral fidelity is ensured by a signal-to-noise ratio (SNR) of $> 400:1$.

- **Spatial Resolution:** The sensor has a resolution of 512×512 pixels. With a $31° \times 31°$ Field of View (FOV), at a 1 m working distance, the camera covers a physical area of approximately 0.55 m $\times$ 0.55 m. This configuration yields a theoretical spatial resolution of approximately **1.08 mm/pixel**, sufficient for resolving individual berries and larger cracks.
- **Radiometric Calibration:** A spectrally flat, diffuse white reference panel (Spectralon®, $\approx 99\%$ reflectance) was included in every captured scene. This allowed subsequent normalization of raw radiance data into relative reflectance values, accounting for varying natural illumination conditions in the vineyard.

RGB Imaging System

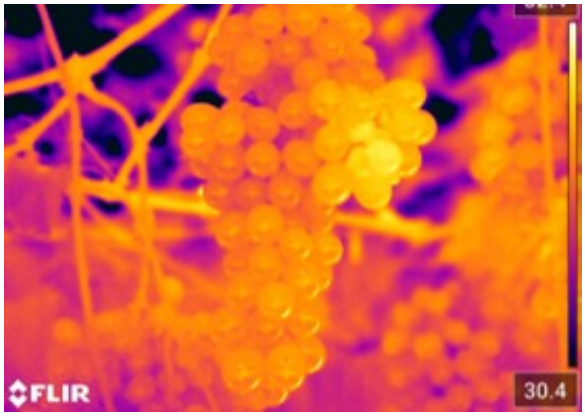
To provide high-resolution visual context and capture fine surface textures that might be missed by the lower spatial resolution of the HSI system, a high-end digital camera was used.

- **Camera & Sensor: Canon EOS R5**, equipped with a 45-megapixel full-frame (36×24 mm) CMOS sensor, producing 14-bit RAW images with a resolution of 8192×4608 pixels.
- **Optics:** The camera was fitted with a **35 mm** prime lens, providing a wide Field of View of 54.4° (horizontal) $\times$ 37.8° (vertical) and sufficient depth of field to keep the entire cluster in focus.
- **Spatial Resolution:** At the 1 m working distance, this system achieved an extremely high theoretical spatial resolution of approximately **0.13 mm/pixel**. This level of detail is critical for detecting micro-cracks (hairline fissures) and minute alterations in the berry skin cuticle before they become visible to the naked eye or the HSI sensor.

Thermal Imaging System

Canopy and grape cluster temperatures, which are widely used as indirect indicators of plant water status, stomatal conductance, and transpirational cooling dynamics, were monitored using a high-definition thermal imaging system.

(a) FLIR T1020 thermal camera used for vineyard measurements.



(b) Thermal image of a grape cluster showing surface temperature variations.

Figure 4.3: Thermal imaging system and example thermal data acquired from grape clusters under vineyard conditions.

- **Camera: FLIR T1020** (Teledyne FLIR LLC, Wilsonville, OR, USA), a high-performance uncooled microbolometer specifically designed for precision research.
- **Sensor & Optics:** The camera has a thermal resolution of 1024×768 pixels with a 17 μm pixel pitch, operating in the Long-Wave Infrared (LWIR) spectral range (7.5–14 μm). It was equipped with a standard **34 mm** lens (adjustable focus including autofocus), providing a FOV of 45° (H) × 34° (V).
- **Performance:** The system boasts a high thermal sensitivity (Noise Equivalent Temperature Difference, NETD) of < 20 mK (< 0.02°C) at 30°C, enabling the detection of subtle temperature gradients across the grape cluster.
- **Spatial Resolution:** At the 1 m working distance, the theoretical thermal spatial resolution was approximately **0.81 mm/pixel**, allowing for accurate temperature extraction from individual berries while minimizing mixed-pixel effects from the surrounding air or leaves.

Due to the high dimensionality of the acquired data, individual hyperspectral images were substantially larger than the corresponding RGB and thermal images. On average, a single hyperspectral datacube occupied approximately **200 MB**, compared to approximately **11 MB** for a high-resolution RGB image and **2.5 MB** for a thermal image. In total, the imaging campaign resulted in a dataset of approximately **575 GB** of raw data, highlighting the computational and storage challenges associated with multi-modal hyperspectral analysis.

4.1.3 Data Collection Protocol

Data collection was conducted during concentrated field days, typically once a week (Mondays or Thursdays), throughout the growing season from June to the end of September. To ensure consistent lighting conditions, field operations were strictly scheduled between 08:00 and 17:00.

Daily Schedule and Logistics

Each collection day began at 05:30 at the Volcani Institute with equipment organization and verification. Departure to the vineyard in Lachish occurred at approximately 06:00. Upon arrival at 07:00–08:00, the imaging systems were mounted on their respective tripods and calibrated.

Sampling Strategy

Due to significant differences in acquisition time between modalities, a split sampling strategy was employed:

- **RGB Imaging:** All 240 marked grape clusters (12 clusters $\times$ 20 plots) were imaged first during each field day to capture them under relatively uniform early-morning lighting.
- **Hyperspectral Imaging (HSI):** Due to the complex acquisition process—which required re-positioning the camera and white reference panel for each cluster, followed by a dedicated spectral focusing procedure—each HSI scan took approximately 2-3 minutes. Consequently, only a subset of 120 clusters was imaged per session. These clusters were selected to represent a balanced cross-section of the experimental blocks, ensuring full dataset coverage over alternating sessions. HSI acquisition typically commenced around 10:00 AM when sunlight was more stable and intense.
- **Thermal Imaging:** Thermal image acquisition was purposely scheduled for later in the day (starting around 12:00 PM), after RGB imaging was complete. This timing leverages the high ambient temperatures typical of midday in the semi-arid Negev region to maximize thermal contrast on the berry surface. Specifically, cracked or decaying areas, often characterized by exposed moist tissue or conversely by localized dehydration, exhibit different thermal signatures compared to healthy, intact skin due to variations in evaporative cooling rates. Capturing images during peak heat enhances these temperature differentials, improving the detectability of such defects.

Field Procedure

To ensure accurate re-imaging, a consistent labeling system was implemented. Target clusters were marked with laminated, color-coded tags attached via plastic cable ties. The tags featured a numeric identifier (1–60), which was repeated across treatments but distinguished by the specific color assigned to each irrigation regime. This design allowed for unambiguous identification and durability against humidity. Furthermore, vine locations were georeferenced using a GPS module integrated into the thermal camera.

Although only the RGB and HSI datasets were analyzed in the present study, all imaging systems were spatially co-aligned and temporally synchronized during field acquisition to facilitate potential multi-modal data fusion in future analyses.

4.2 Preprocessing and Dataset Preparation

Table 4.1: Overview of datasets used for crack detection and prediction experiments.

Dataset	**Level**	**Labels / Classes**	**Purpose**	**RGB / SAM2 Role**	**Hyp.**
Pixel-Level Binary	Pixel	Healthy Berry; Cracked Berry	Learn physical spectral differences	SAM2-assisted segmentation and visual context	H1–H2
Pixel-Level Multi-Class	Pixel	Cracked; Healthy; Leaf; Wood; Soil; Background	Learn spectral context	SAM2-assisted segmentation and class disambiguation	H3
Whole-Image (Late)	Image	Symptomatic (≥ 1 visible crack)	Evaluate robustness under realistic vineyard conditions	RGB used for visual alignment during inference	H3
Whole-Image (Early)	Image	Asymptomatic (no visible crack at acquisition)	Evaluate detection at early, non-visible stages	RGB used for temporal reference and validation	H4
Temporal Spectral	Cluster	Mean spectral signature (no explicit labels)	Analyze temporal trends and identify wavelengths	SAM2 used for cluster localization and ROI extraction	H5

Note: All datasets use paired HSI+RGB acquisitions. RGB imagery supports SAM2-assisted annotation and validation.

4.2.1 Data Management and Pre-processing

Given the large volume and multi-modal nature of the data collected throughout the season, a robust and automated data management pipeline was essential. Custom Python scripts were developed to orchestrate the organization, validation, and preprocessing of the raw imagery.
The pipeline operated in several stages:

- **Automated Organization:** Scripts parsed acquisition timestamps and file metadata to automatically sort thousands of raw images into a structured directory hierarchy, organized by acquisition date, sensor type (RGB, Thermal, HSI), and unique cluster ID.

- **Metadata Verification (HSI):** For the hyperspectral data, a dedicated verification script scanned the XML metadata files associated with each datacube. This ensured that every acquired image was correctly tagged with its corresponding cluster ID in the field, and generated detailed CSV reports mapping successful acquisitions versus missing data points across all dates.

- **Quality Assurance:** Following automated sorting, a rigorous manual inspection was conducted. Images were reviewed to confirm proper focus, appropriate lighting, and for the HSI data, the presence of a valid white reference panel within the frame for subsequent radiometric calibration. Only complete, verified, and correctly labeled image sets were advanced to the downstream analysis and labeling stages.

4.2.2 Pixel-Level Dataset Preparation

To enable efficient creation of the pixel-level dataset—required to test spectral differences between healthy and cracked grape tissues (Hypothesis 1) and to assess pixel-level classification feasibility using machine learning (Hypothesis 2)—a custom graphical user interface (GUI) was developed in Python. The tool integrates data visualization, interactive annotation, and model inference in a single platform (see Figure 4.4). The dataset generation workflow proceeded as follows:

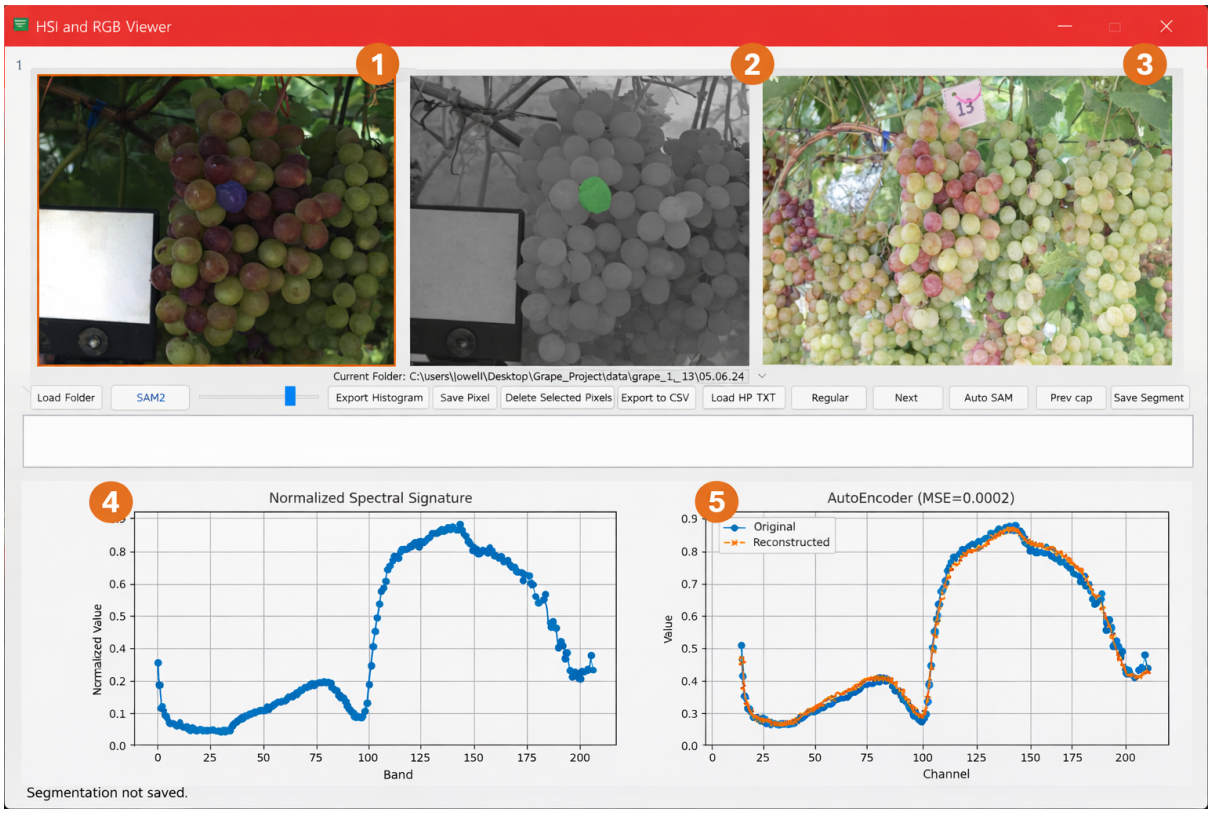


Figure 4.4: Graphical user interface of the HSI – RGB annotation tool. (1) Hyperspectral image visualized as an RGB composite with a SAM2-generated segmentation overlay, (2) wavelength-selectable grayscale HSI band with the same SAM2 segmentation mask, and (3) RGB image of the same grape cluster captured on the same day using a Canon EOS R5. (4) Spectral signature extracted from a single pixel within the SAM2-segmented region, and (5) autoencoder reconstruction of the same spectral signature, discussed later in this work.

1. **Image Loading and Visualization:** The tool loaded and displayed a grape cluster's hyperspectral (HSI) data cube alongside its corresponding high-resolution RGB context images (both from the HSI camera and a separate Canon camera). A slider mechanism allowed for interactive exploration of the HSI cube by cycling through its 204 spectral bands (400–1000 nm).

2. **Interactive Pixel Sampling:** The user manually selected a single "seed pixel" by clicking on a region of interest—either a clearly cracked berry or a clearly healthy berry—within any of the displayed images.

3. **Automated Segmentation with SAM2:** The selected seed pixel was passed as a point prompt to the **Segment Anything Model 2 (SAM2)** Ravi et al., 2024, which was integrated directly into the tool via its Python API. SAM2 then performed instance segmentation, generating a precise binary mask of the entire berry corresponding to the selected point. This semi-automated approach significantly accelerated the tedious process of manual segmentation.

4. **Data Extraction and Labeling:** Subsequent to segmentation, the tool extracted the full spectral signatures for all pixels encompassed by the binary mask. The user then manually assigned a class label—'Cracked' or 'Regular'—to the entire berry instance. Upon verification, the annotated spectral data were stored for inclusion in the training dataset.

This process was repeated for 50 cracked and 50 healthy berries, resulting in a dataset of approximately 125,000 labeled pixels (∼45,000 cracked, ∼80,000 regular), forming the basis for the pixel-level classification models.

4.2.3 Expansion to Multi-Class Dataset

The initial binary pixel-level classification between healthy and cracked grape tissues was not sufficient for direct application to full-image analysis, due to the presence of complex background elements in vineyard scenes. To enable a controlled transition toward whole-image semantic segmentation, the dataset was expanded to include additional background-related classes. This step allows the model to learn spectral context, reduces background-induced misclassification, and supports robust image-level inference required for addressing Hypothesis 3.

Using the same annotation workflow described above, spectral signatures were extracted from representative background elements commonly present in the vineyard environment. The resulting multi-class dataset included the following categories:

- **Cracked Berry:** Grape tissue exhibiting visible cracking symptoms.

- **Healthy Berry:** Intact grape tissue without visible defects.
- **Woody Tissue:** Vine branches, stems, twigs, and other lignified structures.
- **Leaf:** Vine foliage, including both sunlit and shaded leaves.
- **Soil:** Exposed ground surfaces within the vineyard, including soil and small debris.
- **Artificial Background:** Non-biological vineyard elements such as plastic clips, irrigation lines, and tags.

This expansion resulted in a diverse spectral library aimed at training models capable of distinguishing between actual fruit damage and environmental noise.

4.2.4 Whole-Image Dataset Construction and Experimental Design

To evaluate crack detection under realistic field conditions, a whole-image hyperspectral dataset was constructed. Unlike the pixel-level datasets, which rely on local spectral samples, this dataset consists of complete hyperspectral images of grape clusters acquired in the vineyard. This setup preserves real-world factors such as spatial structure, background clutter, and natural illumination variability, which are essential for assessing system-level robustness.

Dataset Composition

The whole-image dataset consists of a total of 70 hyperspectral images representing two distinct physiological states:

- **Symptomatic (Cracked):** 10 images containing grape clusters with macroscopically visible cracking.
- **Asymptomatic (Healthy / Pre-symptomatic):** 60 images containing grape clusters with no visible cracking, acquired during an earlier phenological stage prior to the documented onset of cracking symptoms. Specifically, asymptomatic images were restricted to acquisition dates no later than July 25, 2024, corresponding to five imaging sessions conducted on June 24, July 1, July 10, July 15, and July 25, 2024.

Spatial Partitioning and Data Splits

To prevent information leakage caused by spatial correlation and to ensure biological independence between training and evaluation samples, a spatially stratified partitioning strategy was employed:

- **Training Set:** Images acquired exclusively from **Row 1**, including both cracked and non-cracked clusters.
- **Validation Set:** Images acquired exclusively from **Row 2**, including both cracked and non-cracked clusters.

This row-based separation enforces a conservative evaluation protocol, ensuring that models are tested on grape clusters that are spatially distinct from those used during training.

Temporal Stratification for Early Detection Analysis

To explicitly evaluate sensitivity to disease progression and early crack formation, the dataset was further stratified temporally based on vineyard monitoring records. Two evaluation subsets were defined:

1. **Late-Stage (Symptomatic):** Images acquired after visible cracking symptoms were formally documented for the corresponding grape clusters.
2. **Early-Stage (Pre-Symptomatic):** Images acquired approximately one week prior to the first documented visual annotation of cracking for the same clusters.

Importantly, **Early-Stage** samples were visually annotated as non-cracked at the time of acquisition. Consequently, successful detection in this subset corresponds to identifying latent or pre-symptomatic spectral signatures before cracks become observable to the human eye.

This temporal stratification was applied exclusively during the evaluation phase and did not influence model training, parameter tuning, or aggregation logic. As a result, performance differences between early- and late-stage subsets directly reflect the inherent difficulty of early crack detection rather than methodological variation.

4.2.5 Automated Signature Extraction for Temporal Analysis

To investigate the prediction objective and analyze temporal changes in spectral characteristics across different irrigation regimes, a separate, fully automated data extraction pipeline was developed. This approach was designed to circumvent the labor-intensive process of manual segmentation and to capture a holistic spectral representation of the entire grape cluster over time. The methodology focused on analyzing the central region of each hyperspectral image, where the main body of the grape cluster was consistently located.

The process was implemented via a Python script that systematically processed all collected images. The core logic of the script is outlined below:

Methodological steps

1. **Batch Image Loading:** The script iterates through the main data directory, identifying all subfolders corresponding to different clusters and capture dates. For each entry, it utilizes a helper function, similar to the `load_hsi_image_from_folder` function found in the project's backend, to load the hyperspectral (HSI) data cube into memory.

2. **Defining a Central Region of Interest (ROI):** For each HSI data cube (with dimensions $512 \times 512 \times 204$), a square Region of Interest (ROI) was defined at the geometric center of the image. A typical ROI size was set to 50×50 pixels. This size was chosen as a conservative estimate to ensure that the extracted pixels primarily belonged to the grape cluster, minimizing the inclusion of background elements like leaves or the sky, regardless of the cluster's specific position in the frame.

3. **Mean Signature Calculation:** All pixel vectors within the defined central ROI were extracted. The script then computed the element-wise mean of these vectors, resulting in a single, average spectral signature of 204 bands for that specific cluster at that point in time. This averaging process effectively filters out spatial noise and provides a robust representation of the cluster's overall spectral state.

4. **Data Aggregation and Grouping:** The calculated mean signature was stored along with its corresponding metadata, including the cluster ID, capture date, and the specific irrigation treatment group it belonged to. This process was repeated for every image collected throughout the season.

5. **Final Analysis:** After processing all images, the aggregated data was grouped by irrigation treatment and capture date. For each group at each time point, the final mean spectral signature and its standard deviation were calculated. This allowed for the direct comparison of temporal trends between the different irrigation regimes, as presented in Chapter 4.

This automated approach enabled the efficient processing of thousands of images and formed the foundation of the temporal analysis aimed at predicting cracking susceptibility based on physiological stress indicators.

4.3 Modeling and Evaluation Methodology

This section describes the modeling framework and evaluation protocols employed to analyze hyperspectral data for grape crack detection and prediction. The methodology is structured to progressively assess spectral discriminability, define appropriate problem

formulations, and evaluate model generalization from pixel-level classification to whole-image detection, while maintaining reproducibility and methodological consistency across experiments.

4.3.1 Global Spectral Preprocessing

All hyperspectral data were subjected to a unified and fixed preprocessing pipeline prior to any modeling or evaluation. This preprocessing was applied consistently across all datasets, models, and experimental settings to ensure methodological uniformity, reproducibility, and fair comparability between experiments Thenkarail et al., 2016.

Spectral Range Selection

To avoid the inclusion of spectral regions characterized by consistently low signal-to-noise ratios, the raw spectral signatures were restricted to a predefined wavelength range of 450–925 nm. Preliminary analysis of the data revealed high variance at the spectral extremities (400–450 nm and 925–1000 nm). This observation aligns with previous performance evaluations of the Specim IQ camera; notably, Behmann et al. Behmann et al., 2018 reported increased noise levels in the 400–450 nm range and significant reflectance artifacts in the near-infrared region (specifically above 925 nm) under outdoor conditions. Consequently, this effective operating range was selected once and retained throughout the study to avoid experiment-specific preprocessing decisions and to ensure consistency across all analyses Bioucas-Dias et al., 2013.

Quality Control and Outlier Handling

Hyperspectral datacubes were analyzed in normalized reflectance units, where spectral values are expected to lie within the physically valid range $[0, 1]$. Due to field acquisition conditions and imperfect radiometric normalization, isolated violations of this range may occur and indicate technically corrupted measurements rather than biological variability.

Outlier detection was performed using a robust multivariate statistical approach based on the Mahalanobis distance, computed within each spectral class separately. The Minimum Covariance Determinant (MCD) estimator **rousseeuw1999fast** was employed to obtain a robust estimate of the multivariate location and scatter matrix, thereby reducing the influence of the very outliers targeted for removal.

For each class, the MCD estimator was fitted to a random subsample of up to 20,000 spectra (support fraction = 0.8) to ensure computational tractability. The squared Mahalanobis distance d_i^2 was then computed for every pixel spectrum $\mathbf{x}_i$ within the class:

$$d_i^2 \;=\; (\mathbf{x}_i - \hat{\boldsymbol{\mu}})^\top \, \hat{\boldsymbol{\Sigma}}^{-1} \, (\mathbf{x}_i - \hat{\boldsymbol{\mu}})$$

where $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ denote the robust location and covariance estimates, respectively.

Spectra exceeding the $(1-p)$-th percentile of the per-class distance distribution were flagged as outliers. The outlier fraction p was treated as a configurable parameter. In the final experiments reported in this thesis, outlier removal was not applied ($p = 0$), because the classification models employed — gradient-boosted decision trees (XGBoost) — are inherently robust to moderate spectral noise and isolated anomalous pixels Thenkarail et al., 2016. Retaining all extracted spectra ensured maximal representation of rare physiological variability, including early-stage stress signatures.

All outlier flags were computed and stored alongside the spectral data to enable transparent post-hoc analysis and reproducibility of alternative exclusion criteria.

Per-Spectrum Normalization

Following quality control, each remaining spectral vector was normalized independently using *Standard Normal Variate* (SNV) normalization applied on a per-pixel basis. SNV normalization is widely used in close-range hyperspectral imaging to mitigate additive and multiplicative effects caused by illumination variability, surface geometry, and scattering, while preserving the relative spectral shape of individual samples Barnes et al., 1989; Mishra et al., 2017; Späth et al., 2024.

Per-spectrum normalization was deliberately selected to avoid reliance on dataset-level statistics and to eliminate the risk of information leakage between training and evaluation sets, particularly under cross-validation schemes such as Leave-One-Group-Out (LOGO). Information leakage during preprocessing has been identified as a critical issue in hyperspectral image classification, leading to over-optimistic performance estimates when not properly controlled Kaltenborn et al., 2025; Li et al., 2023.

4.3.2 Spectral Discriminability and Informative Wavelength Analysis

Before model construction, an explicit analysis was conducted to determine whether cracked and healthy grape tissues exhibit distinguishable spectral characteristics and to identify the spectral regions contributing most strongly to class separation. This analysis aimed to answer two foundational questions: (i) whether a statistically meaningful spectral difference exists, and (ii) where along the spectrum this difference is most pronounced.

To quantify spectral discriminability, multiple complementary criteria were employed, including class-wise separation measures and information-theoretic relevance scores. These analyses formed the basis for subsequent feature selection procedures Cover and Thomas, 2006; Thenkarail et al., 2016.

Feature Selection as a Derived Outcome

Feature selection was treated as a derivative of the spectral discriminability analysis rather than as a purely algorithmic optimization step. To ensure the biological and physical relevance of the selected bands, three complementary feature selection methods were employed, following established protocols in agricultural hyperspectral analysis Liu et al., 2022:

- **Fisher Score**: Used to quantify inter-class separation relative to intra-class variance. This filter method is widely adopted in plant phenotyping for identifying wavelengths that maximize the margin between healthy and stressed tissues J. Zhang et al., 2021.
- **Mutual Information (MI)**: Used to measure the dependency between individual wavelengths and class labels. Unlike linear methods, MI captures non-linear relationships between spectral features and physiological states, which are common in complex biological datasets Ding et al., 2023.
- **Random Forest Feature Importance**: Used to assess wavelength relevance based on ensemble-based classification performance. This embedded method is considered a standard in agricultural remote sensing due to its robustness against overfitting and its ability to handle high-dimensional spectral data Benos et al., 2021.

The outputs of the different feature selection methods were not merged into a single optimized subset. Instead, they were analyzed jointly to identify spectral regions that consistently exhibited high discriminative relevance across multiple criteria. Wavelengths that were ranked highly by more than one method were interpreted as robust and biologically meaningful candidates, while discrepancies between methods were used to highlight potential model-dependent effects.

This comparative analysis was intended to support physical interpretability and methodological robustness, rather than to enforce a single model-specific band selection scheme. Accordingly, feature selection was used as an analytical tool for understanding spectral separability, and not as a strict dimensionality reduction step.

4.3.3 Pixel-Level Classification Framework

Pixel-level classification constituted the primary modeling task and served as the foundation for all subsequent analyses. Operating at the pixel scale enabled a direct assessment of intrinsic spectral separability between cracked grape tissue and healthy tissue, while avoiding assumptions imposed by higher-level spatial aggregation or scene-level context.

Hyperspectral data are inherently high-dimensional and characterized by strong correlation between neighboring spectral bands. While such redundancy can pose challenges for certain modeling paradigms, it reflects the underlying physical continuity of plant reflectance spectra and should not be prematurely removed.

Accordingly, the initial modeling stage focused on evaluating class separability using the full spectral representation, without imposing dimensionality reduction or feature selection constraints. Non-linear decision functions were adopted to allow the modeling framework to capture complex spectral relationships while remaining robust to correlated inputs.

Feature selection and dimensionality reduction were therefore intentionally deferred to later stages of the analysis, where they were treated as interpretative and operational refinement steps rather than prerequisites for establishing detection feasibility..

Having established that cracked and healthy grape tissues exhibit stable and physically meaningful spectral differences across specific wavelength regions, the next stage investigates whether this spectral information can be represented more compactly without degrading classification performance.

Binary Pixel-Level Classification

In the binary formulation, spectral pixels were classified as either *Cracked* or *Healthy*. This formulation was adopted as a minimal baseline to assess whether cracked tissue exhibits a distinct spectral signature under controlled conditions, without explicit modeling of background complexity or non-target materials.

By restricting the problem to two biologically meaningful classes, this setting isolates the spectral characteristics of cracked tissue and provides a reference point for subsequent, more realistic formulations.

Consistent with the modeling considerations outlined above, classification was performed using non-linear decision functions, without assuming linear separability in the spectral domain.

Multi-Class Pixel-Level Classification

To account for background variability and non-berry materials commonly present in field-acquired hyperspectral images, the pixel-level framework was extended to a multi-class formulation.

In addition to the two berry-related classes (*Cracked* and *Healthy*), the multi-class formulation included explicit background classes representing common non-berry materials

encountered in the vineyard environment (e.g., leaves, branches, soil, and artificial elements). This extension provides a more realistic classification setting and enables the evaluation of model robustness under spectrally complex field conditions.

The same preprocessing pipeline and general modeling principles were retained to ensure comparability with the binary classification setting. Consequently, performance differences between the binary and multi-class formulations reflect the increased task complexity rather than changes in data processing or methodological assumptions.

The exact definition and annotation procedure of these background classes are described in the data preparation section.

4.3.4 Anomaly Detection Framework

In parallel to supervised classification, an anomaly detection framework was explored to evaluate the feasibility of identifying cracked tissue as deviations from normal spectral behavior. This formulation is particularly relevant in scenarios where labeled examples of damage are scarce, heterogeneous, or difficult to define, such as early-stage crack development.

Unlike supervised classification, anomaly detection does not assume the availability of representative samples for all target conditions. Instead, the task is formulated as learning a model of *normal* spectral behavior and identifying samples that deviate from this learned distribution.

One-Class Pixel-Level Formulation. The anomaly detection framework was implemented at the pixel level using a one-class learning paradigm. Models were trained exclusively on spectral signatures extracted from visually healthy berry tissue, which was assumed to represent normal physiological conditions.

During inference, unseen spectral pixels were evaluated based on their deviation from the learned normality model. Pixels exhibiting sufficiently large deviations were flagged as anomalous and interpreted as potential indicators of cracking-related spectral alterations.

This formulation enables the detection of abnormal spectral patterns without requiring explicit supervision for cracked tissue during training and provides a complementary perspective to supervised binary and multi-class classification approaches.

Importantly, anomaly detection was treated as a distinct analytical framework rather than an alternative classifier. While both approaches ultimately produce a binary decision (normal vs. abnormal), anomaly detection focuses on modeling the distribution of healthy tissue and identifying deviations, rather than learning explicit decision boundaries between predefined classes.

A one-class model was trained exclusively on healthy spectral signatures and evaluated on unseen samples.

4.3.5 Dimensionality Reduction Strategy

Following the establishment of stable and well-performing pixel-level models, dimensionality reduction was investigated as a secondary analysis step. The objective of this stage was to assess whether the spectral dimensionality could be reduced without compromising detection performance, motivated by practical considerations related to sensor design, computational efficiency, and operational deployment.

Dimensionality reduction was applied exclusively at the pixel level and was evaluated separately for two supervised learning settings: (i) binary pixel-level classification, and (ii) multi-class pixel-level classification.

Two distinct pixel-level datasets were therefore considered: one corresponding to the binary formulation (*Cracked* vs. *Healthy*) and one corresponding to the multi-class formulation, which included explicit background classes. In both cases, dimensionality reduction was performed after the baseline models had been validated using the full spectral representation.

Importantly, dimensionality reduction was not applied within the whole-image detection pipeline. The whole-image experiments relied on pixel-level models trained on the original or reduced spectral representations, but no additional learning or adaptation of dimensionality reduction was performed directly at the image level. As such, the dimensionality reduction analysis is interpreted as an investigation of spectral efficiency at the pixel scale, rather than as a standalone optimization of the whole-image detection pipeline.

Three complementary wavelength selection strategies were employed, each offering a different trade-off between computational cost, optimality guarantees, and interpretability:

1. **Backward Feature Selection (BFS):** A deterministic wrapper method that starts from the full set of retained wavelengths and iteratively removes the least informative band at each step, re-evaluating the pixel-level classifier after every removal. The evaluation metric was the Crack class PR–AUC (area under the precision–recall curve), assessed on a held-out validation split stratified by grape cluster identity. BFS provides a monotonic elimination path and is well suited for identifying the minimal subset of bands that preserves near-optimal detection performance.

2. **Minimum Redundancy–Maximum Relevance (mRMR):** A filter-based ranking method that selects wavelengths by jointly maximizing relevance to the target

label and minimizing redundancy among selected features. Feature ranking was performed using a Linear Discriminant Analysis (LDA) classifier with a redundancy weight of 1.0. The ranked feature list was evaluated for subsets ranging from 5 to 50 wavelengths.

3. **Genetic Algorithm (GA):** A stochastic wrapper method inspired by evolutionary optimization. A population of candidate wavelength subsets (population size = 60) was evolved over 40 generations using tournament selection (tournament size = 4), single-point crossover (rate = 0.8), and bit-flip mutation (rate = 0.05). Subset sizes were constrained to the range $[5, 50]$ wavelengths, and a small penalty ($\alpha = 0.002$) was applied to the number of selected features to encourage parsimony. Fitness was defined as the Crack class F1-score on the validation fold.

All three methods were evaluated using the best-performing pixel-level classifier identified in the preceding classification experiments (XGBoost), and results were compared in terms of detection performance as a function of the number of retained wavelengths.

4.3.6 Whole-Image Detection and Aggregation Pipeline

Building upon the pixel-level classification outputs, a whole-image detection and localization pipeline was developed to enable robust crack identification in realistic vineyard scenes. Rather than introducing an additional image-level classifier, the pipeline aggregates pixel-level predictions into a system-level framework that enforces spatial coherence and suppresses isolated false positive responses.

The pipeline operates on pixel-wise crack probability maps generated by the trained pixel-level classifiers and progressively aggregates them into region-level, patch-level, and image-level decisions, bridging the gap between noisy pixel-level predictions and biologically meaningful image-level interpretation.

Pipeline Architecture and Spatial Aggregation

The whole-image aggregation pipeline follows a sequential process, illustrated in Figure 4.5, and consists of the following stages:

- **Input Hyperspectral Image:** Each input consists of a full hyperspectral datacube ($H \times W \times B$) acquired under field conditions.
- **Pixel-Level Probability Map Generation:** The trained pixel-level classifier produces a dense pixel-wise crack probability map with values in the range $[0, 1]$.

- **Morphological Closing and Region Formation:** Spatially adjacent pixels with elevated crack probabilities are grouped into connected regions (BLOBs) using morphological closing operations, enforcing spatial coherence.
- **Blob Size Filtering:** Small connected regions below a minimum size threshold are removed to suppress isolated false positive detections.
- **Patch-Based Division and Decision:** The image is divided into fixed-size patches. A patch is marked as crack-positive if the proportion of crack-related pixels within the patch exceeds a predefined threshold.
- **Image-Level Decision:** The final image-level decision is obtained by aggregating patch-level outcomes. An image is classified as *Cracked* if at least one patch is identified as crack-positive; otherwise, it is classified as *Non-Cracked*.

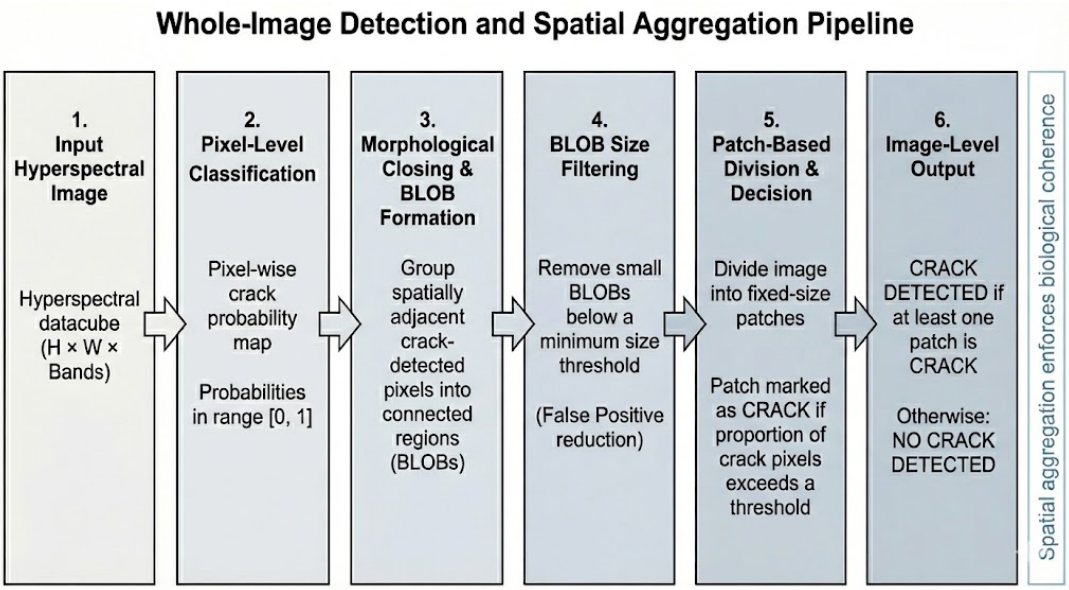


Figure 4.5: Schematic overview of the whole-image detection and spatial aggregation pipeline.

Whole-Image Dataset and Evaluation Setup

Whole-image evaluation was conducted on a dedicated dataset of full hyperspectral images, comprising 10 images containing grape clusters with visible cracking and 60 images containing clusters with no visible cracking. The aggregation pipeline was applied to patch-based representations extracted from the full images, enabling localized detection while preserving image-level context.

Evaluation Across Disease Progression Stages

To assess robustness across different stages of crack development, whole-image performance was evaluated separately for early-stage and late-stage samples. Late-stage images contained clusters with clearly visible macroscopic cracking, whereas early-stage images were annotated as exhibiting no visible cracking based on expert visual inspection. This stratification was applied exclusively at the evaluation stage and did not affect model training or aggregation design.

4.3.7 Evaluation Protocols and Performance Metrics

Model performance was evaluated using task-specific protocols and metrics tailored to each stage of the analysis, reflecting the differing objectives and operational requirements of pixel-level classification, anomaly detection, and whole-image detection.

Evaluation metrics were selected according to the objective of each experimental stage. Pixel-level and anomaly detection metrics are summarized in Table 4.2, while system-level evaluation for whole-image detection is reported in Table 4.3.

Cross-Validation Strategy

To assess generalization across biological samples and prevent spatial or spectral data leakage, a Leave-One-Group-Out (LOGO) cross-validation strategy was employed. Each group corresponded to a single grape cluster, such that all pixels extracted from the same cluster were assigned exclusively to either the training or test set in each fold.

This strategy ensures that evaluation is performed on previously unseen biological samples and provides a realistic estimate of model performance under field deployment conditions.

Pixel-Level Classification Metrics

For binary pixel-level classification, performance was evaluated primarily using the F1-score for the *Cracked* class ($F1_{CRACK}$) and the Precision–Recall Area Under the Curve (PR-AUC). These metrics were selected due to the strong class imbalance and the emphasis on correctly identifying cracked tissue.

Additional metrics, including precision, recall, and ROC-AUC, were reported as supporting indicators to facilitate comprehensive performance analysis.

For multi-class pixel-level classification, Macro-F1 and per-class F1-scores were used as primary metrics to ensure balanced evaluation across all classes, with particular attention to the *Cracked* class. Confusion matrices were used to analyze class-specific error patterns and background-related confusions.

Anomaly Detection Evaluation

For anomaly detection experiments, model performance was evaluated using ROC-AUC and PR-AUC, reflecting the ranking-based nature of one-class detection and the absence of a fixed decision threshold. These metrics quantify the ability of the model to distinguish anomalous (cracked) spectra from normal (healthy) spectral behavior across a range of operating points.

Whole-Image and Patch-Based Evaluation

Whole-image detection performance was evaluated at the patch and image levels, reflecting the system-level objective of identifying cracked grape clusters under realistic field conditions.

Given the higher cost associated with missed detections, recall-oriented metrics were emphasized. In particular, Recall and the F2-score for the *Cracked* class ($F2_{CRACK}$) were used as primary evaluation criteria, assigning greater weight to recall than precision.

Precision and false positive rates were reported as complementary metrics to characterize trade-offs and system behavior. Quantitative evaluation was supported by qualitative visual inspection of detection overlays to analyze spatial consistency, typical failure modes, and false positive patterns.

Table 4.2: Evaluation metrics for pixel-level and anomaly detection experiments

Task	**Primary Metrics**	**Secondary Metrics**
Binary pixel-level classification	$F1_{CRACK}$, PR-AUC	$Recall_{CRACK}$, $Precision_{CRACK}$, ROC-AUC
Multi-class pixel-level classification	Macro-F1, $F1_{CRACK}$	Per-class Precision/Recall, Confusion Matrix
Anomaly detection (one-class)	ROC-AUC, PR-AUC	Recall@Fixed-FPR, Precision@Top-K

Table 4.3: Evaluation metrics for whole-image and patch-based detection

Task	**Primary Metrics**	**Secondary Metrics**
System-level detection of cracked grape clusters	Recall, $F2_{CRACK}$	Precision, Image-level Accuracy, False Positive Rate

Chapter 5

Experimental Results

5.1 Overview of Experimental Results

This chapter presents the quantitative results obtained from the experimental evaluation of hyperspectral-based crack detection and prediction methods. The results are organized to follow the logical progression of the experimental pipeline, beginning with spectral characterization, continuing through pixel-level and whole-image modeling, and concluding with temporal and longitudinal analyses.

Only empirical findings are reported in this chapter, without interpretation or comparison to prior studies.

5.2 Spectral Signature Characterization

This section examines the spectral properties of grape tissue to establish whether measurable differences exist between healthy and cracked samples prior to any modeling stage.

5.2.1 Mean Spectral Signatures of Healthy and Cracked Tissue

Mean spectral signatures were computed for pixels sampled from healthy and cracked grape tissue across the full VIS–NIR range. The analysis focuses on class-level spectral trends and variability, without applying any classification or predictive modeling.

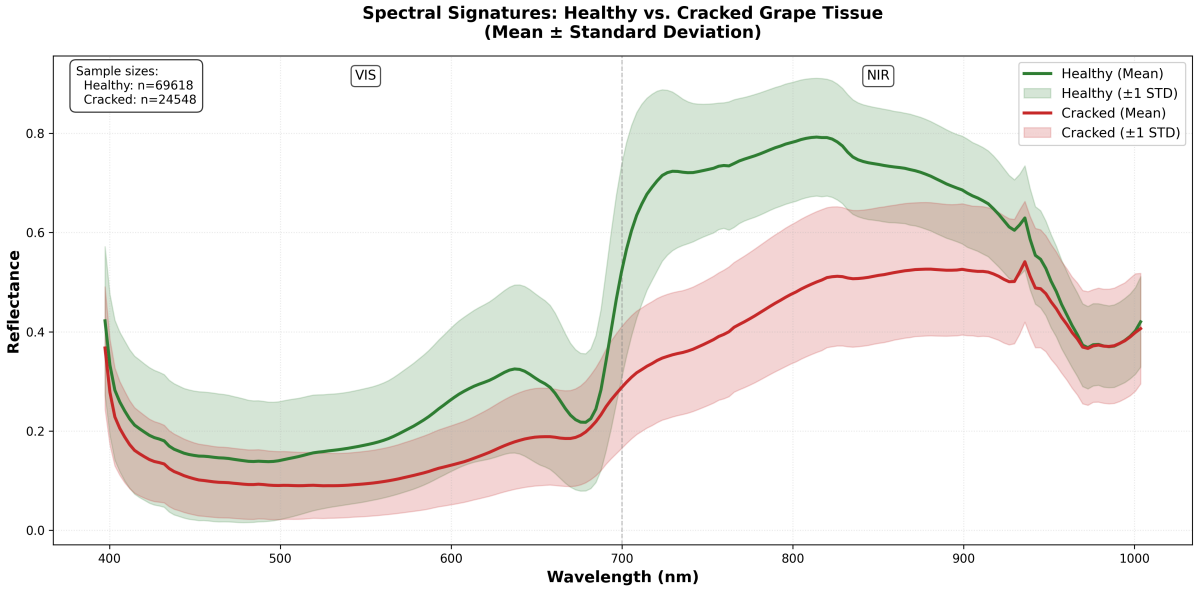


Figure 5.1: Mean spectral signatures (± standard deviation) for healthy and cracked tissue across the VIS–NIR range.

To provide a compact quantitative summary, reflectance values were aggregated over two broad spectral regions corresponding to the visible (VIS) and near-infrared (NIR) ranges.

Table 5.1: Summary statistics of reflectance values across spectral regions for healthy and cracked tissue.

Spectral Region	Class	Mean Reflectance	Std. Dev.	Median
VIS (400–700 nm)	Healthy	0.218	0.128	0.186
	Cracked	0.137	0.076	0.123
NIR (700–1000 nm)	Healthy	0.660	0.114	0.630
	Cracked	0.450	0.121	0.425

5.2.2 Spectral Separability Across Wavelength Regions

This subsection quantifies the degree of spectral separation between healthy and cracked tissue across predefined wavelength regions. The goal is to identify spectral ranges exhibiting the strongest inter-class contrast, without invoking classification models.

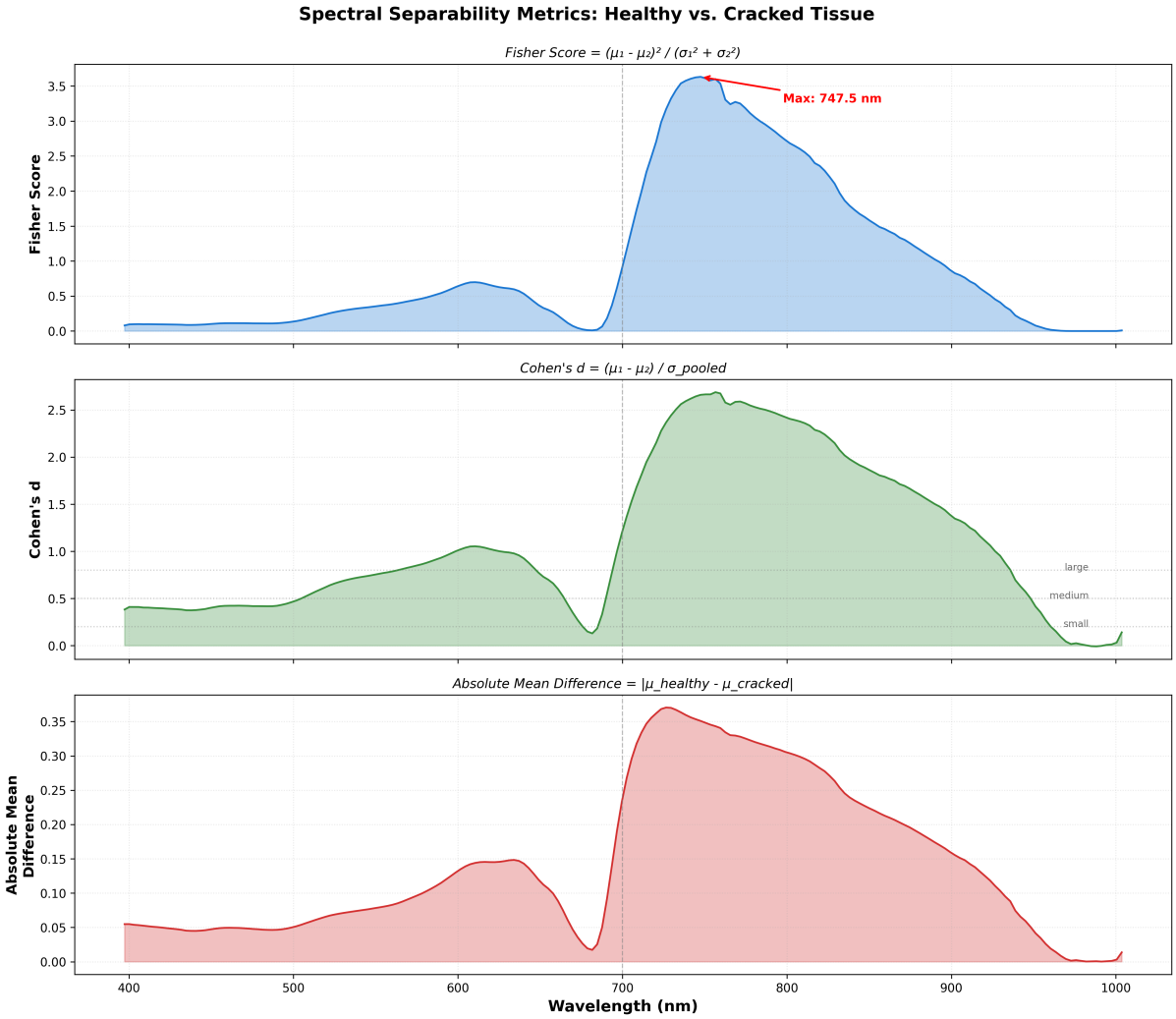


Figure 5.2: Wavelength-wise spectral separability metrics between healthy and cracked tissue, including absolute mean difference, Cohen's d, and Fisher score.

To facilitate regional-level comparison, the wavelength-wise separability measures were aggregated over predefined spectral regions.

Table 5.2: Regional spectral separability metrics between healthy and cracked tissue. Mean values are averaged across wavelengths within each range, while peak values indicate the maximum separability and the corresponding wavelength.

Metric	VIS Range	NIR Range	Full Spectrum
Mean $\lvert\Delta\mu\rvert$ (Abs. Mean Diff.)	0.0810	0.2101	0.1422
Max $\lvert\Delta\mu\rvert$ (nm)	0.2325 (708.6)	0.3704 (726.5)	0.3704 (726.5)
Mean Cohen's d	0.6330	1.6217	1.1014
Max Cohen's d (nm)	1.1910 (699.5)	2.6877 (756.6)	2.6877 (756.6)
Mean Fisher Score	0.2895	1.6590	0.9436
Max Fisher Score (nm)	0.8821 (694.5)	3.6311 (747.5)	3.6311 (747.5)

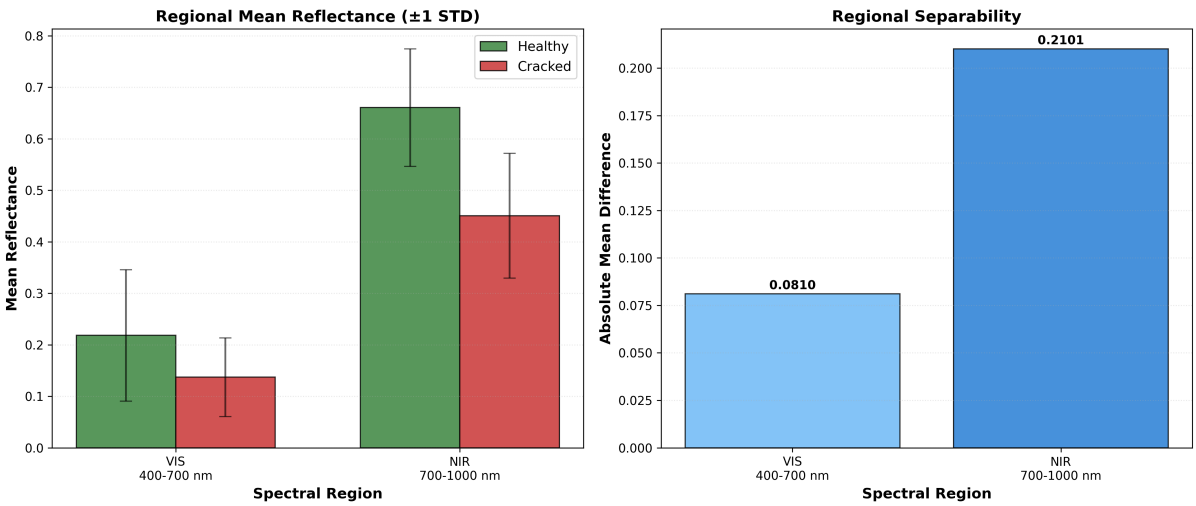


Figure 5.3: Regional visual summary of mean reflectance (±1 STD) and absolute mean difference between healthy and cracked tissue in the VIS and NIR ranges. This figure provides an intuitive visualization of the regional trends quantified in Table 5.2.

5.3 Pixel-Level Classification Performance

This section reports the performance of supervised pixel-level classification models trained on hyperspectral signatures.

5.3.1 Experimental Setup and Evaluation Protocol

Pixel-level classification performance was evaluated using a cross-validation framework designed to assess generalization across grape clusters. Standard classification metrics were computed for all experiments.

5.3.2 Binary Pixel-Level Classification Results

Binary pixel-level classification was evaluated for distinguishing healthy from cracked tissue. Performance was assessed using Accuracy, ROC–AUC, and PR–AUC, together with class-specific Precision and Recall for the *Cracked* class (positive class).

Two sampling regimes were considered: the original naturally imbalanced distribution and a class-balanced configuration.

Unbalanced vs. Balanced Training Sets

Results obtained under both sampling regimes are summarized jointly on the following landscape page. Table 5.3 reports the classification performance obtained using the original unbalanced dataset, while Table 5.4 presents the corresponding results obtained under balanced sampling. The identical evaluation protocol and model set are used in both cases, enabling a direct comparison between regimes.

Table 5.3: Pixel-level classification metrics (unbalanced training set). Values are reported as mean ± standard deviation across folds where available. Precision and recall are reported for the *Cracked* class; F1 is the weighted F1-score.

Model	Accuracy	Precision	Recall	F1	ROC–AUC
PLS-DA	0.9927 ± 0.0168	0.9999	0.9682	0.9924	1.0000 ± 0.0000
Logistic Regression (L1)	0.9962 ± 0.0074	0.9890	0.9948	0.9963	0.9996 ± 0.0012
SVM (RBF)	0.9938 ± 0.0094	0.9943	0.9787	0.9938	0.9997 ± 0.0010
Random Forest	0.9968 ± 0.0055	0.9964	0.9904	0.9968	0.9990 ± 0.0035
XGBoost	0.9947 ± 0.0103	0.9953	0.9798	0.9946	1.0000 ± 0.0000

Table 5.4: Pixel-level classification metrics (balanced training set). Values are reported as mean ± standard deviation across folds where available. Precision and recall are reported for the *Cracked* class; F1 is the weighted F1-score.

Model	Accuracy	Precision	Recall	F1	ROC–AUC
PLS-DA	0.9961 ± 0.0086	0.9998	0.9929	0.9961	1.0000 ± 0.0000
Logistic Regression (L1)	0.9930 ± 0.0148	0.9994	0.9861	0.9930	0.9999 ± 0.0002
SVM (RBF)	0.9876 ± 0.0202	0.9948	0.9771	0.9875	0.9998 ± 0.0007
Random Forest	0.9935 ± 0.0110	0.9970	0.9882	0.9934	0.9990 ± 0.0038
XGBoost	0.9934 ± 0.0107	0.9972	0.9882	0.9933	1.0000 ± 0.0001

5.3.3 Pixel-Level Classification Results Under Different Labeling and Balancing Regimes (10 classes)

This subsection reports pixel-level classification performance under three labeling strategies: (i) *crack_regular_rest*, (ii) *crack_vs_rest*, and (iii) a *multi_class* formulation. All models were evaluated using identical domain-aware LOGO cross-validation splits. For each labeling strategy, two sampling regimes were considered: a class-balanced configuration and the original naturally imbalanced distribution, resulting in six experimental configurations in total.

Performance is reported as mean ± standard deviation across folds. Reported metrics include global measures (Accuracy, Balanced Accuracy and Macro-F1), CRACK-specific metrics (Precision, Recall, F1, ROC–AUC and PR–AUC), and total runtime per experiment. The evaluated classifiers include Logistic Regression (L1), Random Forest, XGBoost, a shallow Multi-Layer Perceptron (MLP), and PLS-DA.

Balanced Training Sets

Results obtained under balanced sampling are summarized in Table **??**. The table reports performance for both binary labeling schemes and for the multi-class configuration.

Unbalanced Training Sets

Results obtained on naturally imbalanced data are summarized in Table **??**. The same set of models and metrics are reported for direct comparison with the balanced regime.

Table 5.5: Pixel-level classification metrics (**unbalanced** training set). Values are reported as mean ± standard deviation across folds.

Label setup	Model	Accuracy	Balanced Acc.	Precision (Cracked)	Recall (Cracked)	Macro-F1	ROC–AUC (Cracked)	PR–AUC (Cracked)
crack_regular_rest								
	PLS-DA	0.5087 ± 0.0044	0.6879 ± 0.0239	0.8825 ± 0.1174	0.7537 ± 0.2385	0.5819 ± 0.0199	0.9977 ± 0.0049	0.8630 ± 0.1242
	Logistic Regression (L1)	0.8537 ± 0.0021	0.8683 ± 0.0075	0.9654 ± 0.0185	0.9663 ± 0.0748	0.8010 ± 0.0049	0.9993 ± 0.0021	0.9839 ± 0.0116
	Random Forest	0.9175 ± 0.0056	0.9179 ± 0.0116	0.8942 ± 0.0644	0.9394 ± 0.0701	0.8904 ± 0.0106	0.9975 ± 0.0037	0.9622 ± 0.0495
	XGBoost	0.9212 ± 0.0040	0.9162 ± 0.0094	0.9804 ± 0.0118	0.9493 ± 0.0894	0.8832 ± 0.0074	0.9992 ± 0.0017	0.9902 ± 0.0160
	MLP (Small)	0.9175 ± 0.0059	0.9166 ± 0.0122	0.9530 ± 0.0272	0.9729 ± 0.0398	0.8759 ± 0.0176	0.9994 ± 0.0017	0.9911 ± 0.0122
crack_vs_rest								
	PLS-DA	0.5120 ± 0.0047	0.6940 ± 0.0189	0.8843 ± 0.1117	0.7623 ± 0.2303	0.5893 ± 0.0164	0.9979 ± 0.0045	0.8669 ± 0.1186
	Logistic Regression (L1)	0.8574 ± 0.0022	0.8704 ± 0.0076	0.9669 ± 0.0190	0.9651 ± 0.0759	0.8049 ± 0.0050	0.9993 ± 0.0022	0.9847 ± 0.0114
	Random Forest	0.9147 ± 0.0060	0.9159 ± 0.0126	0.8871 ± 0.0701	0.9368 ± 0.0751	0.8857 ± 0.0116	0.9975 ± 0.0041	0.9610 ± 0.0535
	XGBoost	0.9200 ± 0.0043	0.9154 ± 0.0098	0.9809 ± 0.0120	0.9508 ± 0.0888	0.8825 ± 0.0078	0.9992 ± 0.0018	0.9900 ± 0.0165
	MLP (Small)	0.9163 ± 0.0062	0.9158 ± 0.0129	0.9521 ± 0.0282	0.9736 ± 0.0390	0.8741 ± 0.0188	0.9994 ± 0.0017	0.9912 ± 0.0122
multi_class								
	PLS-DA	0.9670 ± 0.0211	0.9787 ± 0.0121	0.7204 ± 0.1749	0.9927 ± 0.0127	0.9013 ± 0.0709	0.9978 ± 0.0025	0.9677 ± 0.0372
	Logistic Regression (L1)	0.9972 ± 0.0031	0.9959 ± 0.0097	0.9675 ± 0.0225	0.9977 ± 0.0078	0.9920 ± 0.0085	0.9996 ± 0.0016	0.9971 ± 0.0079
	Random Forest	0.9965 ± 0.0030	0.9940 ± 0.0120	0.9675 ± 0.0234	0.9969 ± 0.0090	0.9906 ± 0.0100	0.9994 ± 0.0026	0.9970 ± 0.0077
	XGBoost	0.9978 ± 0.0021	0.9946 ± 0.0103	0.9812 ± 0.0149	0.9908 ± 0.0206	0.9923 ± 0.0074	0.9995 ± 0.0021	0.9977 ± 0.0069
	MLP (Small)	0.9905 ± 0.0249	0.9890 ± 0.0224	0.9244 ± 0.1648	0.9866 ± 0.0259	0.9699 ± 0.0704	0.9965 ± 0.0083	0.9898 ± 0.0247

Table 5.6: Pixel-level classification metrics (**balanced** training set). Values are reported as mean ± standard deviation across folds.

Label setup	Model	Accuracy	Balanced Acc.	Precision (Cracked)	Recall (Cracked)	Macro-F1	ROC–AUC (Cracked)	PR–AUC (Cracked)
crack_regular_rest								
	PLS-DA	0.4771 ± 0.0088	0.7061 ± 0.0107	0.8433 ± 0.0910	0.9451 ± 0.1069	0.5875 ± 0.0090	0.9982 ± 0.0046	0.9500 ± 0.0899
	Logistic Regression (L1)	0.8629 ± 0.0033	0.8653 ± 0.0073	0.9865 ± 0.0067	0.9671 ± 0.0739	0.8097 ± 0.0054	0.9994 ± 0.0018	0.9874 ± 0.0090
	Random Forest	0.9277 ± 0.0046	0.9199 ± 0.0110	0.9101 ± 0.0533	0.9194 ± 0.1124	0.9047 ± 0.0087	0.9975 ± 0.0035	0.9725 ± 0.0438
	XGBoost	0.9242 ± 0.0026	0.9111 ± 0.0084	0.9932 ± 0.0042	0.9437 ± 0.0954	0.8839 ± 0.0067	0.9993 ± 0.0016	0.9919 ± 0.0175
	MLP (Small)	0.9219 ± 0.0053	0.9140 ± 0.0125	0.9831 ± 0.0123	0.9832 ± 0.0223	0.8869 ± 0.0167	0.9982 ± 0.0057	0.9934 ± 0.0097
crack_vs_rest								
	PLS-DA	0.4782 ± 0.0090	0.7078 ± 0.0106	0.8445 ± 0.0904	0.9455 ± 0.1063	0.5889 ± 0.0093	0.9982 ± 0.0046	0.9507 ± 0.0886
	Logistic Regression (L1)	0.8642 ± 0.0032	0.8662 ± 0.0074	0.9866 ± 0.0067	0.9674 ± 0.0738	0.8109 ± 0.0055	0.9994 ± 0.0018	0.9876 ± 0.0089
	Random Forest	0.9261 ± 0.0049	0.9191 ± 0.0113	0.9092 ± 0.0546	0.9192 ± 0.1128	0.9038 ± 0.0091	0.9974 ± 0.0036	0.9722 ± 0.0444
	XGBoost	0.9238 ± 0.0027	0.9110 ± 0.0085	0.9931 ± 0.0042	0.9439 ± 0.0954	0.8836 ± 0.0068	0.9993 ± 0.0016	0.9919 ± 0.0176
	MLP (Small)	0.9221 ± 0.0052	0.9143 ± 0.0126	0.9830 ± 0.0124	0.9833 ± 0.0222	0.8871 ± 0.0165	0.9982 ± 0.0057	0.9934 ± 0.0097
multi_class								
	PLS-DA	0.9666 ± 0.0210	0.9786 ± 0.0120	0.7179 ± 0.1759	0.9926 ± 0.0129	0.9010 ± 0.0707	0.9978 ± 0.0025	0.9674 ± 0.0375
	Logistic Regression (L1)	0.9973 ± 0.0031	0.9960 ± 0.0097	0.9677 ± 0.0224	0.9977 ± 0.0078	0.9921 ± 0.0085	0.9996 ± 0.0016	0.9971 ± 0.0079
	Random Forest	0.9966 ± 0.0030	0.9941 ± 0.0120	0.9676 ± 0.0233	0.9969 ± 0.0090	0.9907 ± 0.0100	0.9994 ± 0.0026	0.9970 ± 0.0077
	XGBoost	0.9979 ± 0.0021	0.9947 ± 0.0103	0.9813 ± 0.0149	0.9908 ± 0.0206	0.9924 ± 0.0074	0.9995 ± 0.0021	0.9977 ± 0.0069
	MLP (Small)	0.9906 ± 0.0249	0.9891 ± 0.0224	0.9246 ± 0.1647	0.9866 ± 0.0259	0.9700 ± 0.0704	0.9965 ± 0.0083	0.9898 ± 0.0247

5.3.4 Autoencoder-Based Pixel-Level Anomaly Detection

In addition to supervised classifiers, pixel-level crack detection was evaluated using reconstruction-based anomaly detection with autoencoders. The experiments were designed in a progressive manner. First, a binary formulation was considered, in which the autoencoder was trained exclusively on healthy (non-cracked) pixels and evaluated against cracked tissue. Subsequently, the setting was extended to a multi-class environment, in which background and nuisance categories (e.g., leaves, branches, plastic and calibration targets) were explicitly included during testing.

Within the multi-class regime, two complementary training strategies were investigated. In the first, the autoencoder was trained solely on pixels belonging to the *Cracked* class, such that crack signatures define the reconstruction manifold and all remaining classes are treated as anomalous. In the second strategy, the model was trained on all semantic classes except *Cracked*, thereby enforcing accurate reconstruction of diverse non-crack appearances while promoting elevated reconstruction errors for crack pixels at inference time.

Across all configurations, three fully-connected architectures were examined: *Autoencoder (64–32–16)*, *Autoencoder (128–64–32)*, and *Autoencoder (64–32–8)*. All models were trained and evaluated using identical domain-aware LOGO cross-validation splits employed in the supervised experiments.

Anomaly scores were computed from reconstruction errors and converted into binary predictions using a percentile-based threshold estimated on the training folds. Performance is reported as mean ± standard deviation across folds. Reported metrics include Accuracy, Precision, Recall, F1, ROC–AUC and PR–AUC for the *Cracked* class, together with aggregate confusion counts across folds and computational cost indicators. For each autoencoder formulation, two sampling regimes were considered: the original naturally imbalanced distribution and a class-balanced configuration.

Table 5.7: Pixel-level anomaly detection performance using autoencoders trained on non-cracked samples (unbalanced sampling). Values are reported as mean ± standard deviation across folds. Metrics correspond to the *Cracked* class.

Architecture	Accuracy	Precision	Recall	F1	ROC–AUC	PR–AUC	Train Time [s]	Infer Time [s]
Autoencoder (64–32–16)	0.8969 ± 0.1646	0.8226 ± 0.1925	0.9971 ± 0.0016	0.9017 ± 0.1047	0.9814 ± 0.0327	0.9554 ± 0.0900	63.34 ± 2.80	0.043 ± 0.004
Autoencoder (128–64–32)	0.8597 ± 0.1813	0.7544 ± 0.2308	0.9974 ± 0.0015	0.8628 ± 0.1270	0.9673 ± 0.0469	0.9253 ± 0.1190	60.90 ± 1.48	0.057 ± 0.001
Autoencoder (64–32–8)	0.8923 ± 0.1648	0.8126 ± 0.1938	0.9978 ± 0.0013	0.8965 ± 0.1050	0.9799 ± 0.0334	0.9511 ± 0.0913	61.02 ± 1.46	0.041 ± 0.001

Table 5.8: Pixel-level anomaly detection performance using autoencoders trained on non-cracked samples (balanced sampling). Values are reported as mean ± standard deviation across folds. Metrics correspond to the *Cracked* class.

Architecture	Accuracy	Precision	Recall	F1	ROC–AUC	PR–AUC	Train Time [s]	Infer Time [s]
Autoencoder (64–32–16)	0.9343 ± 0.0989	0.9122 ± 0.1348	0.9906 ± 0.0340	0.9437 ± 0.0820	0.9823 ± 0.0457	0.9702 ± 0.0592	65.35 ± 1.90	0.008 ± 0.002
Autoencoder (128–64–32)	0.9083 ± 0.1332	0.8771 ± 0.1576	0.9982 ± 0.0058	0.9256 ± 0.1016	0.9919 ± 0.0178	0.9833 ± 0.0357	58.49 ± 3.87	0.007 ± 0.001
Autoencoder (64–32–8)	0.9353 ± 0.1027	0.9074 ± 0.1338	0.9977 ± 0.0100	0.9449 ± 0.0833	0.9858 ± 0.0344	0.9757 ± 0.0501	52.04 ± 2.44	0.006 ± 0.002

Table 5.9: Pixel-level anomaly detection performance using autoencoders trained exclusively on *Cracked* pixels and evaluated in a multi-class setting (unbalanced sampling). Values are reported as mean across LOGO folds. Metrics correspond to the *Cracked* class.

Architecture	Accuracy	Precision	Recall	F1	ROC–AUC	PR–AUC	Train Time [s]	Infer Time [s]
Autoencoder (64–32–16)	0.9660	0.9995	0.9662	0.9826	0.9858	0.9999	26.96	0.148
Autoencoder (128–64–32)	0.9861	0.9991	0.9869	0.9930	0.9846	0.9999	27.12	0.207
Autoencoder (64–32–8)	0.9629	0.9994	0.9632	0.9809	0.9834	0.9999	26.20	0.155

Table 5.10: Pixel-level anomaly detection performance using autoencoders trained exclusively on *Cracked* pixels and evaluated in a multi-class setting (balanced sampling). Values are reported as mean ± standard deviation across LOGO folds. Metrics correspond to the *Cracked* class.

Architecture	Accuracy	Precision	Recall	F1	ROC–AUC	PR–AUC	Train Time [s]	Infer Time [s]
Autoencoder (64–32–16)	0.9660 ± 0.0106	0.9995 ± 0.0008	0.9662 ± 0.0108	0.9826 ± 0.0055	0.9858 ± 0.0281	0.9999 ± 0.0001	61.06 ± 7.56	0.045 ± 0.013
Autoencoder (128–64–32)	0.9861 ± 0.0052	0.9991 ± 0.0017	0.9869 ± 0.0051	0.9930 ± 0.0027	0.9846 ± 0.0274	0.9999 ± 0.0001	63.03 ± 8.09	0.056 ± 0.004
Autoencoder (64–32–8)	0.9629 ± 0.0149	0.9994 ± 0.0011	0.9632 ± 0.0150	0.9809 ± 0.0079	0.9834 ± 0.0298	0.9999 ± 0.0002	59.80 ± 8.55	0.041 ± 0.003

Table 5.11: Pixel-level anomaly detection performance using autoencoders trained on all semantic classes except *Cracked* and evaluated in a multi-class setting (unbalanced sampling). Values are reported for the held-out test split. Metrics correspond to the *Cracked* class.

Architecture	Accuracy	Precision	Recall	F1	ROC–AUC	PR–AUC	Train Time [s]	Infer Time [s]
Autoencoder (64–32–16)	0.9969	0.0096	0.1110	0.0177	0.4575	0.0083	1429.22	0.388
Autoencoder (128–64–32)	0.9986	0.0118	0.0126	0.0122	0.4513	0.0084	999.73	0.255
Autoencoder (64–32–8)	0.9984	0.0207	0.0316	0.0250	0.4542	0.0110	916.85	0.141

Table 5.12: Pixel-level anomaly detection performance using autoencoders trained on all semantic classes except *Cracked* and evaluated in a multi-class setting (balanced sampling). Values are reported for the held-out test split. Metrics correspond to the *Cracked* class.

Architecture	Accuracy	Precision	Recall	F1	ROC–AUC	PR–AUC	Train Time [s]	Infer Time [s]
Autoencoder (64–32–16)	0.9791	0.0000	0.0000	0.0000	0.2472	0.0083	2567.17	0.381
Autoencoder (128–64–32)	0.9789	0.0000	0.0000	0.0000	0.2847	0.0084	2260.73	0.299
Autoencoder (64–32–8)	0.9781	0.0000	0.0000	0.0000	0.4513	0.0110	1853.00	0.227

5.3.5 Full-Image Inference: Early vs. Late Season Experiments

To evaluate the temporal robustness of the proposed full-image crack-detection pipeline, two independent optimization and evaluation scenarios were considered, corresponding to early-season and late-season acquisition periods. In both cases, post-processing and aggregation parameters were optimized on a calibration split using Optuna and subsequently assessed on an unseen test set.

Each experiment employed identical inference logic, including pixel-level probability thresholding, morphological filtering, patch-level aggregation and global crack-percentage criteria. In addition to the supervised pixel-level classifier, the same full-image pipeline was evaluated using an autoencoder-based anomaly detector (architecture 128–64–32, trained exclusively on *Cracked* pixels). For the autoencoder, reconstruction errors (MSE) were converted to pseudo-probabilities using a sigmoid transformation centered on the training-set threshold, after which the identical post-processing and aggregation logic was applied.

Performance was quantified on both calibration and test splits using a comprehensive set of metrics, encompassing accuracy-based measures, class-specific precision and recall, harmonic means (F1 and F2), specificity, negative predictive value, and the Matthews correlation coefficient, together with the underlying confusion matrices.

The dataset composition for the two temporal scenarios is reported in Table 5.13. Quantitative detection performance is summarized in Table 5.14, enabling a direct comparison between early- and late-season conditions and between calibration and test generalization behavior.

Table 5.13: Dataset composition for full-image crack-detection experiments under early- and late-season conditions. Row 1 was used for calibration, while Row 2 served as an independent test set.

Row	Season	Split	Total	Unique	Pos	Neg
Row 1	Early	Calibration	172	60	32	140
Row 1	Late	Calibration	160	60	35	125
Row 2	Early	Test	60	60	32	28
Row 2	Late	Test	60	60	34	26

Table 5.14: Full-image crack detection performance for early- and late-season experiments. Metrics are reported separately for calibration and test splits. Confusion-matrix entries correspond to the number of grape clusters classified as cracked or healthy. All derived statistics refer to the *Cracked* class.

Season	Split	TP	FP	TN	FN	Acc	BalAcc	Prec	Rec	F1	F2	Spec	NPV	MCC
Late	Calibration	34	5	120	1	0.963	0.966	0.872	0.971	0.919	0.950	0.960	0.992	0.897
Late	Test	27	5	21	7	0.800	0.801	0.844	0.794	0.818	0.804	0.808	0.750	0.598
Early	Calibration	23	6	134	9	0.913	0.838	0.793	0.719	0.754	0.732	0.957	0.937	0.703
Early	Test	19	3	25	13	0.733	0.743	0.864	0.594	0.704	0.633	0.893	0.658	0.504

Table 5.15: Full-image crack detection using autoencoder-based anomaly detection (128–64–32, trained on *Cracked* pixels). Reconstruction errors (MSE) were converted to pseudo-probabilities via sigmoid transformation; the same post-processing pipeline (thresholding, morphological filtering, patch-based aggregation) was applied. Metrics are reported for calibration and test splits under early- and late-season conditions. All derived statistics refer to the *Cracked* class.

Season	Split	TP	FP	TN	FN	Acc	BalAcc	Prec	Rec	F1	F2	Spec	NPV	MCC
Late	Calibration	26	7	118	9	0.900	0.843	0.788	0.743	0.765	0.751	0.944	0.929	0.702
Late	Test	22	7	19	12	0.683	0.689	0.759	0.647	0.698	0.667	0.731	0.613	0.375
Early	Calibration	19	17	123	13	0.826	0.736	0.528	0.594	0.559	0.579	0.879	0.904	0.452
Early	Test	13	13	15	19	0.467	0.471	0.500	0.406	0.448	0.422	0.536	0.441	−0.058

5.3.6 Wavelength selection (BFS) setup

Backward Feature Selection (BFS) was applied to the hyperspectral channels to evaluate full-image crack-detection performance as a function of the number of selected wavelengths. At each BFS step, the pipeline performance was recorded using the CRACK class metrics.

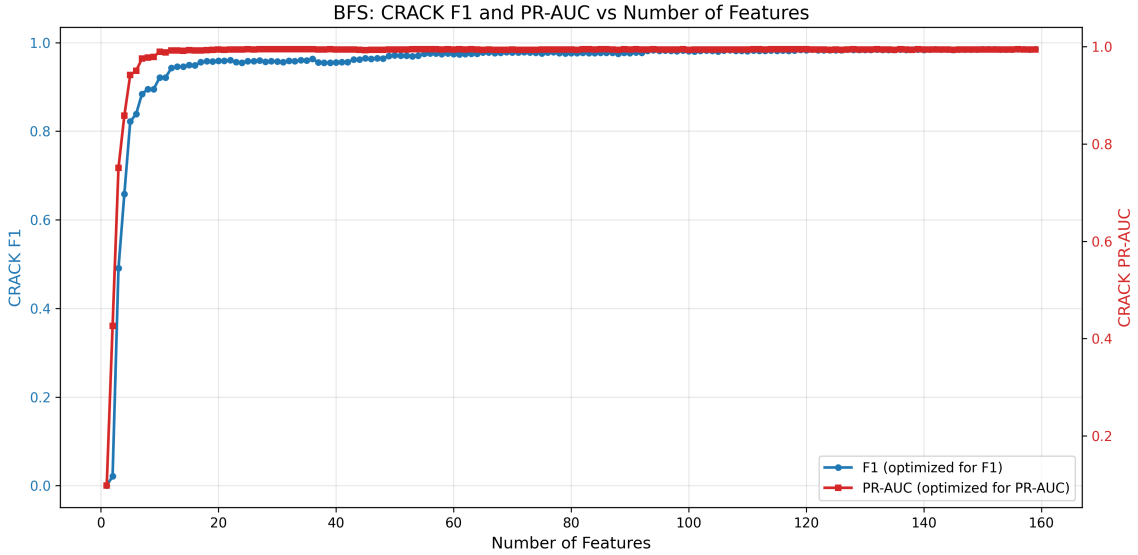
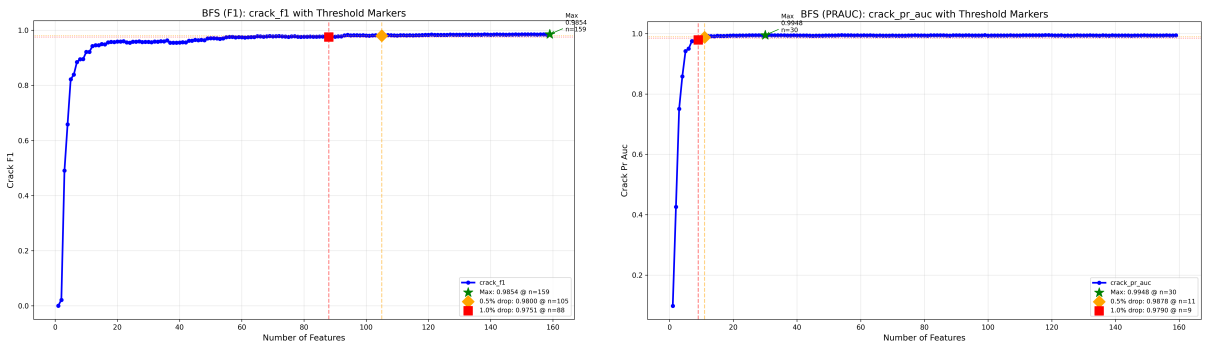


Figure 5.4: BFS results: CRACK F1 and CRACK PR–AUC as a function of the number of selected wavelengths.



(a) CRACK F1 with threshold markers.

(b) CRACK PR–AUC with threshold markers.

Figure 5.5: BFS performance curves with markers indicating the maximum score and the first wavelength counts where the score drops beyond 0.5% and 1.0% from the maximum.

Table 5.16: Summary of BFS wavelength-count thresholds for CRACK metrics.

Metric	Max score	$n_{\max}$	0.5% drop: n	score	1.0% drop: n	score
CRACK PR–AUC	0.9948	30	11	0.9878	9	0.9790
CRACK F1	0.9854	159	105	0.9800	88	0.9751

5.3.7 Full-image wavelength-selection results

Full-image crack-detection experiments were conducted for several wavelength subsets after complete spatial post-processing, including probability thresholding, morphological filtering, and blob-based aggregation. Performance was evaluated separately for early- and late-season acquisitions on an independent test set.

Table 5.17 reports the obtained CRACK-class metrics for representative wavelength counts.

Table 5.17: Full-image crack-detection performance for different wavelength subsets under early- and late-season conditions.

Season	# Wavelengths	Accuracy	Bal. Acc.	Precision	Recall	F1	MCC
Early	9	0.7833	0.7902	0.8800	0.6875	0.7719	0.5873
Late	9	0.8000	0.8054	0.8667	0.7647	0.8125	0.6054
Early	11	0.7500	0.7567	0.8400	0.6562	0.7368	0.5195
Late	11	0.8167	0.8111	0.8286	0.8529	0.8406	0.6254
Early	30	**0.8167**	**0.8214**	**0.8889**	**0.7500**	**0.8136**	**0.6447**
Late	30	**0.8333**	**0.8348**	**0.8750**	0.8235	0.8485	**0.6652**
Early	159	0.8000	0.8036	0.8571	**0.7500**	0.8000	0.6071
Late	159	**0.8333**	0.8258	0.8333	**0.8824**	**0.8571**	0.6591

Table 5.18: Model Generalization Analysis: Comparison between Calibration and Test MCC.

Season	# Features	MCC Calibration	MCC Test	**Gap (Δ)**
Early	9	0.7381	0.5873	0.1508
	30	0.7232	**0.6447**	**0.0785**
	159	0.6631	0.6071	0.0560
Late	9	0.8968	0.6054	0.2914
	30	0.8771	**0.6652**	**0.2119**
	159	0.8931	0.6591	0.2340

5.3.8 BFS Wavelength-Selection Stability Analysis

To assess the robustness of the Backward Feature Selection (BFS) procedure, the full elimination sequence was repeated five times using different random seeds. In each run, both the train/validation data split seed and the XGBoost model seed were set to the same value (seeds 1–5), so that every repetition uses a different data partition *and* a different model initialisation. The objective metric was CRACK PR–AUC, and elimination

proceeded from 159 down to 1 wavelength. For each repetition, the top-30 and top-11 surviving wavelengths were recorded.

Table 5.19 summarises the per-seed results. Best PR–AUC ranged from 0.812 to 0.998, and the optimal number of features varied between 64 and 153, revealing substantial sensitivity to the data partition.

Table 5.19: BFS stability: per-seed results under varying train/test splits.

Seed	Split Seed	Model Seed	Best PR–AUC	Best n_{features}
1	1	1	0.9014	153
2	2	2	0.9933	66
3	3	3	0.8714	111
4	4	4	0.9979	64
5	5	5	0.8119	149
Mean ± Std			0.915 ± 0.077	108.6 ± 42.2

Wavelength Overlap Across Seeds

Tables 5.20 and 5.21 report the pairwise Jaccard similarity between the wavelength subsets selected by each seed for the top-30 and top-11 configurations, respectively.

Table 5.20: Pairwise Jaccard similarity of top-30 wavelengths across seeds (varying split). Mean Jaccard = 0.493.

	Seed 1	Seed 2	Seed 3	Seed 4	Seed 5
Seed 1	—	0.409	0.590	0.722	0.442
Seed 2	0.409	—	0.378	0.409	0.476
Seed 3	0.590	0.378	—	0.590	0.442
Seed 4	0.722	0.409	0.590	—	0.476
Seed 5	0.442	0.476	0.442	0.476	—

Table 5.21: Pairwise Jaccard similarity of top-11 wavelengths across seeds (varying split). Mean Jaccard = 0.398.

	Seed 1	Seed 2	Seed 3	Seed 4	Seed 5
Seed 1	—	0.500	0.412	0.333	0.412
Seed 2	0.500	—	0.412	0.412	0.333
Seed 3	0.412	0.412	—	0.500	0.333
Seed 4	0.333	0.412	0.500	—	0.333
Seed 5	0.412	0.333	0.333	0.333	—

The top-30 subsets share 13 wavelengths across all five seeds (42%), with a mean Jaccard of 0.493. The top-11 subsets share 4 wavelengths across all five seeds (33%), with

a mean Jaccard of 0.398. The four universally stable wavelengths at the 11-feature level (452.25, 548.55, 580.90, and 729.53 nm) are also part of the 13-wavelength intersection at the 30-feature level, confirming them as the most robust spectral features.

Wavelength Popularity and Stability – Top 30

Figures 5.6–5.9 visualise the selection frequency and stability of the top-30 wavelength subsets across the five seeds.

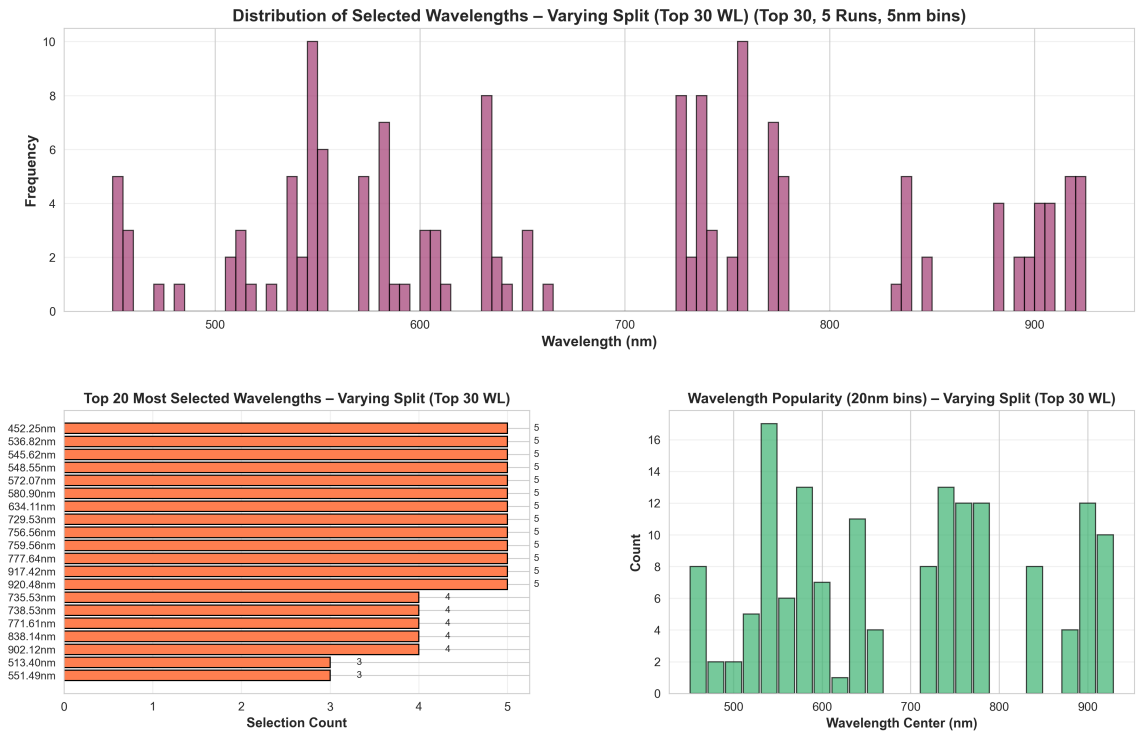


Figure 5.6: Distribution and popularity of the top-30 selected wavelengths across five varying-split seeds.

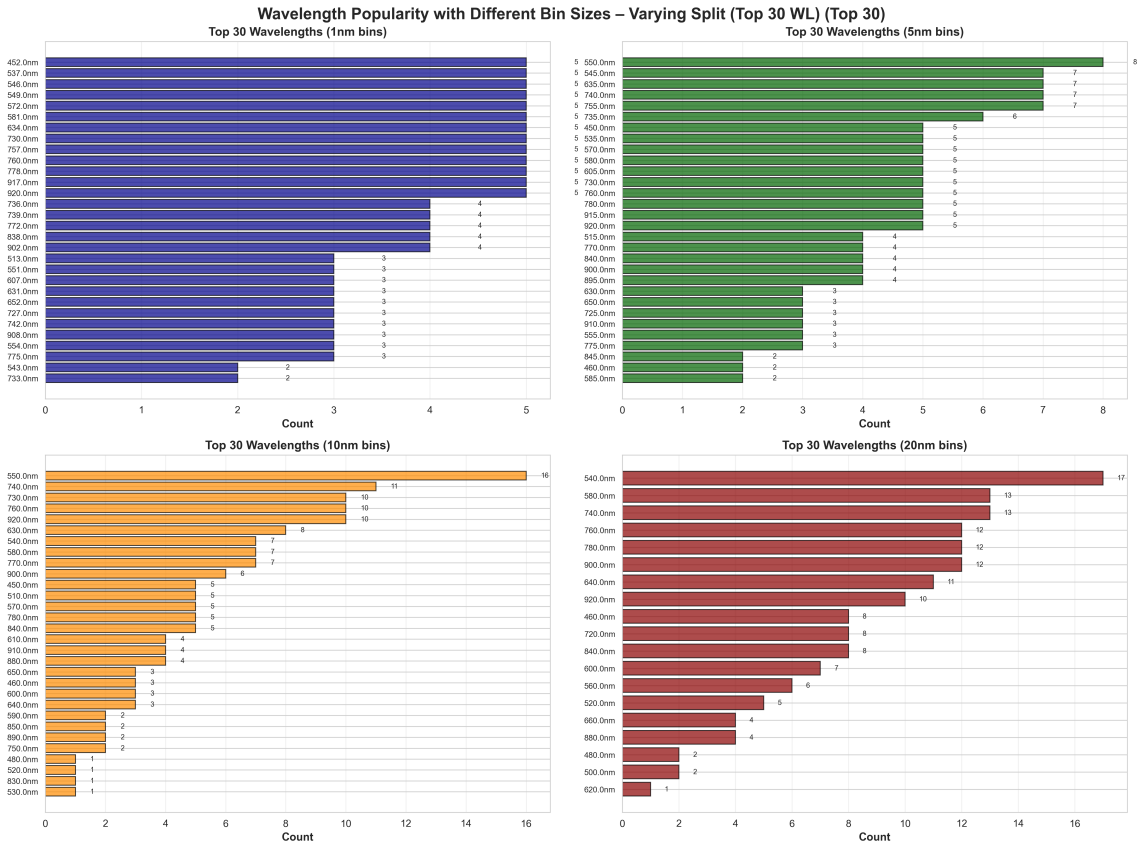


Figure 5.7: Wavelength popularity at different bin sizes (1, 5, 10, 20 nm) for the top-30 subsets.

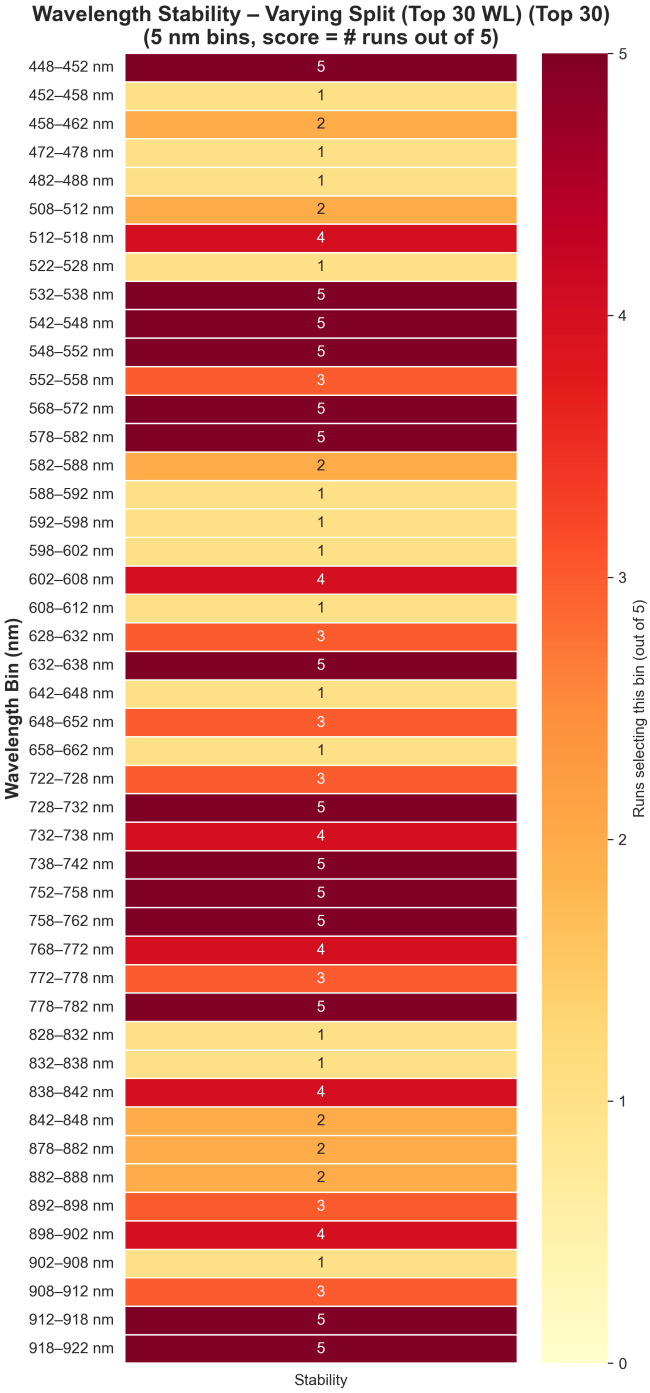


Figure 5.8: Wavelength stability heatmap (5 nm bins) for the top-30 subsets. Each cell shows how many of the five seeds selected a wavelength within that bin.

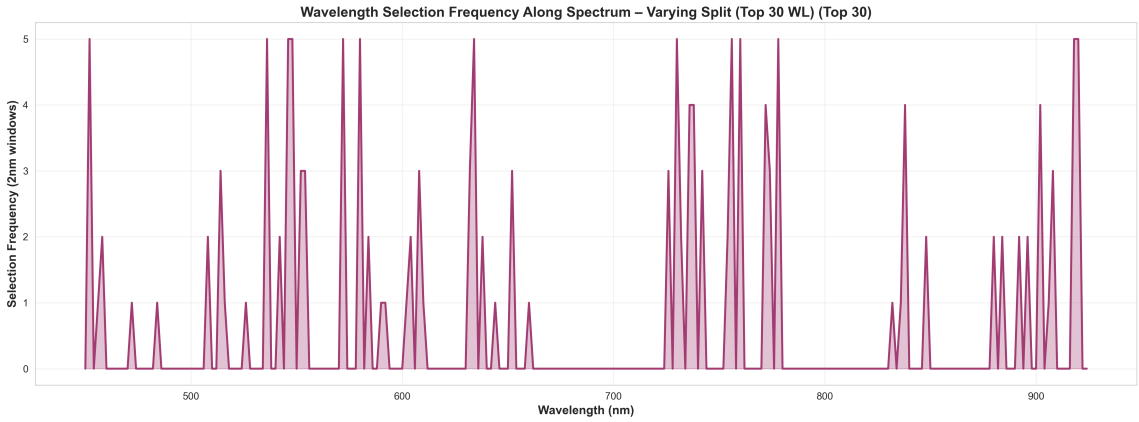


Figure 5.9: Wavelength selection frequency along the spectrum (2 nm sliding window) for the top-30 subsets.

Wavelength Popularity and Stability – Top 11

The same analysis was repeated for the top-11 wavelength subsets (Figures 5.10–5.13), corresponding to the most aggressive dimensionality reduction evaluated (within 1% of the maximum PR–AUC).

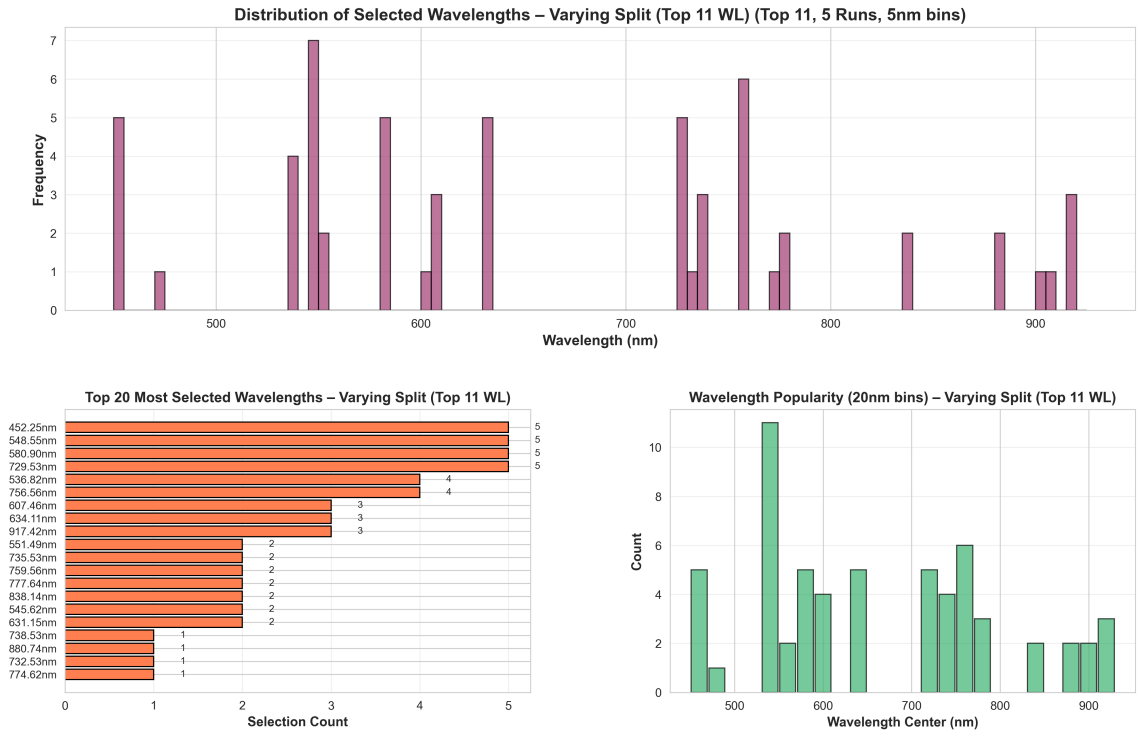


Figure 5.10: Distribution and popularity of the top-11 selected wavelengths across five varying-split seeds.

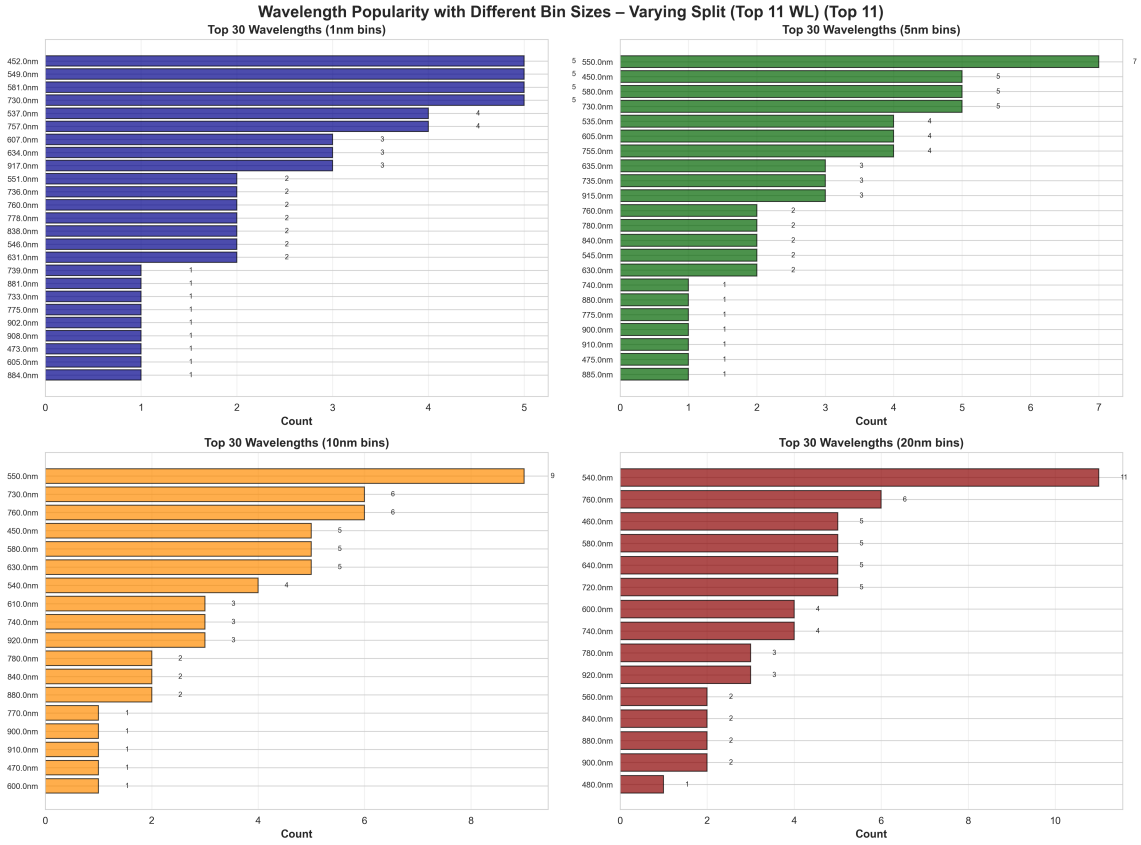


Figure 5.11: Wavelength popularity at different bin sizes (1, 5, 10, 20 nm) for the top-11 subsets.

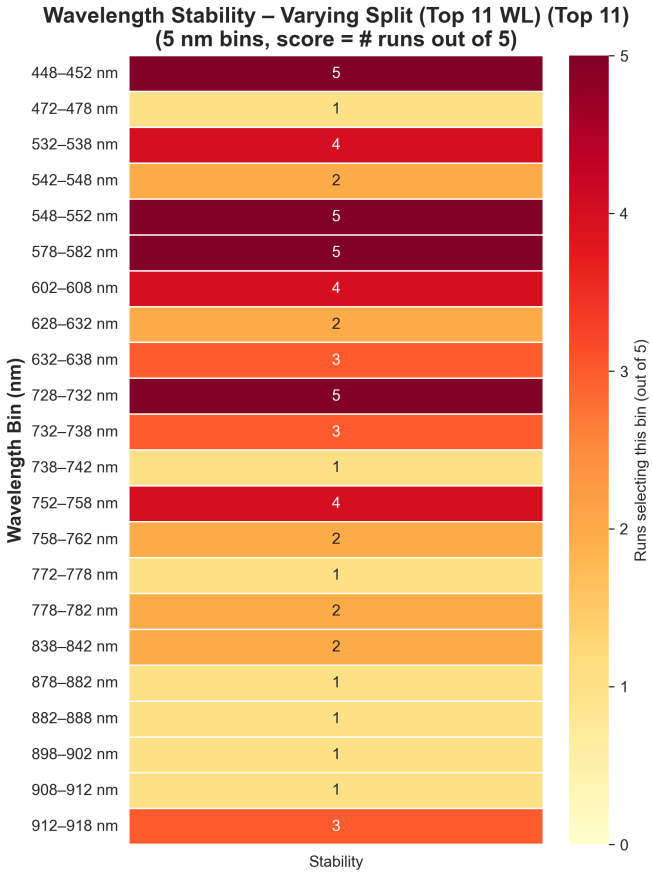


Figure 5.12: Wavelength stability heatmap (5 nm bins) for the top-11 subsets. Each cell shows how many of the five seeds selected a wavelength within that bin.

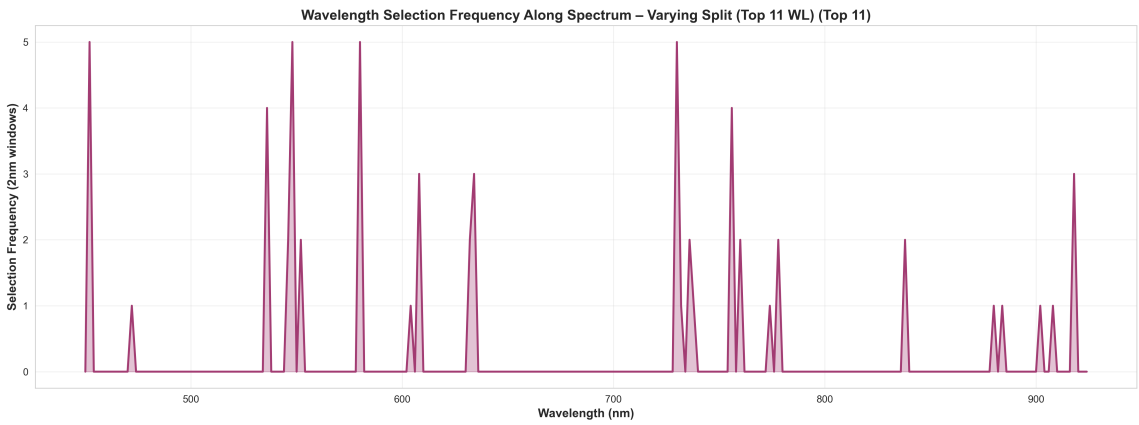


Figure 5.13: Wavelength selection frequency along the spectrum (2 nm sliding window) for the top-11 subsets.

Chapter 6

Discussion

6.1 Summary of Key Findings

This study set out to determine whether hyperspectral imaging in the visible–near-infrared (VIS–NIR) range can detect existing cracks in table grapes and, more ambitiously, predict their occurrence before visual symptoms appear. The experimental programme yielded five principal results that, taken together, advance both the scientific understanding and the practical feasibility of spectral crack detection.

First, the spectral characterisation confirmed that healthy and cracked grape tissues produce markedly different reflectance signatures across the 400–1000 nm range. Healthy tissue exhibited mean VIS reflectance of 0.218 and mean NIR reflectance of 0.660, whereas cracked tissue dropped to 0.137 and 0.450, respectively (Table 5.1). The strongest separability was concentrated in the red-edge and NIR plateau, with Cohen's d reaching 2.69 near 757 nm and a peak Fisher score of 3.63 near 748 nm (Table 5.2). This physical basis validates Hypothesis 1 and establishes that cracking alters the optical properties of berry skin in ways that are measurable with a portable hyperspectral sensor Blackburn, 2007; Y. Zhang et al., 2022.

Second, pixel-level supervised classifiers achieved near-perfect discrimination when trained and tested on manually labelled spectra under the multi-class formulation, with XGBoost reaching accuracy of 0.998 and ROC–AUC of 1.00 (Tables 5.3–5.4), thereby supporting Hypothesis 2. All five classifiers—including PLS-DA, Logistic Regression, SVM, Random Forest, and XGBoost—exceeded 0.99 accuracy under both balanced and unbalanced regimes in the multi-class setting (Tables 5.5–5.6). However, the binary and simplified labelling schemes (crack_regular_rest, crack_vs_rest) that conflate background classes with healthy tissue exhibited substantially lower balanced accuracy (approximately 87–92 %) and higher variance across LOGO folds, foreshadowing the challenges encountered at the whole-image level.

Third, the transition from pixel-level to whole-image inference revealed a meaningful

performance gap. Late-season images, in which cracks are macroscopically visible, yielded a test-set accuracy of 80.0 % and an MCC of 0.598, whereas the early-season scenario—the more practically relevant one—achieved a test accuracy of 73.3 % and an MCC of 0.504 (Table 5.14). These numbers represent a substantial degradation relative to the pixel-level ceiling, yet they also demonstrate that the full-image pipeline retains genuine predictive value, supporting Hypothesis 3 and partially supporting Hypothesis 4.

Fourth, Backward Feature Selection (BFS) showed that the 159-channel hyperspectral cube can be reduced to as few as 11 wavelengths with only 0.5 % loss in CRACK PR–AUC (from 0.9948 to 0.9878), and to 9 wavelengths within 1.0 % of the maximum (Table 5.16). Critically, the 30-wavelength subset matched or exceeded the full spectrum on the independent test set, achieving the highest early-season test MCC of 0.6447 compared with 0.6071 for the full 159-channel cube (Table 5.17). Stability analysis across five random seeds further identified four wavelengths—452.25, 548.55, 580.90, and 729.53 nm—that were selected in every single repetition, confirming their physical importance and supporting Hypothesis 5.

Fifth, the autoencoder-based anomaly-detection branch exposed an instructive asymmetry. Training on cracked pixels alone produced competitive pixel-level performance (accuracy 0.966–0.986, PR–AUC of 0.9999; Table 5.10), whereas training on all non-crack classes failed catastrophically, with near-zero recall and F1 for the cracked class (Table 5.12). At the whole-image level, the autoencoder pipeline achieved a test MCC of only −0.058 in the early-season scenario, compared with 0.504 for the supervised classifier (Tables 5.15 and 5.14).

The sections that follow interpret these findings in the context of grape physiology, sensor design, and operational deployment.

6.2 Spectral and Physiological Interpretation

6.2.1 VIS Region: Pigments and Skin Integrity

The lower mean reflectance of cracked tissue across the entire VIS range (400–700 nm) likely reflects two complementary processes rooted in the biochemistry of berry skin. When the berry cuticle ruptures, the exposed sub-epidermal cells undergo rapid oxidative browning Santos et al., 2023. The resulting accumulation of phenolic oxidation products—including quinones and melanoidins—broadens the absorption across the blue-green region, depressing overall VIS reflectance. Simultaneously, the anthocyanins concentrated in the outer mesocarp—which govern the characteristic red-to-black colouration of 'Scarlotta'—become disordered once the cuticle barrier is breached Castellarin and other, 2016; the loss of the smooth wax layer eliminates the specular component of reflectance and disturbs the internal scattering geometry that normally elevates green-

band reflectance relative to the absorbing flanks of the chlorophyll and anthocyanin peaks Blackburn, 2007.

The fact that the wavelength near 452 nm emerged as one of the four universally stable features (Figure 5.10) corroborates this interpretation: 452 nm falls within the Soret absorption band of chlorophyll and at the short-wavelength edge of anthocyanin absorption Blackburn, 2007. When cracking disrupts the cuticle, enhanced oxidation and pigment leaching alter absorbance at this wavelength more consistently than at neighbouring bands, explaining why BFS repeatedly selects it regardless of the random seed.

Similarly, the stability of the 548 nm and 580 nm features aligns with the green trough and the orange shoulder of anthocyanin reflectance, respectively. These regions are sensitive to anthocyanin concentration, cell-layer organisation, and skin integrity, making them natural indicators of cuticle disruption Blackburn, 2007; Castellarin and other, 2016. Previous work on sweet cherry cracking has reported analogous green-band sensitivity, which was attributed to changes in sub-cuticular scattering following micro-fracture development Peschel and Knoche, 2005; Xu et al., 2025, supporting the generality of these spectral markers across thin-skinned fruit species. La Spada et al. La Spada et al., 2024 further showed that cuticle thickness and quality in table grapes can be modulated by preharvest treatments, reinforcing the link between cuticle integrity and optical properties that underlies the VIS-range discriminability observed here.

6.2.2 NIR and Red-Edge Sensitivity to Water Stress

The most pronounced spectral differences between healthy and cracked tissue were observed in the NIR region, where Cohen's d exceeded 2.5 near 757 nm and the mean absolute reflectance difference reached 0.370 near 727 nm (Table 5.2; Figure 5.2). This large effect size reflects the dominant role of cellular water content and tissue structure in shaping NIR reflectance Thenkarail et al., 2016; Y. Zhang et al., 2022. In intact berries, the mesocarp acts as a powerful diffuse reflector: multiple internal scattering events at air–cell-wall interfaces elevate reflectance above 700 nm, while liquid water in the symplast introduces broad absorption features centred around 760 nm and 970 nm T. Wang and Zhao, 2024. Once cracking occurs, mesocarp cells lose turgor through rapid dehydration and juice leakage, collapsing the scattering interfaces and dramatically reducing NIR reflectance Chang et al., 2019.

The red-edge—the steep rise in reflectance between 680 nm and 740 nm—is particularly informative because it is jointly governed by chlorophyll absorption (which sets the short-wavelength foot) and internal scattering (which sets the shoulder) Blackburn, 2007; Thenkarail et al., 2016. In cracked berries, both contributions are diminished: chlorophyll degrades in exposed tissue and the scattering plateau drops, causing the red-edge to shift toward shorter wavelengths and to flatten. The fourth stable wavelength at 729 nm

sits precisely on this inflection, where the slope and amplitude of the red-edge are most sensitive to simultaneous changes in pigment content and tissue water (Figure 5.9).

These observations align with established water-stress physiology in grape berries. Chang and Keller Chang and Keller, 2021; Chang et al., 2019 demonstrated that turgor pressure dynamics and cuticle permeability modulate berry water relations during ripening, with micro-cracks forming days before macroscopic splitting. Yu et al. Yu et al., 2025 recently showed that aquaporin-mediated water uptake plays a pivotal role in grape berry cracking, with the cracking-susceptible cultivar 'Xiangfei' exhibiting significantly higher mesocarp water content than the resistant 'Zuijinxiang' during the critical cracking period. The spectral changes detected in this study—particularly the depressed NIR plateau and blue-shifted red-edge—can therefore be understood as the optical fingerprint of the biomechanical failure pathway that culminates in visible cracking. This interpretation is consistent with the broader VIS–NIR sensing literature showing that pre-symptomatic stress in plant tissue manifests as altered internal scattering and pigment absorption before macroscopic symptoms appear **mahlein2016plant**; Liu et al., 2023; Panda and Mishra, 2025.

6.3 Why Pixel-Level Detection Works in Controlled Settings but Fails in Field Conditions

The near-perfect pixel-level accuracy reported in Chapter 5 should be interpreted with caution, as it reflects favourable conditions that do not fully represent operational vineyard imaging. Under the multi-class labelling scheme, in which background categories (leaf, wood, soil, plastic) are explicitly modelled, the classifier benefits from a well-structured feature space in which each class occupies a distinct spectral region. The resulting decision boundaries are clean and generalise well across LOGO cross-validation folds drawn from different grape clusters, as evidenced by the low standard deviations across folds for XGBoost (accuracy 0.998 ± 0.002; Table 5.5).

When the same classifiers are deployed on whole-image hyperspectral scenes, however, several confounding factors degrade performance. First, the spatial resolution of the Specim IQ (approximately 1.08 mm per pixel at 1 m distance) Behmann et al., 2018 means that many image pixels are *mixed*: they straddle the boundary between berry and leaf, or between crack and intact skin. Mixed pixels produce hybrid spectral signatures that do not match any of the pure-class prototypes used during training, leading to ambiguous or erroneous predictions Bioucas-Dias et al., 2013. Second, the illumination geometry in the field is highly variable. Although a white reference panel was included in every scene, the hemispherical reflectance from a curved berry surface differs from the Lambertian reference geometry, introducing shading gradients, bidirectional reflectance

distribution function (BRDF) effects, and specular highlights that shift the apparent spectrum Mishra et al., 2017. Third, the presence of thin surface films—water droplets, dust, spray residues—modifies reflectance in ways that are unaccounted for in the pixel-level training set. These field-specific artefacts are well documented in the close-range hyperspectral literature Mishra et al., 2017; Ram et al., 2024 and explain why laboratory-to-field performance degradation is a recurring challenge in agricultural HSI applications García-Vera et al., 2024.

The binary and simplified labelling experiments (crack_regular_rest, crack_vs_rest) anticipated this problem. Under these formulations, the classifier must separate cracked tissue from a heterogeneous "rest" category that encompasses every non-crack pixel, including spectrally diverse background materials. The resulting balanced accuracy of approximately 87–92 % (Table 5.5) still represents good performance for a purely spectral (non-spatial) model, yet the higher variance across folds and the lower precision for the cracked class signal that the decision boundary is less stable when background diversity increases. Notably, PLS-DA suffered the most under these simplified schemes, achieving balanced accuracy of only 0.688–0.706 (Table 5.5), whereas tree-based methods (Random Forest, XGBoost) maintained balanced accuracy above 0.91, suggesting that non-linear classifiers are more resilient to spectral background heterogeneity.

These observations lead to a clear methodological implication: pixel-level spectral classifiers, however accurate in controlled conditions, require an intermediate spatial processing stage—such as segmentation, morphological filtering, or patch-level aggregation—before they can be deployed reliably in uncontrolled vineyard imagery. This finding echoes the broader HSI literature, where spectral–spatial integration strategies have been shown to significantly improve classification robustness in scenes with complex backgrounds Bioucas-Dias et al., 2013; Z. Chen et al., 2024. The full-image pipeline evaluated in this study incorporated exactly such steps, and while the resulting performance did not match the pixel-level ceiling, it demonstrates a viable pathway from laboratory accuracy to field applicability.

6.4 Temporal Aggregation as a Robust Strategy for Early Crack Prediction

One of the most practically significant findings of this study is that the full-image pipeline, which aggregates pixel-level predictions through morphological filtering and spatial thresholding, retains predictive value even in the early-season scenario. Early-season images were acquired before any visually discernible cracking had occurred, yet the pipeline achieved a test accuracy of 73.3 % and a recall of 59.4 % for the cracked class (Table 5.14). While this performance falls short of the late-season benchmark (accuracy

80.0 %, recall 79.4 %), it demonstrates that spectral precursors of cracking are measurable weeks before macroscopic symptoms manifest.

This finding can be understood through the physiological lens developed in the preceding section. Cracking is not an instantaneous event but rather the terminal stage in a progressive sequence: beginning with cuticular micro-fracturing Chang and Keller, 2021; Peschel and Knoche, 2005, continuing through changes in water permeability and turgor dynamics Chang et al., 2019; Yu et al., 2025, and culminating in macroscopic splitting Khadivi-Khub, 2014; Santos et al., 2023. The spectral indicators of the intermediate stages should therefore be detectable before the final mechanical failure. The depressed NIR reflectance and altered red-edge slope observed in the pixel-level analysis (Figure 5.1; Table 5.1) are precisely the optical signatures of the cellular-scale changes that precede visible cracking: loss of internal turgor, reduced scattering, and incipient pigment degradation. This interpretation aligns with recent findings by Dong et al. Dong et al., 2025, who demonstrated that hyperspectral imaging can detect apple bitter pit—another turgor-related physiological disorder—before visual symptoms emerge, and with Haghbin et al. Haghbin et al., 2023, who showed pre-symptomatic detection of grey mould in kiwifruit using VIS–NIR chemometrics.

Temporal aggregation addresses a fundamental limitation of single-pixel classification. Individual pixel predictions are inherently noisy: a single mis-classified pixel in a cluster of several thousand has negligible practical consequence, but a cluster of false positives near a berry edge can generate a spurious crack flag. By aggregating predictions into patch-level and then image-level statistics—and by requiring that the proportion of flagged crack pixels exceeds a calibrated threshold—the pipeline effectively smooths over pixel-level noise and yields a robust cluster-level score Bioucas-Dias et al., 2013.

The experimental design exploited this insight by using Optuna to jointly optimise the probability threshold, morphological kernel sizes, and the global crack-percentage criterion on the calibration split. The resulting pipeline achieved substantially higher precision (0.864 on the early-season test set) than a naive pixel-level approach would, precisely because the spatial aggregation enforces spatial coherence on the predictions. That the early-season pipeline achieved higher precision (0.864) than recall (0.594) on the test split (Table 5.14) indicates a conservative decision boundary, which is appropriate for an early-warning system where false alarms carry higher operational costs than missed detections in the initial screening pass. This paradigm—pixel-level classification followed by spatially informed aggregation—constitutes the core operational strategy recommended by this study for field deployment and aligns with the broader consensus in agricultural HSI that spatial context is critical for robust classification Z. Chen et al., 2024; Ram et al., 2024.

6.5 Feature Selection and Spectral Stability

6.5.1 Dimensionality Reduction via BFS

The BFS elimination sequence revealed a highly non-linear relationship between the number of retained wavelengths and detection performance (Figure 5.4). CRACK PR–AUC remained within 0.5 % of its maximum (0.9948) down to 11 wavelengths, and within 1.0 % down to 9 wavelengths (Table 5.16). These results indicate that the vast majority of the 159 usable channels carry redundant spectral information for the specific task of crack detection, and that the discriminative signal is concentrated in a small number of spectral regions. This high redundancy is a well-known property of contiguous hyperspectral data, where adjacent bands are often correlated at $r > 0.99$ Bioucas-Dias et al., 2013; Liu et al., 2022.

The practical implication is substantial. A full hyperspectral line-scan camera typically costs between $30,000 and $100,000, requires specialised data handling, and generates datasets measured in hundreds of gigabytes per field campaign **BHARGAVA2024e33208**. A multispectral sensor configured with 9–30 narrowband filters can be manufactured for a fraction of this cost, operates at higher frame rates, and produces data volumes that are orders of magnitude smaller Benelli et al., 2020. The BFS results provide a quantitative justification for the transition from hyperspectral research prototype to multispectral operational sensor, echoing the band-reduction strategies that have proven effective in other agricultural HSI applications Khaled et al., 2023; Liu et al., 2022; Z. Zhang et al., 2024.

It is noteworthy that the CRACK F1 metric exhibited a different optimisation landscape from PR–AUC, with its maximum achieved only at 159 wavelengths and a more gradual decline as features were removed (Figure 5.5). This discrepancy reflects the fact that F1 penalises both false positives and false negatives equally, and that finer spectral resolution helps to separate crack signatures from spectrally similar background materials. For early-warning applications in which sensitivity (recall) is prioritised over specificity, the PR–AUC-oriented subset is the more appropriate choice; for applications demanding high precision, a larger wavelength subset may be warranted.

6.5.2 Stability Across Random Seeds and Sensor Design Implications

The stability analysis across five random seeds exposed a methodologically important finding: the *identity* of the selected wavelengths is moderately sensitive to the data partition, even when the *number and performance* of the selected bands remain stable (Table 5.19). The best CRACK PR–AUC ranged from 0.812 to 0.998 across seeds, and the optimal number of features varied between 64 and 153, revealing substantial sensitivity to the calibration–validation partition. Mean pairwise Jaccard similarity was 0.493 for

the top-30 subsets and 0.398 for the top-11 subsets (Tables 5.20–5.21), indicating that roughly half of the selected wavelengths change from one random split to another.

This level of variability is not unexpected in wrapper-based feature selection on high-dimensional, correlated data Liu et al., 2022: when adjacent spectral channels are highly collinear, BFS can substitute one for another without measurable loss, producing different subsets that are functionally equivalent. Similar instability has been reported in filter-based hyperspectral band selection studies, where mutual-information and Fisher-score methods often yield non-unique solutions when inter-band correlation is high **gu2012generalized**; **vergara2014review**; Ding et al., 2023. The key finding is that certain wavelengths resist this substitution entirely. The four universally stable wavelengths—452.25, 548.55, 580.90, and 729.53 nm—were retained by every seed at both the 11-feature and 30-feature levels (Figures 5.10–5.8).

These four wavelengths anchor the sensor design recommendation of this study. Their spectral locations correspond to physiologically interpretable regions: the Soret band (452 nm), the green reflectance trough (548 nm), the anthocyanin shoulder (580 nm), and the red-edge inflection (729 nm), as discussed in Section 6. A multispectral sensor incorporating these four channels, supplemented by additional bands drawn from the broader intersection set (13 common wavelengths at the 30-feature level), would capture the essential information for crack detection while minimising hardware complexity. This approach is consistent with the emerging trend in precision agriculture toward application-specific multispectral sensors that target physiologically meaningful bands rather than employing full hyperspectral coverage Benelli et al., 2020; Ram et al., 2024.

An important caveat is that this stability analysis was conducted on a single cultivar ('Scarlotta') in a single growing region (Lachish, Israel). Other cultivars with different anthocyanin profiles, cuticle thickness, or mesocarp water content may shift the optimal wavelength positions Chang and Keller, 2021; Khadivi-Khub, 2014. Cross-cultivar validation experiments would be needed before recommending a fixed filter configuration for commercial deployment.

6.6 Generalization Gaps Between Calibration and Test Sets

Across every experimental configuration examined in this study, test-set performance was consistently lower than calibration-set performance, revealing a generalization gap that warrants careful interpretation (Table 5.18). The gap was particularly pronounced in the late-season scenario, where the MCC dropped from 0.877–0.897 on the calibration split to 0.605–0.665 on the test split, representing a reduction of more than 20 percentage points. The early-season gap was narrower in absolute terms (approximately 6–15 percentage

points), but the baseline performance was also lower, meaning that the test-set MCC hovered between 0.587 and 0.645.

Several factors contribute to this pattern. The most fundamental is the difference in data provenance between the calibration and test sets. By design, calibration images were drawn from Row 1 of the vineyard, while test images came from Row 2 (Table 5.13). Although both rows belonged to the same experimental vineyard and experienced the same irrigation treatments, they differed in vine age uniformity, canopy architecture, and micro-environmental conditions such as wind exposure and shading. These factors introduce a spatial domain shift—a systematic difference in the joint distribution of spectral features and labels—that is invisible during cross-validation within a single row but manifests as a performance drop when the model is applied to a new spatial domain. Such spatial domain shifts are increasingly recognised as a major challenge in agricultural remote sensing, where models trained on one field or plot often degrade when transferred to another, even within the same farm Fass et al., 2025; Ram et al., 2024. Kaltenborn et al. Kaltenborn et al., 2025 and Li et al. Li et al., 2023 have emphasised that proper data splitting is essential to avoid information leakage in hyperspectral classification, and the row-based spatial split used in this study represents a conservative strategy that exposes rather than conceals the generalization challenge **kaufman2012leakage**.

The interaction between the number of retained wavelengths and the generalization gap adds nuance (Table 5.18). For the early-season scenario, the 30-wavelength subset produced the highest test MCC (0.6447) and a moderate gap (0.0785), whereas the 9-wavelength configuration yielded a larger gap (0.1508) and the full 159-wavelength cube yielded a gap of 0.0560 but a lower absolute test MCC (0.6071). This non-monotonic pattern suggests that an intermediate level of spectral dimensionality strikes a balance between under-fitting (too few features to capture the signal) and over-fitting (too many features that encode row-specific noise). The 30-wavelength configuration may therefore represent a practically optimal compromise between predictive power and generalization stability, an observation consistent with well-known bias–variance trade-off principles **menze2009comparison**.

For the late-season scenario, the generalization gaps were uniformly larger (0.21–0.29; Table 5.18), likely because the higher calibration MCC reflects over-optimisation of the post-processing thresholds on the more separable late-season data. When these thresholds are applied to the test row, the mismatch between the calibration and test spectral distributions becomes more consequential.

For future operational systems, these generalization gaps have direct implications. Models calibrated on one section of a vineyard—or one season—should not be deployed without adaptation to a new section or a new season. Transfer-learning strategies, domain-adaptation techniques, or ensemble approaches that incorporate data from multiple rows and time points could mitigate this limitation and are recommended as a priority

for future research.

6.7 Supervised Classification versus Anomaly Detection

A deliberate objective of this study was to compare supervised and unsupervised approaches to crack detection, both at the pixel level and within the full-image pipeline. The comparison yielded a clear and instructive conclusion: supervised classifiers consistently outperformed autoencoder-based anomaly detection, and the margin of superiority widened as the task complexity increased.

At the pixel level, the autoencoder trained exclusively on cracked pixels performed surprisingly well in the multi-class setting, achieving accuracy above 0.96 and PR–AUC of 0.9999 (Table 5.10). This success can be attributed to the narrow spectral manifold learned by the model: crack signatures occupy a well-defined region of the spectral space, and all other classes (leaf, wood, soil, plastic, healthy tissue) produce larger reconstruction errors precisely because they are spectrally dissimilar. This finding is consistent with the spectral separability analysis (Table 5.2), which showed that cracked tissue occupies a distinct low-reflectance region in both VIS and NIR, well separated from the other vineyard materials.

The reverse configuration—training on all classes except cracked and flagging high reconstruction error as anomalous—failed almost completely, with cracked-class recall of 0.000 under balanced sampling and 0.011–0.111 under unbalanced sampling (Table 5.12). This failure is not incidental; it reflects a fundamental property of the spectral space. Because the non-crack classes span a much wider spectral range (from soil and plastic in the VIS to leaf and wood in the NIR), the autoencoder learns a reconstruction manifold that is broad enough to accommodate crack spectra as well. The reconstruction error for crack pixels is therefore no larger than for normal tissue, eliminating the anomaly-detection signal. This finding aligns with the theoretical analysis of anomaly detection in high-dimensional spaces: when the normal class is heterogeneous, the learned representation often generalises to anomalous samples that fall within the convex hull of the training manifold **gewali2022machine**.

At the whole-image level, the autoencoder-based pipeline achieved a test MCC of only -0.058 in the early-season scenario (Table 5.15), indicating performance indistinguishable from—or slightly worse than—chance. Even in the late-season scenario, where cracks are macroscopically visible, the autoencoder pipeline underperformed the supervised classifier by approximately 22 MCC points (0.375 vs. 0.598; cf. Tables 5.15 and 5.14). The degradation from pixel to whole-image level was substantially more severe for the autoencoder than for the supervised model, because the autoencoder's pseudo-

probability scores—derived from sigmoid-transformed reconstruction errors—are less calibrated than the true posterior probabilities produced by the supervised XGBoost classifier **chen2016xgboost**. When these poorly calibrated scores are propagated through the same spatial post-processing pipeline, the resulting image-level decisions become unreliable.

These results lead to a clear recommendation: for crack detection in table grapes, the supervised classification paradigm is strongly preferred. Anomaly detection approaches may remain relevant for applications in which labelled training data are genuinely unavailable, but in the context of this study—where SAM2-assisted annotation Ravi et al., 2024 made large-scale pixel-level labelling feasible—the investment in labelled data pays dividends in detection accuracy.

An additional insight emerges from the asymmetry between the two autoencoder training strategies. The crack-trained autoencoder's success suggests that cracked tissue possesses a coherent, low-dimensional spectral signature that is well-separated from other vineyard materials. This property is itself a validation of the spectral distinguishability hypothesis (Hypothesis 1) and reinforces the argument that cracking induces a characteristic and physically interpretable spectral alteration.

6.8 Practical Implications for Precision Viticulture

The results of this study translate into several actionable recommendations for the grape industry, situated within the broader context of precision agriculture and data-driven crop management Mgendi, 2024; Soussi et al., 2024.

The first and most immediate implication is that hyperspectral imaging, even in its current portable configuration (Specim IQ; Behmann et al., 2018), can provide a reliable early-warning signal for cracking susceptibility. A grower who systematically images grape clusters during the early ripening period and processes them through the temporal aggregation pipeline described in this study would gain a lead time of several weeks relative to visual scouting. This lead time could be exploited to adjust irrigation scheduling—a well-established lever for managing berry turgor and cracking risk Chang et al., 2019; Khadivi-Khub, 2014—apply targeted calcium or sorbitol-based spray treatments Guirao et al., 2025; La Spada et al., 2024, or advance the harvest of at-risk sections. The economic significance of such advance warning is considerable: cracking causes not only direct yield losses but also facilitates secondary fungal infection, particularly sour rot, which further reduces marketability **OWOYEMI2024113013**; Hall et al., 2018.

The second implication concerns sensor cost. The BFS analysis demonstrates that 9–30 carefully chosen wavelengths capture essentially the same information as the full 159-channel spectrum (Table 5.16). This finding opens the door to custom multispectral cameras—potentially mounted on unmanned aerial vehicles (UAVs) or tractor-mounted

platforms—that combine high throughput with low per-unit cost Benelli et al., 2020; Ram et al., 2024. The four universally stable wavelengths identified by the stability analysis (452, 548, 580, 729 nm) provide a concrete starting point for filter specification. This transition from hyperspectral to multispectral is an active frontier in precision agriculture, where application-specific band selection has been shown to yield cost-effective monitoring solutions without sacrificing predictive performance Fass et al., 2025; Khaled et al., 2023. The CrackSense project Alchanatis et al., 2022, a Horizon Europe initiative focused on real-time monitoring and prediction of fruit cracking, shares this vision of scalable sensing for cracking management.

The third implication relates to the full-image pipeline architecture. The study demonstrates that pixel-level classifiers, however accurate in isolation, require spatial post-processing (morphological filtering, patch aggregation, threshold optimisation) to translate spectral predictions into actionable vineyard-level decisions. This multi-stage architecture—pixel classification, spatial aggregation, and calibrated thresholding—constitutes a deployable decision-support system that can be embedded in existing vineyard management platforms Ahmed and Shakoor, 2025; Krisnawijaya et al., 2022.

Finally, the observed sensitivity to water-related spectral features suggests that hyperspectral monitoring could serve a dual purpose: detecting cracking risk and simultaneously assessing vine water status. These two applications are physiologically linked through berry turgor and cuticle permeability Chang et al., 2019; Yu et al., 2025, and a single imaging survey could inform both irrigation management and harvest-timing decisions. This synergy between cracking prediction and water-status assessment represents a compelling value proposition for growers considering investment in spectral sensing infrastructure.

6.9 Limitations of the Study

Despite the breadth of the experimental programme, several limitations should be acknowledged.

The study was conducted on a single cultivar ('Scarlotta Seedless') in a single vineyard and over a single growing season. While 'Scarlotta' is commercially important and known for its cracking susceptibility **OWOYEMI2024113013**, the spectral markers identified here may not transfer directly to cultivars with different anthocyanin profiles, cuticle thicknesses, or berry architectures Chang and Keller, 2021; Khadivi-Khub, 2014. A multi-cultivar, multi-season validation is essential before the results can be generalised.

The spatial resolution of the Specim IQ (approximately 1.08 mm per pixel at 1 m working distance; Behmann et al., 2018) imposes a lower bound on the size of detectable cracks. Micro-fissures smaller than two pixels may not produce a sufficiently distinct spectral signal, and the mixed-pixel problem at crack boundaries introduces spectral

ambiguity Bioucas-Dias et al., 2013. Higher-resolution pushbroom sensors, or close-range macro-hyperspectral configurations Mishra et al., 2017, could address this limitation in future work.

The calibration–test split followed a spatial design in which one vineyard row served for calibration and another for testing. While this design tests spatial generalization within the same vineyard and season—and is arguably more rigorous than random splitting **kaufman2012leakage**; Kaltenborn et al., 2025—it does not test temporal generalization across seasons. The spectral baseline of healthy tissue, the timing and severity of cracking, and the environmental forcing variables all differ between years, and the models developed here have not been validated under these conditions.

The autoencoder anomaly-detection experiments used relatively simple fully connected architectures (three variants with 64–32–16, 128–64–32, and 64–32–8 bottleneck widths). More expressive architectures—such as variational autoencoders, convolutional autoencoders operating on spectral patches, or normalizing-flow models—might achieve better separation between normal and anomalous reconstruction errors, particularly in the challenging non-crack training configuration where all three architectures tested in this study failed to detect any cracked pixels under balanced sampling (Table 5.12).

The whole-image pipeline relied on post-processing parameters (probability thresholds, morphological kernel sizes, crack-percentage criteria) that were optimised on the calibration split using Optuna. Although these parameters generalised to the test split with reasonable performance, they were tuned on a relatively small calibration set (60–172 images; Table 5.13), and their stability under larger and more diverse datasets has not been assessed.

Finally, the study did not incorporate complementary data modalities. Thermal imaging, which was acquired in parallel during the field campaign, was not included in the modelling pipeline. The potential synergy between spectral and thermal features—the former sensitive to pigment and water content, the latter to evaporative cooling and tissue necrosis—remains an unexplored avenue that could further improve detection accuracy and robustness **mahlein2016plant**.

6.10 Future Research Directions

The findings and limitations of this study suggest several concrete directions for future investigation, ordered by scientific priority and practical impact.

The highest priority is multi-season validation. The models should be re-trained and tested on data from at least two additional growing seasons to establish whether the spectral markers identified here are temporally stable and whether the generalization gap between calibration and test sets can be reduced through temporal domain adaptation. Given that environmental forcing, vine physiology, and cracking incidence vary inter-

annually Chang et al., 2019; Khadivi-Khub, 2014, temporal robustness cannot be assumed from a single-season study.

A second priority is cross-cultivar generalisation. Experiments on cultivars with contrasting skin properties—thin-cuticle varieties prone to cracking versus thick-cuticle varieties that are resistant—would reveal whether the four stable wavelengths have universal relevance or are cultivar-specific. Santos et al. Santos et al., 2023 reviewed the molecular mechanisms underlying fruit cracking across species, and cross-cultivar spectral studies would establish whether common physiological pathways produce transferable spectral signatures.

Third, the integration of spatial models deserves attention. The current pipeline treats each pixel independently and relies on post-hoc morphological filtering for spatial coherence. Convolutional or transformer-based architectures that jointly reason over spatial neighbourhoods and spectral channels Z. Chen et al., 2024; Dhanya et al., 2022 could learn to distinguish true crack boundaries from background edges, potentially eliminating the need for hand-tuned post-processing parameters and improving robustness under variable field conditions.

Fourth, the complementary thermal data acquired during this study offer an opportunity for multi-modal fusion. A combined spectral-thermal model could exploit the distinct failure signatures captured by each modality: spectral changes in reflectance associated with pigment and water loss, and thermal anomalies associated with altered evaporative cooling at crack sites **mahlein2016plant**.

Fifth, the BFS-derived wavelength subsets should be validated using a physical multispectral sensor rather than by computationally subsetting the hyperspectral cube. A true hardware validation would account for differences in filter bandwidth, signal-to-noise ratio, and radiometric calibration that are abstracted away in a simulation-based feature selection Benelli et al., 2020. This validation step is essential before recommending the four-wavelength configuration for commercial sensor manufacturing.

Sixth, the autoencoder-based anomaly detection framework could be revisited with more expressive generative models. Variational autoencoders with spectral-specific priors, normalizing flows, or diffusion-based models could provide sharper reconstruction boundaries and improved anomaly scoring, particularly in the non-crack training configuration that failed comprehensively in the present study.

Finally, the operational deployment of the pipeline on UAV-mounted or tractor-mounted multispectral platforms should be piloted. Such a deployment would test the end-to-end system under realistic throughput, altitude, and illumination conditions, and would provide the evidence base needed for commercial adoption by the viticulture industry Alchanatis et al., 2022; Mgendi, 2024.

6.11 Concluding Perspective

This study began with a straightforward question: can hyperspectral imaging detect grape cracking? The answer—emphatically supported at the pixel level and cautiously affirmed at the whole-image level—opens a broader vista.

The spectral markers identified here are not arbitrary correlates; they reflect the physical and physiological processes that precede and accompany cuticle failure: pigment degradation Blackburn, 2007; Castellarin and other, 2016, water loss Chang et al., 2019; Yu et al., 2025, altered scattering, and red-edge suppression Thenkarail et al., 2016. The fact that these markers are detectable before macroscopic cracking appears transforms the paradigm from reactive detection to proactive prediction—a shift with immediate consequences for vineyard management.

The pathway from a 159-channel research camera to a compact multispectral sensor with fewer than a dozen filters is now quantitatively justified by the BFS and stability analyses (Tables 5.16–5.21; Figures 5.10–5.9). The four universally stable wavelengths provide a minimal spectral fingerprint that is both physiologically grounded and computationally robust, offering a bridge between laboratory science and vineyard operations.

At the same time, the generalization gaps documented in this study (Table 5.18) serve as a necessary reminder that spectral models, like all empirical tools, require continued validation as cultivars, seasons, and management practices change García-Vera et al., 2024; Ram et al., 2024. The contribution of the present work lies not only in the specific models and wavelengths reported, but in the methodological framework—from pixel-level spectral analysis through spatially informed aggregation to stability-guided sensor design—that can be adapted and extended as new data become available.

In the broader context of precision viticulture, this research demonstrates that early, non-destructive crop monitoring using spectral sensing is not merely a theoretical possibility but an operational strategy within reach Alchanatis et al., 2022; Mgendi, 2024. Closing the remaining gaps—temporal validation, cross-cultivar testing, multi-modal fusion, and hardware integration—will bring this strategy from the research vineyard to the commercial field.

Bibliography

Abade, A., et al. (2025). From spectrum to yield: Advances in crop photosynthesis with hyperspectral imaging. *Photosynthetica*, *63*(2). https://doi.org/10.32615/ps.2025.012

Ahmed, N., & Shakoor, N. (2025). Advancing agriculture through iot, big data, and ai: A review of smart technologies enabling sustainability. *Smart Agricultural Technology*, *10*, 100848. https://doi.org/https://doi.org/10.1016/j.atech.2025.100848

Akter, T., Faqeerzada, M. A., Kim, Y., Pahlawan, M. F. R., Aline, U., Kim, H., Kim, H., & Cho, B.-K. (2025). Hyperspectral imaging with multivariate analysis for detection of exterior flaws for quality evaluation of apples and pears. *Postharvest Biology and Technology*, *223*, 113453. https://doi.org/10.1016/j.postharvbio.2025.113453

Alchanatis, V., et al. (2022). *High throughput real-time monitoring and prediction of fruit cracking by utilising and upscaling sensing and digital data technologies* (tech. rep.) (Research and Innovation Action (RIA), HORIZON-CL6-2022-GOVERNANCE-01-11). Horizon Europe.

Attri, I., Awasthi, L. K., Sharma, T. P., & Rathee, P. (2023a). A review of deep learning techniques used in agriculture. *Ecological Informatics*, *77*, 102217.

Attri, I., Awasthi, L. K., Sharma, T. P., & Rathee, P. (2023b). A review of deep learning techniques used in agriculture. *Ecological Informatics*, *77*, 102217. https://doi.org/https://doi.org/10.1016/j.ecoinf.2023.102217

Author, A., et al. (2022). Assessment of scarlotta seedless table grape maturation and quality. *Foods*, *11*(5), 663.

Author, B., et al. (2024). Impact of storage temperature on postharvest quality of scarlotta seedless. *Postharvest Biology and Technology*, *215*, 113013.

Ball, S. R. (2007). *Field guide to spectroscopy*. SPIE Press.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, *43*(5), 772–777.

Behmann, J., Acebron, K., Emin, D., Bennertz, S., Matsubara, S., Thomas, S., Bohnenkamp, D., Kuska, M. T., Jussila, J., Salo, H., et al. (2018). Specim iq: Evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection. *Sensors*, *18*(2), 441.

Benelli, A., Cevoli, C., Fabbri, A., & Ragni, L. (2020). Hyperspectral imaging and mini-video devices for non-destructive analysis of fruit quality: A review. *Biosystems Engineering*, *192*, 142–152. https://doi.org/10.1016/j.biosystemseng.2020.01.018

Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine learning in agriculture: A comprehensive updated review. *Sensors*, *21*(11), 3758.

Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., et al. (2013). Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine*, *1*(2), 6–36.

Blackburn, G. A. (2007). Hyperspectral remote sensing of plant pigments. *Journal of experimental botany*, *58*(4), 855–867.

Cao, Z., Sun, S., & Bao, X. (2025). A review of computer vision and deep learning applications in crop growth management. *Applied Sciences*, *15*(15), 8438.

Castellarin, S. D., & other. (2016). Characterization of major ripening events during softening in grape and their relationship with growth. *Journal of Experimental Botany*, *67*(3), 709–722.

Chang, B.-M., & Keller, M. (2021). Cuticle and skin cell walls have common and unique roles in grape berry splitting. *Horticulture Research*, *8*, 168. https://doi.org/10.1038/s41438-021-00602-2

Chang, B.-M., Zhang, Y., & Keller, M. (2019). Softening at the onset of grape ripening alters fruit rheological properties and decreases splitting resistance. *Planta*, *250*, 1293–1305. https://doi.org/10.1007/s00425-019-03226-y

Chen, L., Wu, Y., Yang, N., & Sun, Z. (2025). Advances in hyperspectral and diffraction imaging for agricultural applications. *Agriculture*, *15*.

Chen, Z., Hao, Y., Liu, Q., Liu, Y., Zhu, M., & Xiao, L. (2024). Deep learning for hyperspectral image classification: A critical evaluation via mutation testing. *Remote Sensing*, *16*(24), 4695. https://doi.org/10.3390/rs16244695

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. Wiley.

Dhanya, V., Subeesh, A., Kushwaha, N., Vishwakarma, D. K., Nagesh Kumar, T., Ritika, G., & Singh, A. (2022). Deep learning based computer vision approaches for smart agricultural applications. *Artificial Intelligence in Agriculture*, *6*, 211–229. https://doi.org/https://doi.org/10.1016/j.aiia.2022.09.007

Ding, K., et al. (2023). Hyperspectral band selection based on mutual information and recursive feature elimination for crop classification. *Computers and Electronics in Agriculture*, *204*, 107538.

Dong, C., Yang, T., Liu, L., Wei, Z., Shi, C., & Gao, D. (2025). Early identification of apple bitter pit using hyperspectral imaging technology. *Applied Food Research*, *5*(2), 101166. https://doi.org/10.1016/j.afres.2025.101166

Emmons, C. L., & Scott, J. W. (1998). Ultrastructural and anatomical factors associated with resistance to cuticle cracking in tomato. *International Journal of Plant Sciences*, *159*, 14–22.

Fass, E., Shlomi, E., Ziv, C., Glikman, O., & Helman, D. (2025). Machine learning models based on hyperspectral imaging for pre-harvest tomato fruit quality monitoring. *Computers and Electronics in Agriculture*, *229*, 109788. https://doi.org/10.1016/j.compag.2024.109788

García-Vera, A., Polochè-Arango, F., & Gutiérrez-Bernal, A. (2024). Hyperspectral image analysis and machine learning techniques for crop disease detection and identification: A review. *Sustainability*, *16*(14), 6064. https://doi.org/10.3390/su16146064

Geladi, P., & Kowalski, B. R. (2001). Chemometrics in spectroscopy. part 1. classical chemometrics. *Analytica Chimica Acta*, *439*(1), 1–20.

Ghazal, S., Munir, A., & Qureshi, W. S. (2024). Computer vision in smart agriculture and precision farming: Techniques and applications. *Artificial Intelligence in Agriculture*, *13*, 64–83.

Guirao, A., García-Pastor, M. E., Solana-Guilabert, A., Valverde, J. M., & Martínez-Romero, D. (2025). Enhancing calcium transport in table grapes using sorbitol: A sustainable strategy for promoting fruit quality. *Horticulturae*, *11*(11). https://doi.org/10.3390/horticulturae11111320

Haghbin, N., Bakhshipour, A., Zareiforoush, H., & Mousanejad, S. (2023). Non-destructive pre-symptomatic detection of gray mold infection in kiwifruit using hyperspectral data and chemometrics. *Plant Methods*, *19*(1), 53. https://doi.org/10.1186/s13007-023-01032-4

Hall, M. E., Loeb, G., Cadle-Davidson, L., et al. (2018). Grape sour rot: A four-way interaction involving the host, yeast, acetic acid bacteria, and insects. *Phytopathology*, *108*, 1429–1442.

Hasanzadeh, B., Abbaspour-Gilandeh, Y., Soltani-Nazarloo, A., Hernández-Hernández, M., Gallardo-Bernal, I., & Hernández-Hernández, J. L. (2022). Non-destructive detection of fruit quality parameters using hyperspectral imaging, multiple regression analysis and artificial intelligence. *Horticulturae*, *8*(7), 598. https://doi.org/10.3390/horticulturae8070598

Hong, D., Li, C., Yokoya, N., Zhang, B., Jia, X., Plaza, A., Gamba, P., Benediktsson, J. A., & Chanussot, J. (2025). Hyperspectral imaging. https://arxiv.org/abs/2508.08107

Imaging, S. S. (2024). *What is hyperspectral imaging?: A comprehensive guide* (tech. rep.). Specim, Spectral Imaging Ltd. https://www.specim.com/technology/what-is-hyperspectral-imaging/

Kaltenborn, J., et al. (2025). Data splitting to avoid information leakage with datasail. *Scientific Reports*, *15*(1), 1–15.

Kanwal, N., Kämper, W., Farrar, M. B., Tootoonchy, M., Lynch, C., Nichols, J., Wallace, H. M., Trueman, S. J., & Hosseini Bai, S. (2025). Rapid assessment of lychee and mango fruit quality using hyperspectral imaging. *LWT*, *224*, 117833. https://doi.org/10.1016/j.lwt.2025.117833

Karantzi, A. D., Kafkaletou, M., Tsaniklidis, G., Bai, J., Christopoulos, M. V., Fanourakis, D., & Tsantili, E. (2021). Preharvest foliar salicylic acid sprays reduce cracking of fig fruit at harvest. *Applied Sciences*, *11*(23). https://doi.org/10.3390/app112311374

Khadivi-Khub, A. (2014). Physiological and genetic factors influencing fruit cracking. *Acta Physiologiae Plantarum*, *36*, 1713–1727.

Khaled, A. Y., Ekramirad, N., Donohue, K. D., Villanueva, R. T., & Adedeji, A. A. (2023). Non-destructive hyperspectral imaging and machine learning-based predictive models for physicochemical quality attributes of apples during storage as affected by codling moth infestation. *Agriculture*, *13*(5), 1086. https://doi.org/10.3390/agriculture13051086

Krisnawijaya, N. N. K., Tekinerdogan, B., Catal, C., & van der Tol, R. (2022). Data analytics platforms for agricultural systems: A systematic literature review. *Computers and Electronics in Agriculture*, *195*, 106813.

La Spada, P., Continella, A., Dominguez, E., Heredia, A., & Gentile, A. (2024). Improving cuticle thickness and quality traits in table grape cv. 'italia' using pre-harvest treatments. *Plants*, *13*(17). https://doi.org/10.3390/plants13172400

Li, S., et al. (2023). Information leakage in deep learning-based hyperspectral image classification: A survey. *Remote Sensing*, *15*(15), 3793.

Liang, Y., Li, Z., Shi, J., Zhang, N., Qin, Z., Du, L., Zhai, X., Shen, T., Zhang, R., Zou, X., & Huang, X. (2025). Advances in hyperspectral imaging technology for grain quality and safety detection: A review. *Foods*, *14*(17). https://doi.org/10.3390/foods14172977

Lichter, A., Kaplunov, T., Zutahy, Y., et al. (2014). Effects of gibberellin on cracking and postharvest quality of the seeded table grape 'zainy'. *Acta Horticulturae*, *1079*, 265–271.

Liu, Y., Zhang, H., & Wang, J. (2023). Early detection of plant physiological stress using VIS–NIR hyperspectral imaging. *Sensors*, *23*. https://doi.org/10.3390/s23083915

Liu, Y., et al. (2022). Feature selection for hyperspectral image classification: A survey. *Sensors*, *22*(10), 3782.

Mehta, N. (2018). *Hyperspectral imaging cameras in rapid detection and potential diagnosis of agricultural crop threats* (tech. rep.). HinaLea Imaging. https://hinalea.ai/wp-content/uploads/2018/03/Agriculture-in-Hyperspectral-Imaging-Whitepaper.pdf

Mgendi, G. (2024). Unlocking the potential of precision agriculture for sustainable farming. *Discover Agriculture*, *2*(1), 87.

Min, D., Zhao, J., Bodner, G., Ali, M., Li, F., Zhang, X., & Rewald, B. (2023a). Early decay detection in fruit by hyperspectral imaging–principles and application potential. *Food Control*, *152*, 109830. https://doi.org/10.1016/j.foodcont.2023.109830

Min, D., Zhao, J., Bodner, G., Ali, M., Li, F., Zhang, X., & Rewald, B. (2023b). Early decay detection in fruit by hyperspectral imaging–principles and application potential. *Food Control*, *152*, 109830. https://doi.org/10.1016/j.foodcont.2023.109830

Mishra, P., Asaari, M. S. M., Herrero-Langreo, A., Lohumi, S., Diezma, B., & Scheunders, P. (2017). Close range hyperspectral imaging of plants: A review. *Biosystems Engineering*, *164*, 49–67.

Moeini, A., & Wang, S. (2018). A dc link sensor-less voltage balancing technique for cascaded h-bridge multilevel converters with asymmetric selective harmonic current mitigation-pwm. *IEEE Transactions on Power Electronics*, *33*(9), 7571–7581. https://doi.org/10.1109/TPEL.2017.2770141

Monteiro, H. S. A., Tecchio, M. A., Brito, S. d. N. S., Neto, F. J. D., Sánchez, C. A. P. C., Alonso, J. C., Feliciano, D. E. F., Maniero, C. R., Cunha, P. H. H., & Silva, M. d. S. (2025). Phenological development, thermal requirement, and quality of 'brs núbia' (vitis vinifera l. x vitis labrusca l.) grapes on different rootstocks. *Horticulturae*, *11*(5). https://doi.org/10.3390/horticulturae11050466

Ou, Y., Yan, J., Liang, Z., & Zhang, B. (2024). Hyperspectral imaging combined with deep learning for the early detection of strawberry leaf gray mold disease. *Agronomy*, *14*(11), 2694. https://doi.org/10.3390/agronomy14112694

Panda, S., & Mishra, D. (2025). From spectrum to physiology: Hyperspectral indicators of plant function. *Photosynthetica*. https://doi.org/10.32615/ps.2025.012

Patrício, D. I., & Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and electronics in agriculture*, *153*, 69–81.

Peschel, S., & Knoche, M. (2005). Characterization of microcracks in the cuticle of developing sweet cherry fruit. *Journal of the American Society for Horticultural Science*, *130*, 487–495.

Ram, B. G., Oduor, P., Igathinathane, C., Howatt, K., & Sun, X. (2024). A systematic review of hyperspectral imaging in precision agriculture: Analysis of its current state and future prospects. *Computers and Electronics in Agriculture*, *222*, 109037. https://doi.org/10.1016/j.compag.2024.109037

Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., et al. (2024). Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*.

Santos, M., Egea-Cortines, M., Gonçalves, B., & Matos, M. (2023). Molecular mechanisms involved in fruit cracking: A review. *Frontiers in Plant Science*, *14*, 1130857.

Seo, H.-J., Sawant, S. S., & Song, J. (2022). Fruit cracking in pears: Its cause and management—a review. *Agronomy*, *12*(10). https://doi.org/10.3390/agronomy12102437

Shaw, G., & Manolakis, D. (2002). *Hyperspectral imaging remote sensing*. John Wiley & Sons.

Soussi, A., Zero, E., Sacile, R., Trinchero, D., & Fossa, M. (2024). Smart sensors and smart data for precision agriculture: A review. *Sensors*, *24*(8), 2647.

Späth, N., et al. (2024). Separating surface reflectance from volume reflectance in medical hyperspectral imaging. *Journal of Biomedical Optics*, *29*(8), 086001.

Sun, Q., Li, X., & Zhao, Y. (2024). VIS–NIR hyperspectral imaging for internal quality assessment of fruits. *Postharvest Biology and Technology*. https://doi.org/10.1016/j.postharvbio.2024.112095

Sun, T., Xu, R., Zhang, G.-x., Li, X.-y., Fan, S.-d., & Wang, X. (2020). Non-destructive determination of moisture content of single barley grain using hyperspectral imaging. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *226*, 117623.

Thenkarail, P. S., Lyon, J. G., & Huete, A. (2016). *Hyperspectral remote sensing of vegetation*. CRC Press.

Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation—a review. *Information processing in agriculture*, *7*(1), 1–19.

Tian, P., Meng, Q., Wu, Z., Lin, J., Huang, X., Zhu, H., Zhou, X., Qiu, Z., Huang, Y., & Li, Y. (2023). Detection of mango soluble solid content using hyperspectral imaging technology. *Infrared Physics & Technology*, *129*, 104576. https://doi.org/10.1016/j.infrared.2023.104576

Valme, D., Rassõlkin, A., & Liyanage, D. C. (2025). From adas to material-informed inspection: Review of hyperspectral imaging applications on mobile ground robots. *Sensors*, *25*(8), 2346. https://doi.org/10.3390/s25082346

Vignati, S., Tugnolo, A., Giovenzana, V., Pampuri, A., Casson, A., Guidetti, R., & Beghi, R. (2023). Hyperspectral imaging for fresh-cut fruit and vegetable quality assessment: Basic concepts and applications. *Applied Sciences*, *13*(17). https://doi.org/10.3390/app13179740

Wang, B., Yang, H., Zhang, S., & Li, L. (2023). Detection of defective features in cerasus humilis fruit based on hyperspectral imaging technology. *Applied Sciences*, *13*(5). https://doi.org/10.3390/app13053279

Wang, C., Li, X., Zhang, Z., Luo, X., Cai, J., & Wang, A. (2025). Nondestructive quality detection of characteristic fruits based on vis/nir spectroscopy: Princi-

ples, systems, and applications. *Agriculture*, *15*(20). https://doi.org/10.3390/agriculture15202167

Wang, T., & Zhao, L. (2024). Hyperspectral imaging for monitoring plant physiological status. *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2024.1298743

Xu, J., Chen, L., Dong, J., Jiang, L., & Hong, L. (2025). Overview of fruit cracking in sweet cherry (prunus avium l.): Causes, testing methods, mitigation strategies, and research perspectives. *Frontiers in Sustainable Food Systems*, *9*, 1534778.

Yang, C., Guo, Z., Fernandes Barbin, D., Dai, Z., Watson, N., Povey, M., & Zou, X. (2025). Hyperspectral imaging and deep learning for quality and safety inspection of fruits and vegetables: A review. *Journal of Agricultural and Food Chemistry*, *73*(17), 10019–10035. https://doi.org/10.1021/acs.jafc.4c11492

Yu, J., Li, J., Hong, Z., Zeng, Q., Fu, Y., Deng, R., Xu, K., Huang, Z., & Zhu, M. (2025). Vvpip1;1 plays a role in grape berry cracking by regulating water uptake. *Horticultural Plant Journal*, *11*(4), 1506–1516. https://doi.org/https://doi.org/10.1016/j.hpj.2024.07.010

Zhang, C., Wang, Q., He, H., Wu, Y., Shan, W., & Liu, H. (2025). Postharvest preservation strategies for table grapes: A comprehensive review from practical methods to future developments. *Plants*, *14*(16). https://doi.org/10.3390/plants14162462

Zhang, J., et al. (2021). Hyperspectral band selection for tobacco disease identification based on fisher score. *Spectroscopy and Spectral Analysis*, *41*(4), 1234–1240.

Zhang, X., Li, Y., Chen, H., & Wang, Q. (2024). Application of hyperspectral imaging in fruit quality and safety assessment. *Frontiers in Plant Science*, *15*, 1432120. https://doi.org/10.3389/fpls.2024.1432120

Zhang, X., Vinatzer, B. A., & Li, S. (2024). Hyperspectral imaging analysis for early detection of tomato bacterial leaf spot disease. *Scientific Reports*, *14*(1), 27666. https://doi.org/10.1038/s41598-024-78907-x

Zhang, Y., Chen, J., & Liu, Q. (2022). Physiological interpretation of VIS–NIR hyperspectral reflectance in vegetation. *Remote Sensing*, *14*. https://doi.org/10.3390/rs14215345

Zhang, Y., Li, X., Wang, M., Xu, T., Huang, K., Sun, Y., Yuan, Q., Lei, X., Qi, Y., & Lv, X. (2024). Early detection and lesion visualization of pear leaf anthracnose based on multi-source feature fusion of hyperspectral imaging. *Frontiers in Plant Science*, *15*, 1461855. https://doi.org/10.3389/fpls.2024.1461855

Zhang, Z., Cheng, H., Chen, M., Zhang, L., Cheng, Y., Geng, W., & Guan, J. (2024). Detection of pear quality using hyperspectral imaging technology and machine learning analysis. *Foods*, *13*(23), 3956. https://doi.org/10.3390/foods13233956

Zhu, H., Qin, S., Liang, S., Su, M., Wang, P., & He, Y. (2025). Hyperspectral imaging and machine learning for quality assessment of apples with different bagging

types. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *343*, 126443. https://doi.org/10.1016/j.saa.2025.126443

Zwinkels, J. C. (2015). Spectrophotometry. In *Encyclopedia of color science and technology*. Springer. https://doi.org/10.1007/978-1-4419-8071-7_221