remote sensing

MDPI

Article

Self-Supervised Learning for Soybean Disease Detection Using UAV Hyperspectral Imagery

Mustafizur Rahaman [1], Vasit Sagan [1,2,*], Felipe A. Lopes [1], Haireti Alifu [1], Cagri Gul [1], Hadi Aliakbarpour [2] and Kannappan Palaniappan [3]

[1] Remote Sensing Laboratory, Saint Louis University, St. Louis, MO 63103, USA; mustafizur.rahaman@slu.edu (M.R.)
[2] Department of Computer Science, Saint Louis University, St. Louis, MO 63103, USA
[3] Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO 65211, USA
* Correspondence: vasit.sagan@slu.edu

Highlights

What are the main findings?

- A self-supervised learning framework achieves 92% accuracy in early soybean disease detection using unlabelled UAV hyperspectral data, matching supervised baselines.
- A distance-based spectral pairing technique enables effective feature learning directly from canopy reflectance without manual annotations.

What are the implications of the main findings?

- The framework addresses the annotation bottleneck in remote sensing, enabling scalable early disease detection for large-scale agricultural monitoring.
- The approach reduces reliance on expert-labeled field data while maintaining high accuracy, making precision agriculture more accessible and cost-effective.

Abstract

The accuracy of machine learning models in plant disease detection significantly relies on large volumes of knowledge-based labeled data; the acquisition of annotation remains a significant bottleneck in domain-specific research such as plant disease detection. While unsupervised learning alleviates the need for labeled data, its effectiveness is constrained by the intrinsic separability of feature clusters. These limitations underscore the need for approaches that enable supervised early disease detection without extensive annotation. To this end, we propose a self-supervised learning (SSL) framework for the early detection of soybean's sudden death syndrome (SDS) using hyperspectral data acquired from an unmanned aerial vehicle (UAV). The methodology employs a novel distance-based spectral pairing technique that derives intermediate labels directly from the data. In addition, we introduce an adapted contrastive loss function designed to improve cluster separability and reinforce discriminative feature learning. The proposed approach yields an 11% accuracy gain over agglomerative hierarchical clustering and attains both classification accuracy and F1 score of 0.92, matching supervised baselines. Reflectance frequency analysis further demonstrates robustness to label noise, highlighting its suitability in label-scarce settings.

Keywords: self-supervised learning; hyperspectral; plant disease; clustering; remote sensing

check for updates

Academic Editor: Ehsan Khoramshahi

Received: 10 October 2025
Revised: 26 November 2025
Accepted: 28 November 2025
Published: 4 December 2025

Citation: Rahaman, M.; Sagan, V.; Lopes, F.A.; Alifu, H.; Gul, C.; Aliakbarpour, H.; Palaniappan, K. Self-Supervised Learning for Soybean Disease Detection Using UAV Hyperspectral Imagery. *Remote Sens.* **2025**, *17*, 3928. https://doi.org/10.3390/rs17233928

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Extreme climate events have intensified plant disease outbreaks over the past two decades, driven by rising temperatures, altered precipitation regimes, and humidity fluctuations [1–3]. Soybeans, a globally important crop for protein and oil, are particularly vulnerable to climate-driven stresses, with heightened disease susceptibility threatening yield stability and long-term food security [4,5]. Conventional visual inspection remains the primary diagnostic method, but is labor-intensive, subjective, and inadequate for early detection when timely intervention is most critical [6]. Advances in remote sensing, integrated with machine learning, now offer scalable, non-invasive monitoring solutions that enable earlier diagnosis, targeted pesticide application, and potential yield improvements, while reducing environmental impacts [7]. However, these approaches predominantly rely on supervised learning frameworks that require large volumes of manually annotated training data.

Recent work on plant disease detection has largely centered on leaf-image analysis with computer vision. Convolutional neural networks (CNNs) achieve high accuracy in controlled settings, predominantly on leaf image datasets [8]. However, leaf-image methods are poorly suited for identifying diseases before visible symptoms emerge, as many stressors do not manifest visibly on the leaf surface [9], and image acquisition at scale is operationally burdensome for commercial agriculture. This limitation extends to self-supervised learning approaches, which have primarily focused on controlled leaf-image datasets [10,11] rather than field-scale canopy monitoring.

To overcome these spatial and operational limitations of leaf-level approaches, unmanned aerial vehicles (UAVs) with hyperspectral sensors offer a field-scale alternative and have shown promise for early disease detection in experimental and commercial settings [12,13]. Narrow bands in the red-edge and near-infrared regions are sensitive to water status and subtle physiological changes that precede visible symptoms [14]. Despite these potential, the practical deployment of UAV-based hyperspectral imaging for early detection faces a critical bottleneck, as most supervised pipelines require large volumes of expertly labeled, in situ data [15,16]. This dependency is especially problematic when symptoms are pre-visual and subtle, making annotation time-consuming, costly, and subjective [17,18]. The challenge is amplified for root-origin diseases such as soybean sudden death syndrome (SDS), where early canopy signals are minimal and demand specialized expertise to identify reliably [19]. As a result, assembling sufficiently large labeled datasets for robust supervised learning remains prohibitively expensive for large-scale monitoring, underscoring the need for methods that reduce annotation burden while preserving accuracy.

Self-supervised learning (SSL) offers a methodology with particular promise for addressing scenarios characterized by limited labeled data. Unlike unsupervised learning, which is hindered by the curse of dimensionality and limited feature separability [20–22], SSL creates pseudo-labels through pretext tasks. This approach enables representation learning without extensive manual annotation while maintaining discriminative power. Recent SSL frameworks, including contrastive learning [23], Siamese networks [24,25], BYOL [26], and SimSiam [27], have demonstrated strong capabilities in extracting meaningful features from unlabeled data [28,29].

Despite SSL's proven effectiveness in agricultural applications such as crop type classification and leaf-based disease detection [30], its potential for large-scale UAV-based hyperspectral monitoring of early-stage field diseases remains untapped. This gap is particularly critical because existing SSL frameworks face three key limitations when applied to field-scale monitoring. First, most existing studies focus on controlled leaf-image datasets [10,11], which are inadequate for detecting early-stage diseases like SDS, where symptoms mani-

fest subtly at the canopy level before becoming visible on individual leaves. Second, SSL architectures designed for spatial features (typically CNN-based) do not directly transfer to hyperspectral data, which require specialized handling of high-dimensional spectral signatures [11,30,31]. Third, there is a lack of SSL frameworks specifically designed to leverage the unique characteristics of hyperspectral reflectance data, such as physiologically meaningful spectral bands in the red-edge and NIR regions, for early disease detection at the plot or field scale.

To address this critical gap, this study proposes a self-supervised learning framework for sudden death syndrome (SDS) detection in soybeans. The framework operates on canopy-level hyperspectral reflectance data acquired from unmanned aerial vehicles, enabling effective disease monitoring under label-scarce conditions. Unlike conventional approaches that depend on extensive manual annotation, particularly problematic for early-symptomatic diseases, our framework derives supervision signals directly from the intrinsic structure of unlabeled spectral data. The main contributions of this work are summarized as follows:

1. An end-to-end self-supervised learning model is developed, specifically tailored for early-stage SDS detection using UAV hyperspectral data.
2. A Euclidean distance-based pseudo-labeling strategy is introduced that leverages the physiologically meaningful separability in hyperspectral space to create high-confidence training pairs, enabling encoder training directly from spectral representations without manual annotation.
3. A comparative evaluation is conducted to benchmark the SSL framework against conventional clustering algorithms and supervised classifiers.

This framework demonstrates that self-supervised learning can achieve performance comparable to supervised baselines while substantially reducing the reliance on labeled data, suggesting its potential as a scalable approach for hyperspectral-based plant disease detection in scenarios where manual annotation is limited or costly.

2. Materials and Methods

2.1. Study Area

The experimental site was situated in Montgomery County, Illinois, USA. The region experiences a continental climate, characterized by warm, humid summers and cold winters. According to data from the National Oceanic and Atmospheric Administration (NOAA), July is the warmest month, with an average temperature of 23 °C, while November is the coldest, averaging 6 °C. Rainfall peaks in May and gradually declines through November. The trials were conducted on Mollisol soils with a loam texture, commencing in May 2022. Experimental plots varied in size, ranging from 1.5 to 3 m in width and 5.5 to 6 m in length, with a row spacing of 0.76 m. Standard tillage practices were followed, and all trials were conducted under rainfed conditions. Disease symptoms emerged naturally and were in their early stages at the time of data collection in mid-August 2022. Harvesting occurred in October and November, in accordance with genotype maturity groups. Visual assessment of disease severity was conducted to confirm the disease status.

Disease status was determined through labor-intensive field assessment conducted on 15 August 2022, by pathological and phenological specialists from GDM Seeds. Evaluators independently inspected individual plots, examining stems and leaves for SDS symptoms, and the assessments were later compiled to establish final classifications. Following a conservative protocol, plots exhibiting yellowing, chlorosis, or damage on any observed stem or leaf were labeled as SDS-affected, while plots with no visible symptoms were labeled as healthy, a threshold that necessarily introduced some classification ambiguity in borderline cases with minimal symptom expression. Figure 1 provides an overview of the

study area and plot layout of the 375 total plots, 206 were classified as healthy (55%) and 169 as SDS-affected (45%), representing a moderate class imbalance that was addressed during SSL training through balanced pair generation (Section 2.6). At this midpoint of the season, the plants were in the early to mid stages of their life cycle, and SDS was in an early development phase.

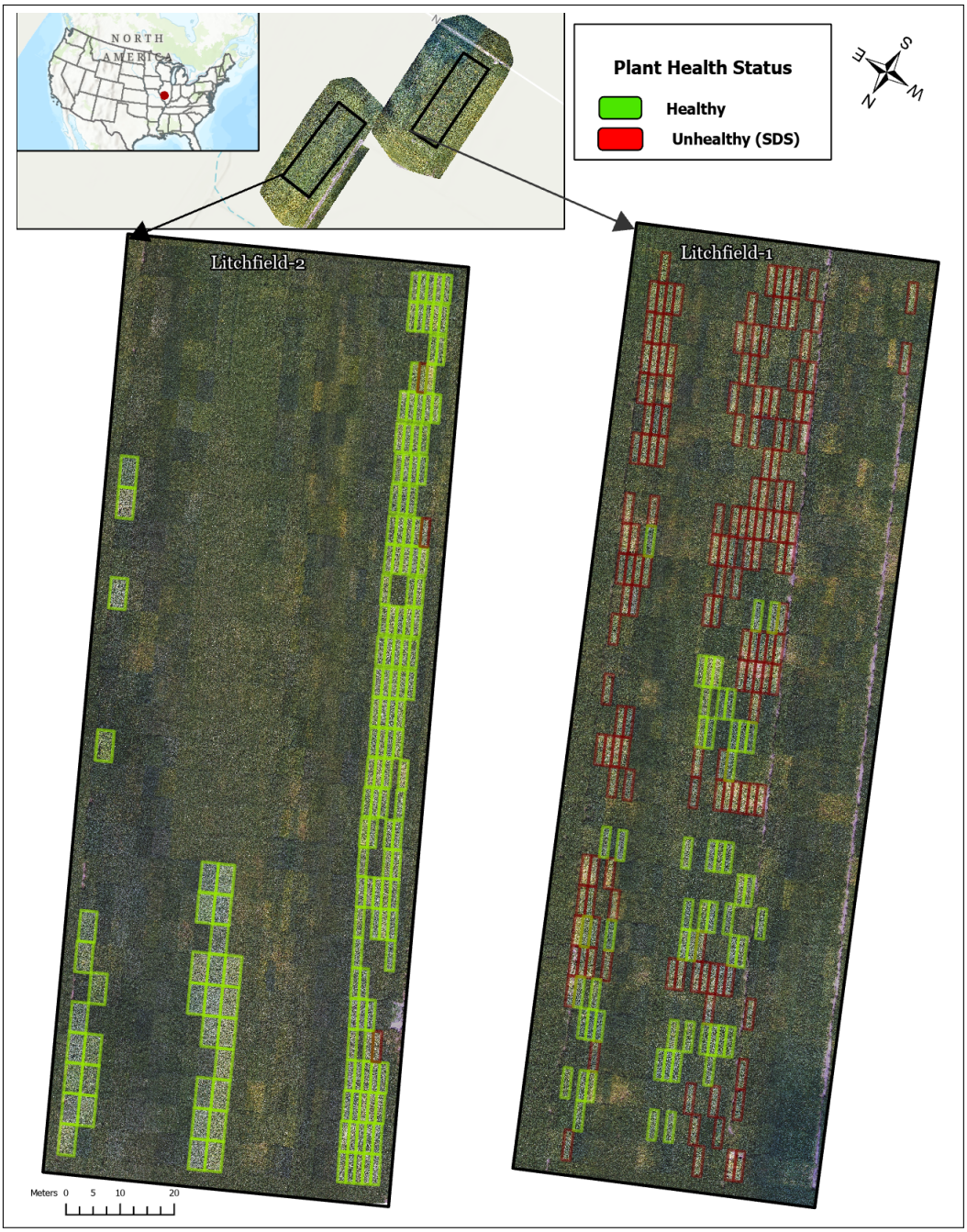

Figure 1. Study area and plot-level health status of the soybean. Plot with any symptom annotated as red, while green indicates no symptoms.

2.2. Materials

Hyperspectral data acquisition was conducted through a single field campaign in mid-August 2022 using a Headwall Nano-Hyperspec VNIR sensor (Headwall Photonics, Bolton, MA, USA) (400–1000 nm, 269 bands) mounted on a Matrice 600 Pro UAV (DJI, Shenzhen, China). The flight was executed at an altitude of 100 m above ground level, yielding a spatial resolution of approximately 4 cm per pixel. Georeferencing was performed using

a mobile R12 Real-Time Kinematic (RTK) GNSS receiver (Trimble, Sunnyvale, CA, USA), which provided sub-centimeter-level positioning accuracy for ground control points. Raw hyperspectral data, initially distorted and non-georeferenced, were calibrated to consistent reflectance values for analysis. The dataset was split into training (80%, n = 300) and testing (20%, n = 75) sets using stratified sampling to maintain class distribution. The self-supervised learning framework was implemented in Python 3.10 using PyTorch 1.12, with data preprocessing and clustering analyses conducted using scikit-learn, NumPy, and pandas libraries.

2.3. Data Pre-Processing

The preprocessing pipeline transforms raw hyperspectral imagery into plot-level spectral signatures suitable for SSL training. Traditional SSL architectures predominantly employ CNNs, which necessitate extensive datasets, substantial computational resources, and rich spatial features to achieve effective feature separation [32–34]. However, the high dimensionality of HSI data and inadequate spatial features in early-stage crop vegetation increase architectural complexity [35]. To mitigate these challenges, this study utilizes plot-level mean spectra as the dataset, enhancing computational efficiency while preserving discriminative spectral information.

The preprocessing pipeline, illustrated in Figure 2, consists of three main stages. First, raw hyperspectral data undergo radiometric calibration to correct sensor-specific errors and convert digital numbers to standard reflectance values, followed by orthorectification to remove geometric distortions. Second, individual hyperspectral image cubes are georeferenced using the R12 RTK GNSS data and mosaicked to generate a seamless field-scale image.

Third, background soil and shadow pixels were removed using an adapted data-driven approach [36] to isolate pure canopy spectral signals. The soil masking technique exploits inherent spectral differences between vegetation and soil across VNIR wavelengths, using rule-based classification that leverages characteristic reflectance patterns of photosynthetically active vegetation versus bare soil. This unsupervised approach eliminates the need for manual training sample generation while maintaining robust performance. Shadow masking was implemented to detect and exclude shaded regions cast by the UAV platform and environmental conditions, ensuring spectral consistency across plots. Shadow removal is critical because shaded areas exhibit incomplete spectral information and reduced intensity values that could confound disease detection.

Following soil and shadow removal, plot-level mean reflectance signatures were computed exclusively from pure canopy pixels, ensuring that extracted spectral features represent genuine vegetation characteristics rather than mixed soil-plant or illumination artifacts. All spectra are then smoothed using the Savitzky-Golay filter to reduce noise. Subsequently, 47 vegetation indices spanning chlorophyll content ($NDVI, GNDVI, MCARI$ variants), water stress ($WBI, WSCT, NDWI$), pigment composition ($SRPI, NPCI, ARI_2$), and structural characteristics ($SAVI, OSAVI, MSAVI$) are computed from the 269 spectral bands using established formulas from the remote sensing literature (Table S2). These indices provide biophysically interpretable features that complement raw spectral signatures by encoding domain knowledge about vegetation stress responses. The 269 spectral bands and 47 derived indices are concatenated to form a 316-dimensional feature vector per plot, which undergoes standardization before model training.

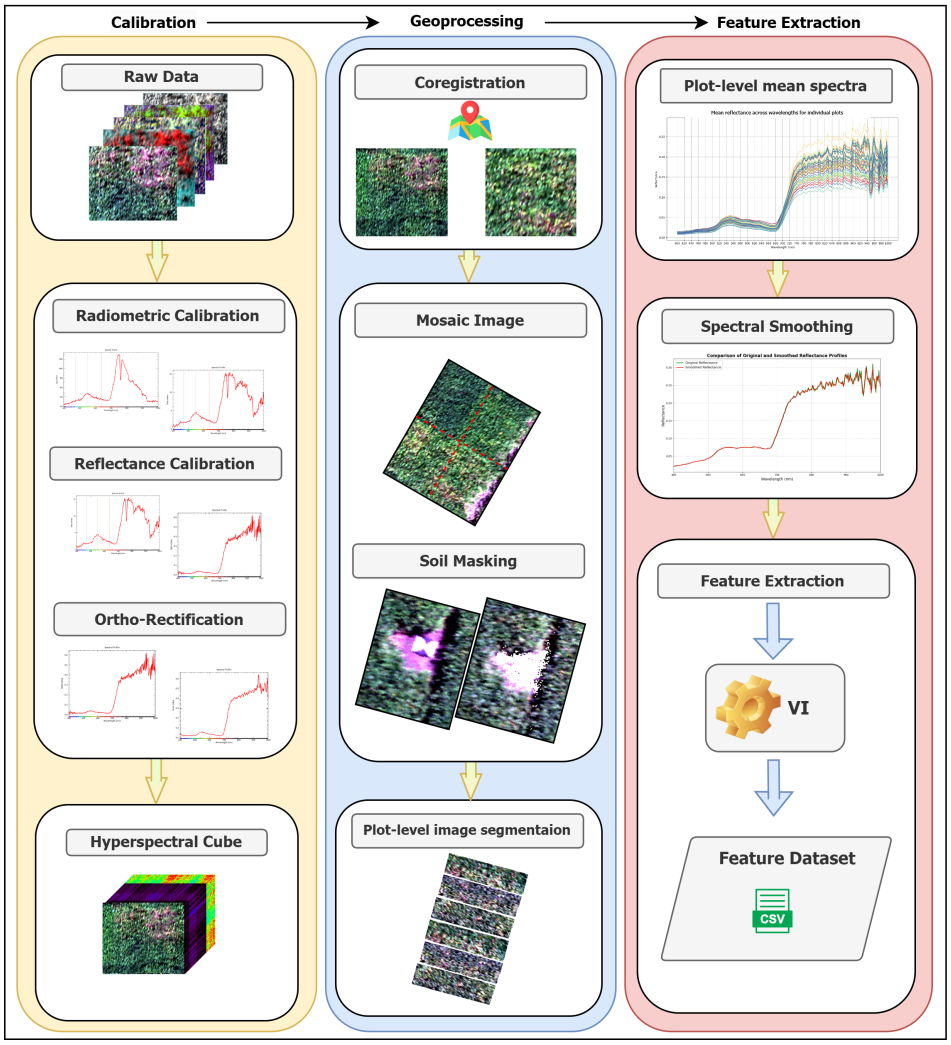

Figure 2. Data processing workflow highlighting learning and inference performed in the spectral domain.

2.4. Plot-Level Mean Spectra Analysis

Despite rigorous filtering and calibration processes that maintain standard procedural integrity, notable edge noise persists within the dataset. Figure 3a illustrates the mean spectral reflectance curves for healthy and unhealthy plots. The spectral reflectance curves for unhealthy plots exhibit a broader horizontal distribution compared to those of healthy plots. This wider dispersion in spectral reflectance values indicates greater variability among unhealthy plots. Consequently, this variability complicates the differentiation between healthy and unhealthy plots when relying solely on spectral information. The increased spectral overlap highlights the challenge of using spectral data alone to accurately classify plot health, underscoring the need for incorporating advanced modeling techniques to enhance classification accuracy.

2.5. SSL Architecture

This section presents the self-supervised learning framework developed for SDS detection. As introduced in Section 1, we develop an SSL framework tailored to plot-level hyperspectral reflectance. Unlike methods predicated on spatial context (CNNs) or semantic tokenization (LLMs), our approach operates directly on tabular spectra, where neither assumption holds. The framework comprises: (i) a Siamese encoder architecture (Section 2.5), (ii) a spectral distance-guided pairing scheme that yields high-confidence positive/negative

pairs from unlabeled data (Section 2.6), and (iii) an exponentiated contrastive loss that modulates the strength of inter/intra-class separation via a hyperparameter exponent (Section 2.7).

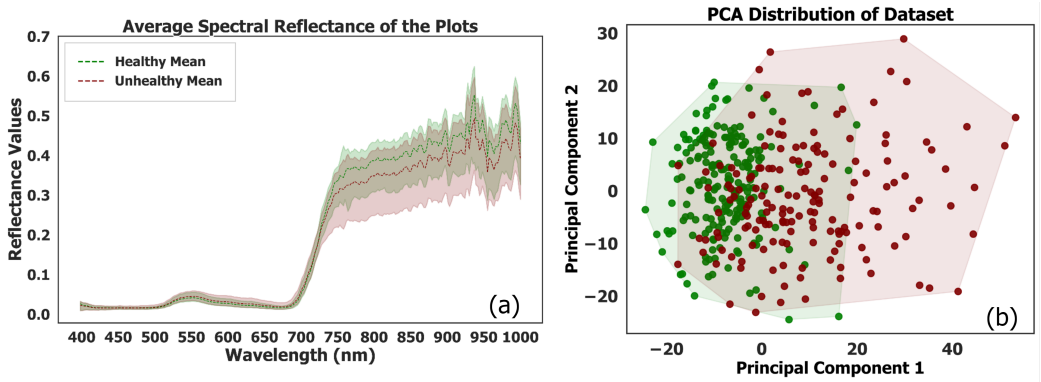

Figure 3. (**a**) All samples' mean spectral reflectance and (**b**) top two principal component distribution of data points. Red color denotes the SDS affected, while green denotes the healthy.

Figure 4 summarizes the training pipeline. Given an unlabeled dataset of per-plot index-reflectance vectors, we (a) standardize per band and construct pairs using the distance-guided data pairing method, (b) optimize a Siamese encoder with an exponentiated margin-based contrastive objective, and (c) deploy the trained encoder to obtain low-dimensional embeddings for evaluation. For the encoder, we use a fully connected feed-forward network with five hidden layers, comprising 316 input neurons, successive layers of 512, 256, 128, and 64 neurons, and a 32-dimensional embedding output (ReLU activations and layer normalization). The 32-D embedding provides a compact representation sufficient for downstream separability while mitigating overfitting under limited labels. Training uses Adam (learning rate 10^{-3}), L_2 regularization, early stopping (patience 25), and 300 epochs. Model training convergence under 30 min. After convergence, the encoder processes the test set to produce test-embeddings.

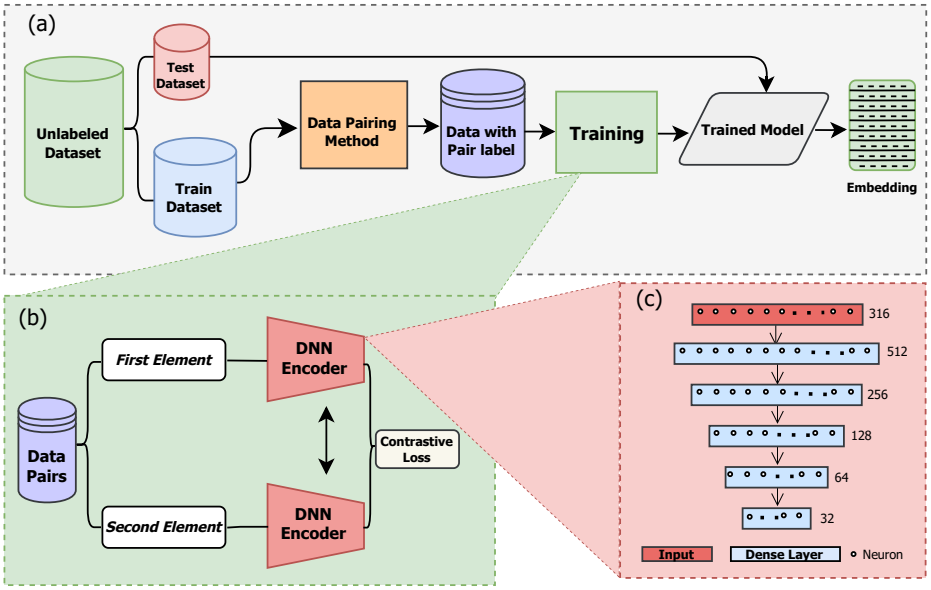

Figure 4. (**a**) Model architecture with workflow, (**b**) training mechanism, and (**c**) encoder architecture.

2.6. Distance-Based Pairing Strategy

We construct training pairs directly in standardized reflectance vectors, following distance-based SSL pretext (e.g., SimCLR [29], MoCo [37]). Let $X = \{x_i\}_{i=1}^{N} \subset \mathbb{R}^B$ denoted normalized vi-spectra with B = 316. For uniformly sampled index pairs (i, j), we compute Euclidean distance $d_{ij} = \|x_i - x_j\|_2$ and assign high-confidence pseudo-labels using two thresholds (α, β) with $0 < \alpha < \beta$:

$$(P, L_S) = \bigcup_{k=1}^{n} \begin{cases} ((x_i, x_j), 0) & \text{if } 0 < \|x_i - x_j\| \leq \alpha, \\ ((x_i, x_j), 1) & \text{if } \|x_i - x_j\| \geq \beta, \end{cases}$$
$$\text{where} \quad 0 \leq \alpha \leq \beta \text{ and } i \neq j \tag{1}$$

Pairs with $\alpha < d_{ij} < \beta$ are discarded to avoid ambiguous supervision. Index pairs are de-duplicated, and similar/dissimilar counts are balanced to reduce label bias. Threshold selection in SSL poses an intrinsic challenge: appropriate cutoffs must be inferred from unlabeled data structure alone, precluding direct reliance on ground-truth validation.

A distribution-driven approach was adopted wherein thresholds are derived from empirical quantiles of the pairwise distance distribution. Distances were computed for a representative subset of training instances, with α and β positioned at the 20th and 80th percentiles, respectively, yielding $\alpha = 0.74$ and $\beta = 1.51$ (Figure S1) in standardized spectral space. A total of 3000 pairs were generated ($n = 3000$, with $n_{\text{similar}} \approx n_{\text{dissimilar}}$) with balanced distribution between similar and dissimilar labels to prevent class bias during training. This percentile-based threshold selection constitutes a data-adaptive strategy that instinctively scales to the intrinsic geometry of the feature space, obviating the need for manual tuning on labeled validation data, which would violate the self-supervised paradigm.

The 20th/80th percentile positions reflect a principled trade-off between pseudo-label confidence and training sample sufficiency: more conservative thresholds (e.g., 10th/90th percentiles) would yield higher-purity pairs at the cost of insufficient training diversity, while more permissive cutoffs (e.g., 30th/70th percentiles) would increase label noise in the contrastive signal. The adopted quartile-based approach ensures that approximately 40% of candidate pairs receive confident pseudo-labels (20% similar, 20% dissimilar), while the ambiguous middle 60% are excluded—a conservative strategy aligned with established self-supervised learning practices that prioritize signal quality over quantity. Crucially, this method generalizes across datasets because percentiles adapt to the observed distance distribution: a dataset exhibiting tighter spectral clustering would produce smaller absolute threshold values, while one with greater inter-class dispersion would yield larger thresholds, yet both would achieve equivalent relative separation in their respective feature spaces.

This adaptive property ensures methodological transferability without requiring dataset-specific hyperparameter tuning. This pairing strategy mitigates the class imbalance present in the raw dataset (55% healthy, 45% SDS-affected) by ensuring equal representation of both relationship types during contrastive learning. The quartile-based thresholds strike a balance between the need for sufficient training pairs and the requirement for high-confidence pseudo-labels, accepting approximately one-fifth of pairs as similar, one-fifth as dissimilar, and rejecting three-fifths as ambiguous. Algorithm 1 operationalizes this strategy with stratified sampling, gray-zone exclusion, de-duplication, and class balancing. The distance distribution exhibits (Figure S1) a unimodal structure with an extended right tail, characteristic of high-dimensional spectral data, where most pairs show moderate dissimilarity, with distributional extrema representing unambiguous cases. Euclidean

distance on normalized vi-spectra preserves physiologically informative amplitude differences (red-edge, NIR regions), constitutes a proper metric for margin-based objectives, and enables computationally efficient pair generation.

Algorithm 1 Distance-Based Pair Generation

Initialize $S \leftarrow \emptyset$, $P \leftarrow []$, $y \leftarrow []$
Set $n_{\text{similar}}, n_{\text{dissimilar}} \leftarrow \lfloor n/2 \rfloor$
Set $I \leftarrow \text{indices}(X)$
while $|P| < n$ **do**
  Randomly sample $i, j \in I$ where $i \neq j$
  $K \leftarrow \text{sorted}(\{i, j\})$ {Create unique pair identifier}
  if $K \notin S$ **then**
    Compute $d_{ij} = \|x_i - x_j\|_2$
    if $d_{ij} \leq \alpha$ **and** $n_{\text{similar}} > 0$ **then**
      $P \leftarrow P \cup \{(x_i, x_j)\}$, $y \leftarrow y \cup \{0\}$
      $n_{\text{similar}} \leftarrow n_{\text{similar}} - 1$
      $S \leftarrow S \cup \{K\}$
    else if $d_{ij} \geq \beta$ **and** $n_{\text{dissimilar}} > 0$ **then**
      $P \leftarrow P \cup \{(x_i, x_j)\}$, $y \leftarrow y \cup \{1\}$
      $n_{\text{dissimilar}} \leftarrow n_{\text{dissimilar}} - 1$
      $S \leftarrow S \cup \{K\}$
    end if
  end if
end while
return P, y

2.7. Contrastive Loss

The contrastive loss function trains the model using negative sampling to distinguish between similar and dissimilar pairs. We introduce a flexible exponential parameter n on top of the traditional contrastive loss to enhance model performance. This enhancement allows for better discrimination between similar and dissimilar pairs. The loss function $L(W, (Y, x_1, x_2))$ is defined as:

$$(W, (Y, x_1, x_2)) = \frac{1}{N} \sum_{i=1}^{N} [(1 - Y_i) \cdot \|x_{1_i} - x_{2_i}\|^n + Y_i \cdot \max(0, \text{margin} - \|x_{1_i} - x_{2_i}\|^n)] \quad (2)$$

Here, x_{1_i} and x_{2_i} represent the i-th pair of embedding vectors, and Y_i is the corresponding label indicating whether the pair is similar ($Y_i = 0$) or dissimilar ($Y_i = 1$). The term $\|x_{1_i} - x_{2_i}\|$ denotes the Euclidean distance between the embeddings. For similar pairs ($Y_i = 0$), the loss is $\|x_{1_i} - x_{2_i}\|^n$, encouraging the embeddings to be close. For dissimilar pairs ($Y_i = 1$), the loss is $\max(0, \text{margin} - \|x_{1_i} - x_{2_i}\|^n)$, encouraging a distance greater than the margin. The overall loss is averaged over the total number of pairs N.

The traditional contrastive loss uses a fixed exponent of $n = 2$, which limits the model's ability to emphasize larger distances. By allowing n to vary, the modified loss can place greater emphasis on larger distances, resulting in better separation of dissimilar pairs in the embedding space. This flexibility leads to improved performance as the model becomes more sensitive to the nuances of the data. The optimal value of n for hyperspectral disease detection was determined through systematic ablation analysis (Section 2.8).

2.8. Evaluation Metrics

The embedding vectors generated from the trained model (Figure 4) encode feature representations for each data point. To evaluate the effectiveness of SSL, both the original and embedded test data are clustered using K-Means [38] and Agglomerative Hierarchical

Clustering (AHC) [39]. Since SSL operates without labeled supervision, clustering performance is assessed using cluster accuracy, where dominant class labels are assigned to clusters using the Hungarian algorithm [40], and the Adjusted Rand Index (ARI), which quantifies clustering agreement with ground truth while accounting for random assignments. ARI is particularly useful in imbalanced datasets, ensuring that performance improvements reflect meaningful feature separability rather than chance.

$$\text{Accuracy} = \frac{\sum_i CM_{ii}}{\sum_i \sum_j CM_{ij}} \tag{3}$$

Equation (3) defines accuracy, where CM_{ii} represents correctly classified instances in the confusion matrix, and $\sum_i \sum_j CM_{ij}$ is the total number of instances. To benchmark SSL against traditional supervised learning, we compare it with Random Forest (RF), Support Vector Machine (SVM), and Deep Neural Network (DNN). Performance is evaluated using Accuracy, Precision, Recall, and F1-score, standard metrics in classification tasks. While accuracy provides a global correctness measure, it can be misleading in class-imbalanced settings. Precision is critical when minimizing false positives, whereas recall is essential when false negatives must be minimized, such as in detecting unhealthy plants. F1-score provides a balanced assessment, integrating both precision and recall. By combining clustering-based and classification-based evaluations, this study ensures a comprehensive assessment of SSL's capability in feature representation learning and its viability as an alternative to fully supervised methods.

2.9. Ablation Study: Contrastive Loss Exponent

The contrastive loss exponent n (Equation (2)) directly controls inter-class separation characteristics in the learned embedding space, making its selection critical to framework performance. As detailed in Section 2.7, the exponential modification extends conventional contrastive loss formulations by introducing a hyperparameter that modulates the penalization magnitude applied to embedding distances. Rather than relying on the standard $n = 2$ from conventional contrastive learning, we conducted a systematic ablation study to determine the optimal value for hyperspectral disease detection.

Five-fold cross-validation was performed on the training set (N = 3000), wherein the exponent n was varied over the discrete set $n \in \{2, 4, 6, 8, 10\}$. For each candidate value, the model was trained independently, and the resulting embeddings were evaluated using K-Means and AHC clustering algorithms. Clustering performance was quantified using cluster accuracy (with Hungarian algorithm alignment) and Adjusted Rand Index (ARI), both established metrics for unsupervised evaluation that account for chance agreement and class imbalance. Table 1 presents the clustering outcomes across different exponent values, reporting results from the best-performing fold. The results demonstrate consistent and substantial improvement in both clustering accuracy and ARI metrics as n increases from 2 to 8. Performance reaches its maximum at $n = 8$, where K-Means achieves a clustering accuracy of 0.88 and ARI of 0.57, while AHC achieves 0.92 accuracy and 0.70 ARI. Beyond $n = 8$, performance plateaus or slightly degrades ($n = 10$: K-Means accuracy 0.85, AHC 0.90), suggesting diminishing returns from excessive penalization.

The observed trend is consistent with the theoretical motivation behind the exponential modification. The standard contrastive loss ($n = 2$) applies quadratic penalization, which may insufficiently discriminate between moderately dissimilar and highly dissimilar pairs in high-dimensional hyperspectral space. Higher exponent values amplify the penalty gradient for smaller inter-class distances, thereby enforcing more pronounced cluster separation. Specifically, for dissimilar pairs with embedding distance d, the penalty term $\max(0, \text{margin} - d^n)$ exhibits steeper gradients as n increases, resulting in stronger

repulsive forces that expand inter-class margins. This geometric restructuring promotes enhanced intra-class compactness and inter-class dispersion, properties particularly advantageous for unsupervised clustering algorithms. The 54% improvement in K-Means accuracy (from 0.57 at $n = 2$ to 0.88 at $n = 8$) and 46% improvement in AHC accuracy (from 0.63 to 0.92) empirically validate this theoretical framework. From both theoretical and empirical perspectives, $n = 8$ represents an optimal configuration that balances intra-class cohesion and inter-class separability without over-penalizing, which could potentially destabilize training convergence. Accordingly, this hyperparameter setting ($n = 8$) was adopted for all subsequent analyses presented in this study, and all results in Section 3 use this optimized configuration.

Table 1. Ablation study results: clustering performance for different contrastive loss exponent values. Results from the best-performing fold of 5-fold cross-validation on the training set (N = 3000). The optimal value $n = 8$ (bold) was selected for all subsequent experiments.

n	K-Means Acc.	K-Means ARI	AHC Acc.	AHC ARI
2	0.57	0.25	0.63	0.20
4	0.61	0.38	0.63	0.37
6	0.71	0.43	0.82	0.55
8	**0.88**	**0.57**	**0.92**	**0.70**
10	0.85	0.65	0.90	0.66

3. Results

This section presents the impact of SSL-based data transformation on model performance. Section 3.1 analyzes the embedding space distribution before and after SSL encoding. Section 3.2 evaluates clustering performance in an unsupervised setting. Section 3.3 compares SSL with supervised learning baselines and examines label efficiency. Finally, Section 3.4 presents spatial analysis of plot-level predictions. Detailed interpretation of these findings is provided in Section 4.

3.1. Distribution of Embedding

Figure 5 presents two scatter plots depicting the clustering of test data points, each characterized by 316 features. Principal Component Analysis (PCA) is employed to reduce the data's dimensionality, enabling a visual representation of its distribution. The left plot illustrates the distribution of the raw test data, while the right plot represents the distribution after applying the SSL embedding technique outlined in Section 2.5. Both visualizations utilize the Agglomerative Hierarchical Clustering (AHC) algorithm to partition the soybean plots into two clusters, facilitating an assessment of clustering performance before and after embedding.

In the raw spectral space (Figure 5 (left)), considerable mixing occurs within cluster boundaries, with healthy and unhealthy samples frequently co-located. The Silhouette coefficient of 0.470 indicates moderate cluster quality, while the 3.1% convex hull overlap quantifies the ambiguous decision region. Following SSL embedding (Figure 5 (right)), cluster quality improves markedly (Silhouette: 0.526), with complete elimination of hull overlap. The 71.8% increase in centroid separation (1.20 to 2.06 units) further confirms enhanced feature discriminability. These quantitative improvements align with the 92% classification accuracy reported in Section 3.3, demonstrating that while SSL substantially enhances class separability, it appropriately maintains some boundary samples that reflect genuine spectral ambiguity in early-stage disease detection.

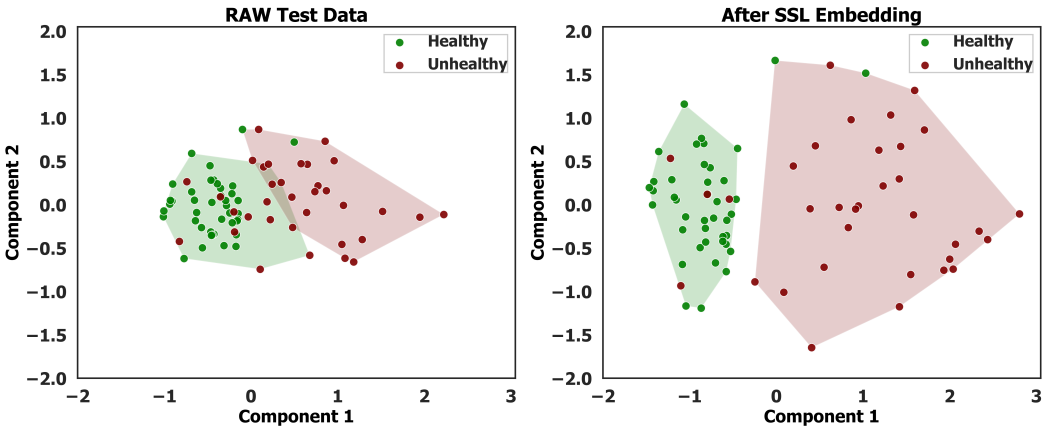

Figure 5. PCA distribution of Agglomerative hierarchical clustering performance before (**left**) and after (**right**) SSL embedding. Convex hulls represent cluster boundaries identified by unsupervised AHC, while point colors indicate ground truth labels (green: healthy, red: unhealthy).

3.2. Clustering Performance Evaluation

Clustering performance is evaluated using K-Means and AHC algorithms to assess the impact of the embedding technique on unsupervised classification. Figure 6 presents a comparative analysis between the raw test dataset and the embedded dataset, utilizing the Adjusted Rand Index (ARI) and Clustering Accuracy as evaluation metrics. The results highlight the substantial improvements achieved through embedding, demonstrating enhanced separability of the clusters.

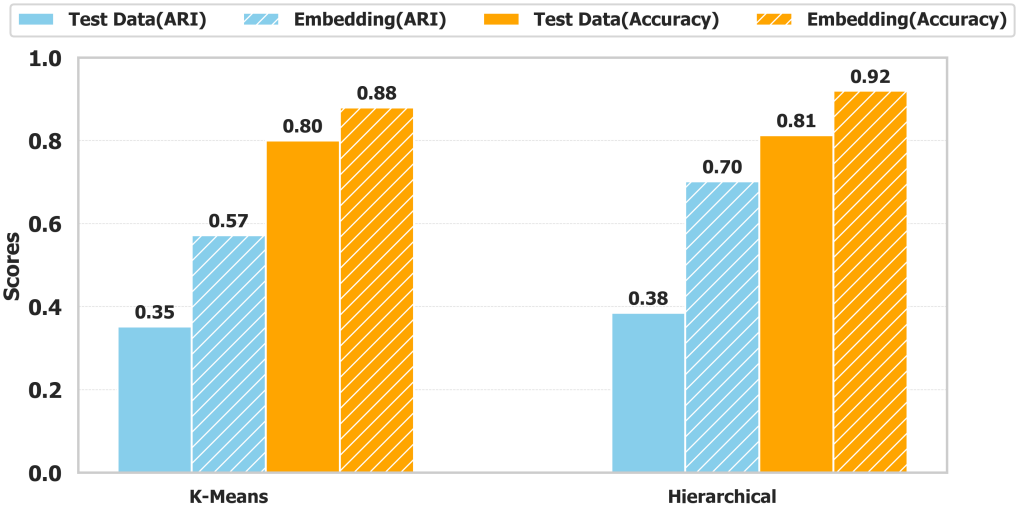

Figure 6. Performance comparison between raw test data and embedding on K-means and AHC.

For K-Means, ARI increases from 0.35 to 0.57, reflecting a 63% improvement, while for AHC, it rises from 0.38 to 0.70, an 84% gain. This enhancement indicates that embedding strengthens structural coherence in the feature space, reducing misclassification. Similarly, clustering accuracy improves notably. K-Means accuracy increases from 0.80 to 0.88, showing a 10% gain, while AHC improves from 0.81 to 0.92, achieving a 14% increase. These results demonstrate that embedding refines cluster compactness and inter-class separability, enabling better discrimination between healthy and unhealthy plots.

By restructuring the data distribution, SSL-based embedding mitigates intra-cluster variance while reinforcing inter-cluster distinctions. This transformation enhances decision boundaries, allowing clustering algorithms to produce more consistent and well-

separated clusters, ultimately improving classification reliability in unsupervised plant disease detection.

3.3. Supervised Classification Performance

To evaluate the efficacy of the proposed SSL method, its performance was benchmarked against supervised learning models, including DNN, RF, and SVM. Hyperparameter configurations were selected to balance model capacity with generalization capability. The Random Forest classifier employed 500 decision trees with unrestricted maximum depth, resulting in an empirical mean depth of 9.1 ± 1.3 across training folds, trained directly on raw features without preprocessing. The SVM utilized an RBF kernel with regularization parameter $C = 2.0$ and scale-based gamma $(1/(n_{features} \times var(X)))$, applied to standardized features (zero mean, unit variance). The DNN implemented with a six-layer fully connected architecture (Figure 4c) with ReLU activation, batch normalization on hidden layers, dropout regularization ($p = 0.5$) on the first two hidden layers, and Adam optimization (learning rate 10^{-3}). DNN training employed early stopping with 25-epoch patience on a validation split (20% of training data), class-weighted cross-entropy loss, and a maximum of 300 epochs with batch size 256. Five-fold stratified cross-validation ensured robust generalization assessment.

As summarized in Table 2, the SSL (AHC) model outperformed all comparators, achieving the highest accuracy (0.92), precision (0.91), and F1-score (0.92), thereby demonstrating enhanced feature extraction and clustering efficacy. The SVM exhibited the highest recall (0.98), indicating superior sensitivity in detecting positive cases. SSL (K-Means) showed comparatively lower performance, highlighting the advantage of AHC feature organization. These findings confirm that SSL-based embeddings yield more discriminative feature representations from unlabeled data, leading to improved generalization. The observed improvements of +3.4% in accuracy, +3.3% in precision, and +2.2% in F1-score compared to DNN and RF underscore SSL's capability to capture structural dependencies effectively. This supports the utility of SSL as a robust alternative to conventional supervised learning, particularly in contexts with limited labeled data.

Table 2. Performance comparison of supervised baseline models.

Model	Accuracy	Precision	Recall	F1-Score
DNN	0.89	0.90	0.89	0.90
RF	0.89	0.88	0.92	0.90
SVM	0.89	0.85	<u>0.98</u>	0.91
SSL (K-Means)	0.88	0.89	0.88	0.88
SSL (AHC)	<u>0.92</u>	<u>0.91</u>	0.92	<u>0.92</u>

Beyond competitive accuracy, SSL's key advantage lies in its annotation requirements. To quantify this, we evaluated supervised baselines trained on progressively reduced labeled subsets (10%, 25%, and 50%). Performance degraded substantially with decreasing labels, with RF accuracy dropping from 89.0% (50% labels) to 84.0% (10% labels), SVM from 89.0% to 86.0%, and DNN from 89.0% to 80.0% (Table S1). These results, averaged across five-fold cross-validation, demonstrate that supervised methods are highly sensitive to the availability of training data. In contrast, SSL achieves comparable accuracy (88–92%, Table 2) while requiring zero labels during training; the ground truth is used only to validate the clusters. This label-free learning addresses a fundamental challenge in early-stage disease detection: when visual symptoms are ambiguous and expert field assessments are resource-intensive, supervised methods struggle with insufficient training labels, whereas SSL learns directly from unlabeled spectral patterns. This operational advantage makes SSL

particularly viable for real-world agricultural monitoring, where annotation bottlenecks limit supervised approaches.

3.4. Plot Level Prediction

Figure 7 presents a comparative analysis of plot-level predictions generated by the SSL and DNN models on a soybean field. To ensure robustness and reliability, predictions were validated using K-fold cross-validation. The field is denoted into distinct plots, each color-coded to visualize prediction accuracy for both models, providing a spatial representation of their performance. This comparative mapping facilitates a comprehensive evaluation of the SSL method relative to a traditional supervised learning approach in assessing soybean plot health.

The results indicate a high degree of agreement between the two models, with both correctly predicting the plot health in 297 instances, as depicted by green plots. However, 47 plots were misclassified by both models, shown in red. Notably, in 16 cases, the SSL model misclassified plots that the DNN correctly predicted (blue), whereas in 15 instances, the DNN misclassified plots that the SSL correctly predicted (yellow). These discrepancies highlight the nuanced differences in model generalization and error patterns. Overall, the SSL model demonstrates performance comparable to the DNN model, despite operating without explicit supervision. This underscores its ability to leverage unlabeled data effectively, making it a viable alternative for classification tasks in scenarios where labeled data is scarce.

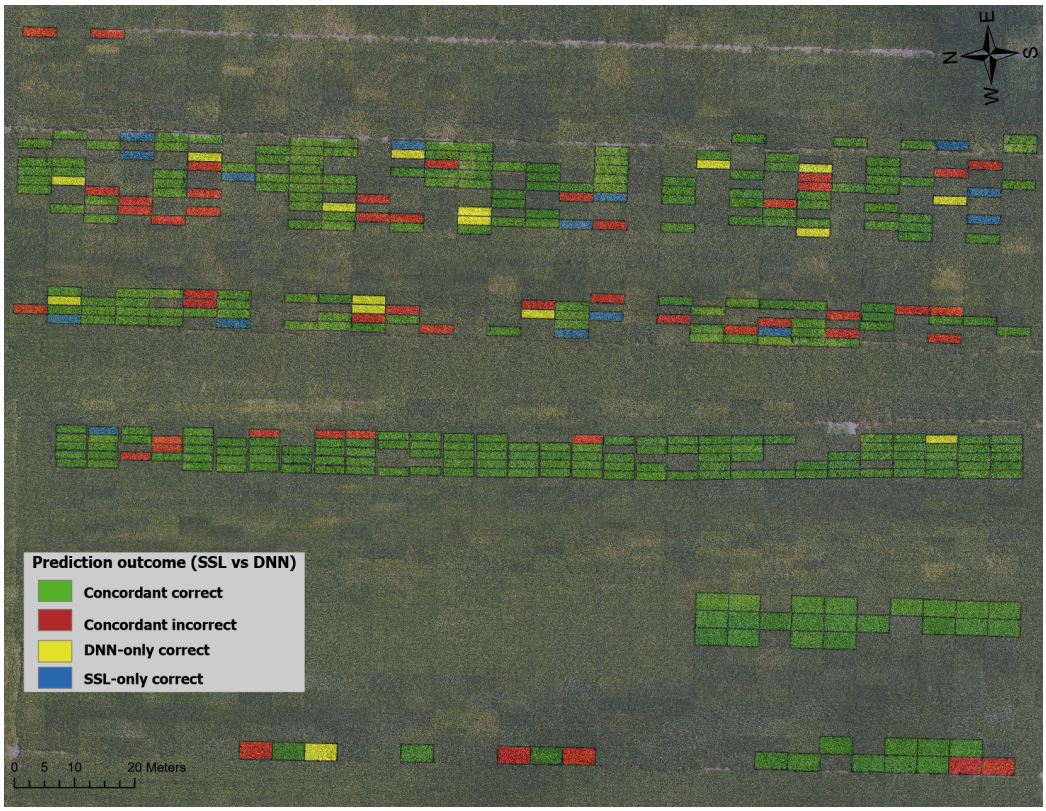

Figure 7. Spatial comparison of plot-level predictions from SSL and DNN models. Concordant correct/incorrect indicate plots where SSL and DNN predictions agree, whereas DNN-only correct and SSL-only correct indicate disagreement where only one model matches the ground truth. Plots are color-coded by prediction category for each model.

4. Discussion

The effectiveness of the proposed SSL framework for early SDS detection stems from its ability to leverage physiologically induced spectral separability without extensive manual annotation. While early-stage SDS produces subtle canopy-level stress signatures before visible symptoms emerge, these pre-visual changes create measurable reflectance shifts that our distance-based pairing strategy can exploit directly from unlabeled data. By learning to discriminate spectral patterns through contrastive representation learning, the framework achieves early detection sensitivity while circumventing the annotation bottleneck that constrains supervised approaches. This is particularly critical for SDS, where ground-truth labels during early infection are expensive and often inconsistent, even among expert observers.

To elucidate the model's decision-making process and identify key features driving soybean disease classification, a SHAP-based feature importance analysis was performed on hyperspectral bands and vegetation indices (VIs). This data-driven approach, unrestricted by predefined spectral categories, revealed the wavelengths and indices most critical for accurate discrimination. Table 3 presents the top-ranked VIs, including $SRPI$, $NPCI$, $WSCT$, ARI_2, and $MCARI_2$. These indices capture distinct physiological responses to SDS infection, such as $WSCT$ detecting water stress from impaired root water uptake, ARI_2 capturing stress-induced anthocyanin accumulation, $MCARI_2$ reflecting chlorophyll degradation from nutrient deficiency, while $NPCI$ and $SRPI$ respond to pigment changes in the early stress stages. These results substantiate that the model's predictions are grounded in physiologically relevant spectral responses specific to SDS pathogenesis.

Table 3. Top 10 most impactful vegetation indices with equations (wavelengths adapted to the sensor's spectral ranges.

Index Name	Equation	Reference
Simple Ratio Pigment Index	$srpi = \frac{R_{430}}{R_{680}}$	[41]
Normalized Pigment Chlorophyll Ratio Index	$npci = \frac{R_{680}-R_{430}}{R_{680}+R_{430}}$	[42]
Water Stress and Canopy Temperature	$wsct = \frac{R_{970}-R_{850}}{R_{970}+R_{850}}$	[43]
Anthocyanin reflectance index 2	$ari_2 = \left(\frac{1}{R_{550}} - \frac{1}{R_{700}}\right) \times R_{800}$	[44]
Modified chlorophyll absorption ratio index 2	$mcari_2 = \frac{1.5\times[2.5\times(R_{800}-R_{670})-1.3\times(R_{800}-R_{550})]}{\sqrt{(2\times R_{800}+1)^2-(6\times R_{800}-5\times\sqrt{R_{670}})-0.5}}$	[45]
Green chlorophyll index	$gci = \frac{R_{800}}{R_{550}} - 1$	[46]
Carotenoid reflectance index	$cri = \frac{1}{R_{550}} - \frac{1}{R_{700}}$	[47]
Modified triangular vegetation index 2	$mtvi_2 = \frac{1.5\times[1.2\times(R_{800}-R_{550})-2.5\times(R_{670}-R_{550})]}{\sqrt{(2\times R_{800}+1)^2-(6\times R_{800}-5\times\sqrt{R_{670}})-0.5}}$	[45]
Photochemical reflectance index	$pri = \frac{R_{531}-R_{570}}{R_{531}+R_{570}}$	[48]
Ratio vegetation index 1	$rvi_1 = \frac{R_{810}}{R_{560}}$	[49]

Figure 8 presents the SHAP attributions for vegetation indices (VIs) and hyperspectral bands. Notably, red-edge (753–771 nm) and near-infrared (924–969 nm) bands were consistently highlighted. These spectral regions are directly linked to SDS pathophysiology, where *Fusarium virguliforme* colonizes soybean roots and produces toxins that disrupt vascular function, impairing nutrient and water transport to the canopy [19]. The red-edge region is particularly sensitive to chlorophyll content degradation resulting from nitrogen deficiency caused by impaired root nutrient uptake [46]. Similarly, NIR reflectance responds to reduced leaf water content and altered mesophyll structure when SDS compromises vascular water transport [50,51]. This correspondence between SHAP-selected features and established physiological stress markers confirms that the model's predictions are

grounded in biologically meaningful spectral responses rather than spurious correlations, reinforcing their diagnostic value for early SDS detection.

The co-occurrence of raw spectral bands and vegetation indices among top-ranked SHAP features indicates that both feature types contribute complementary information to classification performance. While raw bands (red-edge: 753–771 nm, NIR: 924–969 nm) capture direct physiological responses, vegetation indices encode nonlinear transformations (e.g., normalized difference ratios, soil-adjusted formulations) that emphasize specific stress signatures while suppressing confounding factors such as soil background and illumination variability. This suggests that the 316-feature representation (269 bands + 47 indices) provides multiple complementary perspectives on the underlying plant stress state, rather than redundant information. The presence of both feature types in the top-10 discriminative features supports the utility of vegetation index augmentation for enhanced separability in the SSL framework.

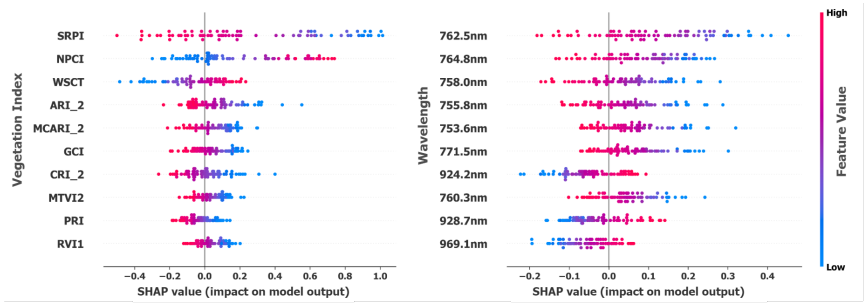

Figure 8. Top 10 SHAP values for Vegetation Index (**left**) and Wavelength (**right**).

Reflectance frequency distributions (Figure 9) support these results by comparing correctly and incorrectly classified plots. Correctly classified samples show clear separation between healthy and diseased states in the red-edge and NIR regions, with healthy plots peaking near 0.09 and 0.40 reflectance, respectively, and diseased plots peaking lower. In contrast, blue and green regions exhibit significant overlap, indicating limited discriminative value. This highlights the importance of red-edge and NIR reflectance in disease detection. Misclassified plots demonstrate spectral overlap and increased variability in the blue, green, and red-edge regions, likely due to label noise and subtle disease progression. Such ambiguity stems from inconsistent ground-truth labels that inadequately reflect continuous spectral changes. Although NIR retains some separability, the increased noise emphasizes the need for uncertainty-aware techniques such as self-supervised denoising or contrastive spectral embeddings to improve feature reliability and reduce misclassification in UAV-based disease monitoring.

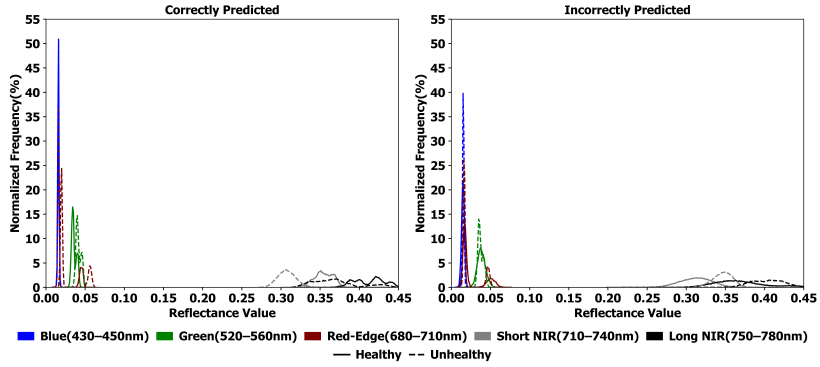

Figure 9. Normalized reflectance distributions for correctly (**left**) and incorrectly (**right**) classified plots, illustrating strong separation between healthy and diseased states in red-edge and NIR bands (correct prediction), but overlap in incorrect prediction.

5. Conclusions

In this study, we developed a self-supervised learning (SSL) framework for the early detection of Sudden Death Syndrome (SDS) in soybean plants using UAV-based hyperspectral data. To enable effective representation learning from plot-level reflectance features, we introduced a distance-based pairing strategy and a modified contrastive loss function tailored to spectral data. The key contributions can be summarized as follows:

- An end-to-end SSL framework for SDS detection using UAV hyperspectral data, reducing reliance on in-situ measurements.
- A distance-based spectral pairing strategy that enhances cluster separability and strengthens feature learning.
- Demonstrated performance gains of 11% over unsupervised methods and 3% over traditional supervised learning, highlighting the efficacy of SSL for hyperspectral plant disease detection.

Model interpretability analysis using SHAP confirmed that the red-edge (753–771 nm) and NIR (924–969 nm) regions were the most critical spectral domains, consistent with known physiological stress markers. Reflectance-based separability further reinforced the diagnostic value of these bands, while errors were linked to spectral overlap and label noise. These findings demonstrate that SSL, coupled with targeted spectral feature selection, provides a scalable, interpretable, and label-efficient approach for precision agriculture.

While this study validates SSL on 375 plots from a single field site during the 2022 growing season, the framework's design principles support broader applicability. The distance-based pairing strategy operates on standardized spectral features and employs adaptive quantile-based thresholds that instinctively adjust to data distribution, enabling methodological transferability across datasets. The reliance on universal physiological markers, red-edge and NIR bands corresponding to chlorophyll degradation and water stress, provides a biophysically grounded foundation that transcends site-specific characteristics. Cross-validation performance and the elimination of manual threshold tuning further demonstrate the framework's robustness within the studied conditions. Nevertheless, systematic validation across diverse environmental contexts would strengthen operational confidence. Specifically, extending the framework to different soil types (Alfisols, Vertisols) would verify that the soil masking approach (Section 2.3) generalizes across pedological contexts. Multi-season deployment across varying phenological windows (V4–R1 early stages, R7–R8 late stages) would demonstrate temporal robustness under different climatic conditions and canopy architectures. Evaluating performance across broader disease severity ranges from pre-visual infection to severe foliar damage (>50%) would establish whether the distance thresholds ($\alpha = 0.74$, $\beta = 1.51$) generalize or require dataset-specific recalibration. Future work will focus on three key directions: (i) incorporating multi-temporal hyperspectral data to improve robustness and enable disease progression monitoring; (ii) extending the framework across diverse crops and environmental conditions to enhance generalizability; and (iii) integrating advanced SSL architectures and multi-modal fusion techniques to refine feature learning and improve classification performance across different disease stages.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs17233928/s1.

Author Contributions: Conceptualization, M.R. and F.A.L.; methodology, M.R. and F.A.L.; software, M.R.; validation, M.R., F.A.L. and H.A. (Haireti Alifu) ; formal analysis, M.R.; investigation, M.R. and V.S.; resources, V.S.; data curation, M.R. and C.G.; writing—original draft preparation, M.R.; writing—review and editing, M.R., V.S., F.A.L., H.A. (Haireti Alifu), K.P. and H.A. (Hadi Aliakbarpour); visualization, M.R.; supervision, V.S.; project administration, V.S. and K.P.;

funding acquisition, V.S. and K.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the U.S. Army Corps of Engineers, Engineering Research and Development Center—Information Technology Laboratory (ERDC-ITL) under Contract W912HZ23C0041 and United Soybean Board grant 2431-201-0101. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Government or agency thereof.

Data Availability Statement: Sample data are available upon request to the corresponding author.

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

References

1. Singh, B.K.; Delgado-Baquerizo, M.; Egidi, E.; Guirado, E.; Leach, J.E.; Liu, H.; Trivedi, P. Climate change impacts on plant pathogens, food security and paths forward. *Nat. Rev. Microbiol.* **2023**, *21*, 640–656. [CrossRef]
2. Bebber, D.P.; Ramotowski, M.A.T.; Gurr, S.J. Crop pests and pathogens move polewards in a warming world. *Nat. Clim. Change* **2013**, *3*, 985–988. [CrossRef]
3. Garrett, K.A.; Dendy, S.P.; Frank, E.E.; Rouse, M.N.; Travers, S.E. Climate Change Effects on Plant Disease: Genomes to Ecosystems. *Annu. Rev. Phytopathol.* **2006**, *44*, 489–509. [CrossRef] [PubMed]
4. Castelao Tetila, E.; Brandoli Machado, B.; Belete, N.A.D.S.; Guimaraes, D.A.; Pistori, H. Identification of Soybean Foliar Diseases Using Unmanned Aerial Vehicle Images. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2190–2194. [CrossRef]
5. Voora, V.; Bermúdez, S.; Le, H.; Larrea, C.; Luna, E. *Global Market Report: Soybean Prices and Sustainability*; The International Institute for Sustainable Development: Winnipeg, MB, Canada, 2024.
6. Brodbeck, C.; Sikora, E.; Delaney, D.; Pate, G.; Johnson, J. Using Unmanned Aircraft Systems for Early Detection of Soybean Diseases. *Adv. Anim. Biosci.* **2017**, *8*, 802–806. [CrossRef]
7. Bradley, C.A.; Allen, T.W.; Sisson, A.J.; Bergstrom, G.C.; Bissonnette, K.M.; Bond, J.; Byamukama, E.; Chilvers, M.I.; Collins, A.A.; Damicone, J.P.; et al. Soybean Yield Loss Estimates Due to Diseases in the United States and Ontario, Canada, from 2015 to 2019. *Plant Health Prog.* **2021**, *22*, 483–495. [CrossRef]
8. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using Deep Learning for Image-Based Plant Disease Detection. *Front. Plant Sci.* **2016**, *7*, 01419. [CrossRef] [PubMed]
9. Trippa, D.; Scalenghe, R.; Basso, M.F.; Panno, S.; Davino, S.; Morone, C.; Giovino, A.; Oufensou, S.; Luchi, N.; Yousefi, S.; et al. Next-generation methods for early disease detection in crops. *Pest Manag. Sci.* **2024**, *80*, 245–261. [CrossRef] [PubMed]
10. Chai, A.Y.H.; Lee, S.H.; Tay, F.S.; Bonnet, P.; Joly, A. Beyond supervision: Harnessing self-supervised learning in unseen plant disease recognition. *Neurocomputing* **2024**, *610*, 128608. [CrossRef]
11. Yilma, G.; Dagne, M.; Ahmed, M.K.; Bellam, R.B. Attentive Self-supervised Contrastive Learning (ASCL) for plant disease classification. *Results Eng.* **2025**, *25*, 103922. [CrossRef]
12. Nguyen, C.; Sagan, V.; Maimaitiyiming, M.; Maimaitijiang, M.; Bhadra, S.; Kwasniewski, M.T. Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors* **2021**, *21*, 742. [CrossRef]
13. Xu, Z.; Zhang, Q.; Xiang, S.; Li, Y.; Huang, X.; Zhang, Y.; Zhou, X.; Li, Z.; Yao, X.; Li, Q.; et al. Monitoring the severity of Pantana phyllostachysae Chao infestation in Moso bamboo forests based on UAV multi-spectral remote sensing feature selection. *Forests* **2022**, *13*, 418. [CrossRef]
14. Terentev, A.; Dolzhenko, V.; Fedotov, A.; Eremenko, D. Current state of hyperspectral remote sensing for early plant disease detection: A review. *Sensors* **2022**, *22*, 757. [CrossRef]
15. Zhang, Z.; Huang, L.; Wang, Q.; Jiang, L.; Qi, Y.; Wang, S.; Shen, T.; Tang, B.H.; Gu, Y. UAV Hyperspectral Remote Sensing Image Classification: A Systematic Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *18*, 3099–3124. [CrossRef]
16. Benfenati, A.; Causin, P.; Oberti, R.; Stefanello, G. Unsupervised deep learning techniques for powdery mildew recognition based on multispectral imaging. *arXiv* **2021**, arXiv:2112.11242. [CrossRef]
17. Ma, G.; Javidan, S.M.; Ampatzidis, Y.; Zhang, Z. A Novel Hybrid Technique for Detecting and Classifying Hyperspectral Images of Tomato Fungal Diseases Based on Deep Feature Extraction and Manhattan Distance. *Sensors* **2025**, *25*, 4285. [CrossRef]
18. Ge, Z.; Fan, X.; Zhang, J.; Jin, S. SegPPD-FS: Segmenting Plant Pests and Diseases in the Wild Using Few-shot Learning. *Plant Phenomics* **2025**, 100121. [CrossRef]
19. Spampinato, C.P.; Scandiani, M.M.; Luque, A.G. Soybean sudden death syndrome: Fungal pathogenesis and plant response. *Plant Pathol.* **2021**, *70*, 3–12. [CrossRef]
20. Bröker, F.; Holt, L.L.; Roads, B.D.; Dayan, P.; Love, B.C. Demystifying unsupervised learning: How it helps and hurts. *Trends Cogn. Sci.* **2024**, *28*, 974–986. [CrossRef] [PubMed]

21. Peng, D.; Gui, Z.; Wu, H. Interpreting the curse of dimensionality from distance concentration and manifold effect. *arXiv* **2023**, arXiv:2401.00422.
22. Wu, Z.; Xiong, Y.; Yu, S.X.; Lin, D. Unsupervised feature learning via non-parametric instance discrimination. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3733–3742.
23. Liu, L.; Zhang, H.; Wang, Y. Contrastive Mutual Learning with Pseudo-Label Smoothing for Hyperspectral Image Classification. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 2520314. [CrossRef]
24. Jing, L.; Tian, Y. Self-Supervised Visual Feature Learning with Deep Neural Networks: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 4037–4058. [CrossRef]
25. Jaiswal, A.; Babu, A.R.; Zadeh, M.Z.; Banerjee, D.; Makedon, F. A Survey on Contrastive Self-Supervised Learning. *Technologies* **2020**, *9*, 2. [CrossRef]
26. Grill, J.B.; Strub, F.; Altché, F.; Tallec, C.; Richemond, P.; Buchatskaya, E.; Doersch, C.; Avila Pires, B.; Guo, Z.; Gheshlaghi Azar, M.; et al. Bootstrap your own latent-a new approach to self-supervised learning. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 21271–21284.
27. Chen, X.; He, K. Exploring simple Siamese representation learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 15750–15758.
28. Chicco, D. Siamese Neural Networks: An Overview. In *Artificial Neural Networks*; Cartwright, H., Ed.; Methods in Molecular Biology; Humana: New York, NY, USA, 2021; Volume 2190, pp. 73–94. [CrossRef]
29. Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G. A simple framework for contrastive learning of visual representations. In Proceedings of the International Conference on Machine Learning, Virtual, 13–18 July 2020; pp. 1597–1607.
30. Güldenring, R.; Nalpantidis, L. Self-supervised contrastive learning on agricultural images. *Comput. Electron. Agric.* **2021**, *191*, 106510. [CrossRef]
31. Monowar, M.M.; Hamid, M.A.; Kateb, F.A.; Ohi, A.Q.; Mridha, M.F. Self-Supervised Clustering for Leaf Disease Identification. *Agriculture* **2022**, *12*, 814. [CrossRef]
32. Zhao, X.; Wang, L.; Zhang, Y.; Han, X.; Deveci, M.; Parmar, M. A review of convolutional neural networks in computer vision. *Artif. Intell. Rev.* **2024**, *57*, 99. [CrossRef]
33. Kalapos, A.; Gyires-Tóth, B. CNN-JEPA: Self-Supervised Pretraining Convolutional Neural Networks Using Joint Embedding Predictive Architecture. *arXiv* **2024**, arXiv:2408.07514.
34. Qin, Y.; Ye, Y.; Zhao, Y.; Wu, J.; Zhang, H.; Cheng, K.; Li, K. Nearest neighboring self-supervised learning for hyperspectral image classification. *Remote Sens.* **2023**, *15*, 1713. [CrossRef]
35. Jarocińska, A.; Kopeć, D.; Kycko, M. Comparison of dimensionality reduction methods on hyperspectral images for the identification of heathlands and mires. *Sci. Rep.* **2024**, *14*, 27662. [CrossRef]
36. Sagan, V.; Maimaitijiang, M.; Paheding, S.; Bhadra, S.; Gosselin, N.; Burnette, M.; Demieville, J.; Hartling, S.; LeBauer, D.; Newcomb, M.; et al. Data-Driven Artificial Intelligence for Calibration of Hyperspectral Big Data. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5510320. [CrossRef]
37. He, K.; Fan, H.; Wu, Y.; Xie, S.; Girshick, R. Momentum contrast for unsupervised visual representation learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 9729–9738.
38. McQueen, J.B. Some methods of classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*; University of California: Berkeley, CA, USA, 1967; pp. 281–297.
39. Ward Jr, J.H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. [CrossRef]
40. Kuhn, H.W. The Hungarian method for the assignment problem. *Nav. Res. Logist. Q.* **1955**, *2*, 83–97. [CrossRef]
41. Peñuelas, J.; Filella, I.; Lloret, P.; Muñoz, F.; Vilajeliu, M. Reflectance assessment of mite effects on apple trees. *Int. J. Remote Sens.* **1995**, *16*, 2727–2733. [CrossRef]
42. Peñuelas, J.; Gamon, J.; Fredeen, A.; Merino, J.; Field, C. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sens. Environ.* **1994**, *48*, 135–146. [CrossRef]
43. Babar, M.; Reynolds, M.; Van Ginkel, M.; Klatt, A.; Raun, W.; Stone, M. Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. *Crop Sci.* **2006**, *46*, 1046–1057. [CrossRef]
44. Gitelson, A.A.; Merzlyak, M.N.; Chivkunova, O.B. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochem. Photobiol.* **2001**, *74*, 38–45. [CrossRef] [PubMed]
45. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [CrossRef]
46. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [CrossRef]
47. Gitelson, A.A.; Zur, Y.; Chivkunova, O.B.; Merzlyak, M.N. Assessing carotenoid content in plant leaves with reflectance spectroscopy. *Photochem. Photobiol.* **2002**, *75*, 272–281. [CrossRef]

48. Gamon, J.; Penuelas, J.; Field, C. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [CrossRef]
49. Birth, G.S.; McVey, G.R. Measuring the color of growing turf with a reflectance spectrophotometer. *Agron. J.* **1968**, *60*, 640–643. [CrossRef]
50. Carter, G.A. Primary and secondary effects of water content on the spectral reflectance of leaves. *Am. J. Bot.* **1991**, *78*, 916–924. [CrossRef]
51. Peñuelas, J.; Pinol, J.; Ogaya, R.; Filella, I. The reflectance at the 950–970 nm region as an indicator of plant water status. *Int. J. Remote Sens.* **1997**, *18*, 2869–2875. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.