AgriEngineering

Article

Agri-DSSA: A Dual Self-Supervised Attention Framework for Multisource Crop Health Analysis Using Hyperspectral and Image-Based Benchmarks

Fatema A. Albalooshi

College of Information Technology, University of Bahrain, Sakhir P.O. Box 32038, Bahrain; falbalooshi@uob.edu.bh

Abstract

Recent advances in hyperspectral imaging (HSI) and multimodal deep learning have opened new opportunities for crop health analysis; however, most existing models remain limited by dataset scope, lack of interpretability, and weak cross-domain generalization. To overcome these limitations, this study introduces Agri-DSSA, a novel Dual Self-Supervised Attention (DSSA) framework that simultaneously models spectral and spatial dependencies through two complementary self-attention branches. The proposed architecture enables robust and interpretable feature learning across heterogeneous data sources, facilitating the estimation of spectral proxies of chlorophyll content, plant vigor, and disease stress indicators rather than direct physiological measurements. Experiments were performed on seven publicly available benchmark datasets encompassing diverse spectral and visual domains: three hyperspectral datasets (Indian Pines with 16 classes and 10,366 labeled samples; Pavia University with 9 classes and 42,776 samples; and Kennedy Space Center with 13 classes and 5211 samples), two plant disease datasets (PlantVillage with 54,000 labeled leaf images covering 38 diseases across 14 crop species, and the New Plant Diseases dataset with over 30,000 field images captured under natural conditions), and two chlorophyll content datasets (the Global Leaf Chlorophyll Content Dataset (GLCC), derived from MERIS and OLCI satellite data between 2003–2020, and the Leaf Chlorophyll Content Dataset for Crops, which includes paired spectrophotometric and multispectral measurements collected from multiple crop species). To ensure statistical rigor and spatial independence, a block-based spatial cross-validation scheme was employed across five independent runs with fixed random seeds. Model performance was evaluated using R^2, RMSE, F1-score, AUC-ROC, and AUC-PR, each reported as mean ± standard deviation with 95% confidence intervals. Results show that Agri-DSSA consistently outperforms baseline models (PLSR, RF, 3D-CNN, and HybridSN), achieving up to $R^2 = 0.86$ for chlorophyll content estimation and F1-scores above 0.95 for plant disease detection. The attention distributions highlight physiologically meaningful spectral regions (550–710 nm) associated with chlorophyll absorption, confirming the interpretability of the model's learned representations. This study serves as a methodological foundation for UAV-based and field-deployable crop monitoring systems. By unifying hyperspectral, chlorophyll, and visual disease datasets, Agri-DSSA provides an interpretable and generalizable framework for proxy-based vegetation stress estimation. Future work will extend the model to real UAV campaigns and in-field spectrophotometric validation to achieve full agronomic reliability.

Keywords: hyperspectral imaging (HSI); precision agriculture; deep learning; dual attention mechanism; spectral–spatial feature extraction; chlorophyll content estimation; crop disease detection; remote sensing; pseudo-ground truth generation

check for updates

Academic Editor: Ray E. Sheriff

Received: 22 August 2025
Revised: 8 October 2025
Accepted: 15 October 2025
Published: 17 October 2025

Citation: Albalooshi, F.A. Agri-DSSA: A Dual Self-Supervised Attention Framework for Multisource Crop Health Analysis Using Hyperspectral and Image-Based Benchmarks. *AgriEngineering* **2025**, *7*, 350. https://doi.org/10.3390/agriengineering7100350

Copyright: © 2025 by the author. 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

The global demand for food is escalating rapidly, driven by a burgeoning population and changing dietary patterns. This increasing demand places immense pressure on agricultural systems to enhance productivity while simultaneously minimizing environmental impact. Traditional agricultural practices, often reliant on broad-scale interventions, are increasingly proving inefficient and unsustainable. In response, precision agriculture has emerged as a paradigm shift, advocating for site-specific management based on real-time data and advanced analytical techniques [1,2]. The core objective of precision agriculture is to optimize resource utilization—such as water, fertilizers, and pesticides—and improve crop yields by addressing variability within fields.

Central to the success of precision agriculture is the accurate and timely monitoring of crop health. Conventional methods for assessing crop health, including visual inspection and destructive laboratory analyses, are labor-intensive, time-consuming, and often provide delayed insights, limiting their utility for proactive management [3]. The advent of remote sensing technologies, particularly hyperspectral imaging (HSI), has revolutionized the ability to acquire detailed information about crop physiological and biochemical properties non-invasively. HSI captures electromagnetic radiation across a continuous spectrum of narrow, contiguous bands, providing a unique spectral signature for each pixel. This rich spectral information allows for the discrimination of subtle variations in plant health, nutrient status, and disease presence that are imperceptible to the human eye or conventional multispectral sensors [4].

Despite the immense potential of HSI in agriculture, its widespread adoption has been hampered by several challenges. Hyperspectral datasets are characterized by high dimensionality, often comprising hundreds of spectral bands. This high dimensionality leads to significant computational complexity, data redundancy, and the "curse of dimensionality," which can degrade the performance of traditional machine learning algorithms [5,6]. Furthermore, the intricate nonlinear relationships between spectral signatures and various crop health indicators necessitate advanced analytical tools capable of extracting meaningful features from this complex data.

In recent years, deep learning (DL) has demonstrated unprecedented success in various computer vision tasks [7], including image classification, object detection, and semantic segmentation. The ability of DL models, particularly Convolutional Neural Networks (CNNs), to automatically learn hierarchical features from raw data makes them exceptionally well-suited for processing complex image data like HSI [8]. However, standard CNN architectures may not fully leverage the spectral richness of HSI data, often treating spectral bands as independent channels or failing to capture inter-band dependencies effectively. This limitation has led to the exploration of more sophisticated DL architectures, including those incorporating attention mechanisms.

Attention mechanisms, inspired by human cognitive processes, enable neural networks to selectively focus on the most relevant parts of the input data. When applied to HSI, attention mechanisms can dynamically weigh the importance of different spectral bands or spatial regions, thereby enhancing the model's ability to extract salient features related to crop health [9]. This selective focus enhances predictive accuracy while also providing greater interpretability, enabling researchers to pinpoint the specific wavelengths or spatial patterns that are most indicative of certain plant conditions.

Nevertheless, the majority of existing attention-based approaches rely on limited datasets or simulated scenes and lack cross-domain generalization between hyperspectral and RGB imagery. Moreover, few studies explicitly address the challenge of proxy-based training—where vegetation indices such as NDVI or MCARI are used as stand-ins for real

chlorophyll or disease measurements—leaving a gap between methodological advances and field-deployable agronomic validation.

To address this limitation, the present study introduces Agri-DSSA, a Dual Self-Supervised Attention framework designed for multimodal crop health analysis across hyperspectral and image-based datasets. The Agri-DSSA architecture integrates two complementary attention branches—spectral and spatial—that cooperatively learn discriminative features and highlight physiologically meaningful regions. Unlike conventional supervised approaches, the framework employs a self-supervised pretraining strategy to enhance generalization under limited labeled data conditions, thereby improving robustness across diverse datasets and sensing modalities.

This work focuses on benchmarking and methodological validation rather than field-deployed UAV acquisition. Seven publicly available datasets were utilized for evaluation: three hyperspectral datasets (Indian Pines, Pavia University, Kennedy Space Center), two plant disease datasets (PlantVillage and New Plant Diseases), and two chlorophyll content datasets (GLCC and the Leaf Chlorophyll Content Dataset for Crops). These datasets collectively represent a diverse set of spectral and spatial conditions, forming a comprehensive foundation for cross-domain learning and proxy-based validation.

A major challenge in agricultural remote sensing is the scarcity of large-scale, co-registered field data with direct biochemical or pathological ground truth. As a result, the present study emphasizes learning from spectral proxies and well-curated public benchmarks as a precursor to real agronomic validation.

The main contributions of this research are summarized as follows:

- Proposes Agri-DSSA, a novel Dual Self-Supervised Attention framework that integrates spectral and spatial attention mechanisms for multimodal (HSI and RGB) crop health analysis.
- Introduces a unified training and evaluation pipeline across seven benchmark datasets, encompassing hyperspectral, chlorophyll, and plant disease domains, to assess model generalization and robustness.
- Presents comprehensive ablation studies and spatial cross-validation to quantify the contribution of each module and ensure statistically robust evaluation (mean $\pm$ SD, 95% CI).
- Provides visual interpretability through spectral and spatial attention maps, demonstrating alignment with known pigment absorption regions (550–710 nm).
- Establishes Agri-DSSA as a methodological foundation for future UAV and in-field deployment, enabling scalable integration of multimodal spectral–spatial learning for precision agriculture.

The remainder of this paper is organized as follows: Section 2 reviews related work in the field. Section 3 describes the proposed methodology, including the data preprocessing pipeline, model architecture, and experimental setup. Section 4 presents and analyzes the experimental results. Section 5 provides a detailed discussion of the findings and their implications. Finally, Section 6 concludes the paper and outlines potential directions for future research.

2. Related Work

Hyperspectral imaging (HSI) has emerged as a pivotal technology in precision agriculture, offering unparalleled capabilities for detailed crop monitoring and management. Unlike traditional RGB or multispectral sensors, HSI captures data across hundreds of narrow, contiguous spectral bands, providing a unique spectral fingerprint for various plant components and conditions, as noted by Vinod et al. [10]. This rich spectral information enables the detection of subtle changes in plant physiology, biochemistry, and morphology,

which are often indicative of stress, disease, or nutrient deficiencies long before visible symptoms appear, as reported by Nguyen et al. [11].

Early applications of HSI in agriculture primarily focused on basic crop classification and yield prediction. Researchers, such as Ang and Seng [12], utilized statistical methods and traditional machine learning algorithms, including Support Vector Machines (SVM) and Random Forests (RF), to analyze hyperspectral data. While these methods demonstrated promising results, they often struggled with the high dimensionality and inherent non-linearity of HSI data, requiring extensive feature engineering and dimensionality reduction techniques. The curse of dimensionality posed a significant challenge, as the number of training samples required for robust model performance increases exponentially with the number of features—a common limitation in HSI datasets.

The advent of deep learning (DL) marked a significant turning point in HSI analysis. Ma and Liu [13] highlighted that Deep Neural Networks (DNNs), particularly Convolutional Neural Networks (CNNs), have proven highly effective in automatically extracting hierarchical features from raw HSI data, circumventing the need for manual feature engineering. El Sakka et al. [14] present a comprehensive review of Convolutional Neural Network (CNN) applications in smart agriculture, with a particular focus on multimodal data integration. By analyzing over 115 studies alongside a large-scale bibliometric survey, the authors highlight CNNs' effectiveness across key domains such as weed and disease detection, crop classification, water management, and yield prediction. The review emphasizes CNNs' superiority over traditional machine learning methods, especially when leveraging diverse data sources like RGB, multispectral, radar, and thermal imagery from UAVs and satellites. Furthermore, the study discusses advanced architectures, including 3D CNNs and hybrid CNN–LSTM/ViT models, which enhance performance in complex agricultural tasks. Looking forward, the authors outline future research directions involving IoT integration, Large Language Models, AR/VR applications, and stronger data security, positioning CNNs as pivotal tools for the advancement of sustainable and intelligent agriculture. For instance, Zhou et al. [15] showed that 3D-CNNs can effectively capture the intricate relationships between spectral bands and spatial contexts, leading to improved classification accuracies in complex agricultural scenes.

Recent studies have demonstrated the effectiveness of deep learning approaches in agricultural disease detection and classification. Ref. [16] conducted a comparative study using pre-trained deep learning models for papaya leaf disease classification, evaluating various architectures to determine optimal performance for agricultural applications.

Building on this foundation, ref. [17] introduced PapNet, an AI-driven approach specifically designed for early detection and classification of papaya leaf diseases. Their methodology emphasized the importance of timely intervention in preventing crop losses.

Extending beyond leaf diseases to fruit quality assessment, ref. [18] developed PlmNet with transfer learning for time-sensitive bruise detection in plums. This research highlighted the critical role of temporal factors in post-harvest quality monitoring and demonstrated how transfer learning can enhance detection accuracy while reducing computational requirements.

Despite the successes of CNNs, the sheer volume and complexity of hyperspectral data still present challenges. Standard CNNs often treat all spectral bands equally, potentially overlooking the varying importance of different wavelengths for specific tasks. This limitation has spurred interest in integrating attention mechanisms into deep learning models for HSI.

Several studies have explored the application of attention mechanisms in HSI for agricultural purposes. For example, Ye et al. [19] utilized a spectral attention module within a 1D-CNN framework to predict lettuce chlorophyll content, demonstrating that the

attention mechanism significantly improved the model's ability to learn the importance of specific hyperspectral bands. Similarly, Kumar et al. [20] employed spatial attention mechanisms to highlight critical regions within the image, such as diseased plant areas, leading to more accurate and localized detection. Pande and Banerjee [21] proposed hybrid attention mechanisms combining both spectral and spatial attention, which have shown promise in capturing comprehensive contextual information from HSI data, leading to state-of-the-art performance in various crop classification and health monitoring tasks.

While significant progress has been made, there remains a need for novel attention-based deep learning architectures that can more effectively integrate spectral and spatial information, reduce computational complexity, and enhance the interpretability of results for real-world precision agriculture applications. Building upon the existing body of literature, the proposed Agri-DSSA model directly addresses several critical research gaps and offers innovative solutions to current challenges in hyperspectral image analysis for precision agriculture. Unlike traditional methods that rely on explicit dimensionality reduction, Agri-DSSA leverages a spectral attention mechanism to adaptively learn and emphasize the most informative spectral bands. This intrinsic feature selection process effectively mitigates the curse of dimensionality and spectral redundancy, leading to more efficient and accurate models without loss of critical information. Moreover, while many deep learning models utilize 3D CNNs for combined spectral–spatial feature extraction, our dual-attention module (spectral and spatial attention) provides a more refined and targeted approach. This allows the model to not only learn features but also dynamically focus on the most relevant spectral and spatial cues, leading to superior performance in complex agricultural scenarios.

3. Materials and Methods

This section details the proposed attention-based deep learning model for high-resolution hyperspectral image analysis in crop health monitoring. The data acquisition and preprocessing steps, the architecture of the attention-based deep learning model, and the training and evaluation procedures are described in the following subsections.

3.1. Data Acquisition

To ensure the robustness, diversity, and generalizability of the proposed Agri-DSSA framework, we employed a comprehensive set of publicly available datasets encompassing both hyperspectral benchmarks and physiologically validated plant datasets. This combination allows for a dual evaluation—assessing the model's methodological performance on hyperspectral imagery and its biological relevance through real-world chlorophyll and disease data.

(a) Hyperspectral Benchmark Datasets

Three widely used hyperspectral datasets from the remote sensing community were utilized: Indian Pines [22], Pavia University [23], and Kennedy Space Center (KSC) [24]. These datasets represent heterogeneous land-cover classes, vegetation types, and agricultural structures, providing a diverse and challenging testbed for evaluating the spectral–spatial learning capability of Agri-DSSA. Each dataset contains hundreds of spectral bands covering the visible to near-infrared range, enabling fine-grained analysis of crop and vegetation reflectance properties. Figure 1 illustrates representative samples from each dataset, highlighting their variability in scene composition and spectral characteristics. Figure 1 presents sample images from each of the three datasets.



(**a**) (**b**) (**c**)

Figure 1. Sample images from three datasets. (**a**) The Indian Pines dataset showcases various agricultural fields and natural vegetation. (**b**) The Pavia University dataset features urban areas mixed with agricultural fields. (**c**) The Kennedy Space Center dataset illustrates diverse vegetation types and land cover.

1. Indian Pines Dataset [22]: This dataset was gathered by the AVIRIS sensor over the Indian Pines test site in Northwestern Indiana. It consists of 145 × 145 pixels and 220 spectral reflectance bands in the wavelength range 0.4–2.5 µm. The dataset primarily features agricultural crops, forests, and other natural perennial vegetation. The Indian Pines hyperspectral dataset ground truth [22] is a labeled land-cover map that assigns each pixel in the hyperspectral image to one of 16 predefined classes.
2. Pavia University Dataset [23]: Acquired by the ROSIS sensor over Pavia, northern Italy. This dataset comprises 1096 × 715 pixels and 103 spectral bands, covering the wavelength range 0.43–0.86 µm. It includes urban areas with a significant presence of vegetation, making it suitable for evaluating the model's performance in mixed land-cover scenarios.
3. Kennedy Space Center (KSC) Dataset [24]: This dataset was collected by the AVIRIS sensor over the Kennedy Space Center, Florida. It contains 512 × 614 pixels and 224 spectral bands (reduced to 176 after removing water absorption and low SNR bands). The KSC site is dominated by various types of wetland vegetation, providing a challenging environment for species classification and health monitoring.

(b) Integration of Physiological Datasets

To extend the evaluation beyond methodological testing and incorporate biologically meaningful ground truth, additional datasets related to plant disease detection and chlorophyll content estimation were integrated. These datasets include verified physiological and pathological annotations, allowing the proposed model to be validated against real agronomic indicators of plant health and stress.

(1) Plant Disease Detection Datasets:

To address the lack of physiological ground truth in hyperspectral benchmarks, two large-scale and publicly available datasets commonly used for plant pathology research were added:

- PlantVillage Dataset [25]: Contains over 54,000 expertly labeled images of healthy and diseased plant leaves from 14 crop species and 38 disease classes. Each image was collected under controlled laboratory conditions, providing a clean, well-annotated dataset ideal for baseline model training and disease classification benchmarking.
- New Plant Diseases Dataset [26]: Extends the original PlantVillage dataset with high-resolution leaf images captured under natural field conditions and variable illumination. It encompasses multiple crop species (e.g., apple, corn, grape, potato, tomato, and wheat), covering both healthy and diseased classes. Its environmental variability introduces real-world complexity, making it valuable for evaluating model generalization and robustness beyond controlled settings.

(2) Chlorophyll Content Estimation Datasets:

To incorporate datasets with real chlorophyll concentration measurements, two complementary resources combining remote sensing and spectrophotometric data were utilized:

- Global Leaf Chlorophyll Content Dataset (GLCC) [27]: Provides global-scale chlorophyll estimates derived from MERIS (Envisat) and OLCI (Sentinel-3) satellite observations between 2003–2020. The dataset includes validation against extensive in situ field measurements across multiple vegetation types, offering a robust reference for evaluating large-scale chlorophyll retrieval accuracy.
- Leaf Chlorophyll Content Dataset for Crops [28,29]: Combines spectrophotometric chlorophyll measurements with co-registered multispectral imagery collected across diverse crop species under varying physiological conditions. This dataset bridges the gap between laboratory-based chlorophyll quantification and image-based spectral analysis, making it ideal for training and validating predictive models linking spectral indices to actual biochemical parameters.

The integration of these six datasets ensures that Agri-DSSA is evaluated across multiple dimensions—ranging from synthetic spectral benchmarks to field-verified biological indicators—thereby validating both its methodological soundness and its agronomic applicability.

3.2. Preprocessing of Hyperspectral Datasets

To ensure the reliability and consistency of the experimental results, standardized and comprehensive preprocessing was applied to all three hyperspectral datasets. These steps aimed to improve data quality, reduce redundancy, and prepare the inputs for deep learning-based analysis, as described in the following subsections.

3.2.1. Pseudo-Ground Truth Generation for Chlorophyll Content

The Indian Pines hyperspectral dataset does not provide direct chlorophyll content measurements. To enable chlorophyll estimation experiments, a pseudo-ground truth map was generated using vegetation indices sensitive to chlorophyll absorption features. First, non-vegetation pixels were excluded using the provided land-cover ground truth labels, retaining only agricultural and vegetation-related classes. Hyperspectral bands affected by water absorption noise were removed to improve spectral quality.

From the remaining spectral bands, reflectance values were extracted at wavelengths corresponding to the green (≈550 nm), red (≈670 nm), red-edge (≈700 nm), and near-infrared (≈800 nm) regions, which are known to be correlated with chlorophyll content. Three commonly used vegetation indices sensitive to chlorophyll were computed for each pixel: the Normalized Difference Vegetation Index (NDVI, calculated as in Equation (1)), the Modified Chlorophyll Absorption Ratio Index (MCARI, calculated as in Equation (2)), and the Transformed Chlorophyll Absorption Ratio Index (TCARI, calculated as in Equation (3)).

$$\mathrm{NDVI} = \frac{R_{\mathrm{NIR}} - R_{\mathrm{Red}}}{R_{\mathrm{NIR}} + R_{\mathrm{Red}}}, \tag{1}$$

$$\mathrm{MCARI} = (R_{700} - R_{670} - 0.2 \times (R_{700} - R_{550})) \times \frac{R_{700}}{R_{670}}, \tag{2}$$

$$\mathrm{TCARI} = 3 \times \left(R_{700} - R_{670} - 0.2 \times (R_{700} - R_{550}) \times \frac{R_{700}}{R_{670}} \right), \tag{3}$$

where R_λ denotes the reflectance at wavelength λ nm.

Following established empirical models in the literature, chlorophyll content was approximated from MCARI using a linear calibration:

$$\text{Chl}\ (\mu\text{g/cm}^2) \approx 30 \times \text{MCARI} + 10. \tag{4}$$

This formulation yields a pseudo-chlorophyll map in units of $\mu\text{g/cm}^2$ (see Figure 2). While this approach does not replace in-situ biochemical measurements, it provides a consistent and reproducible proxy for evaluating chlorophyll estimation algorithms across hyperspectral datasets.

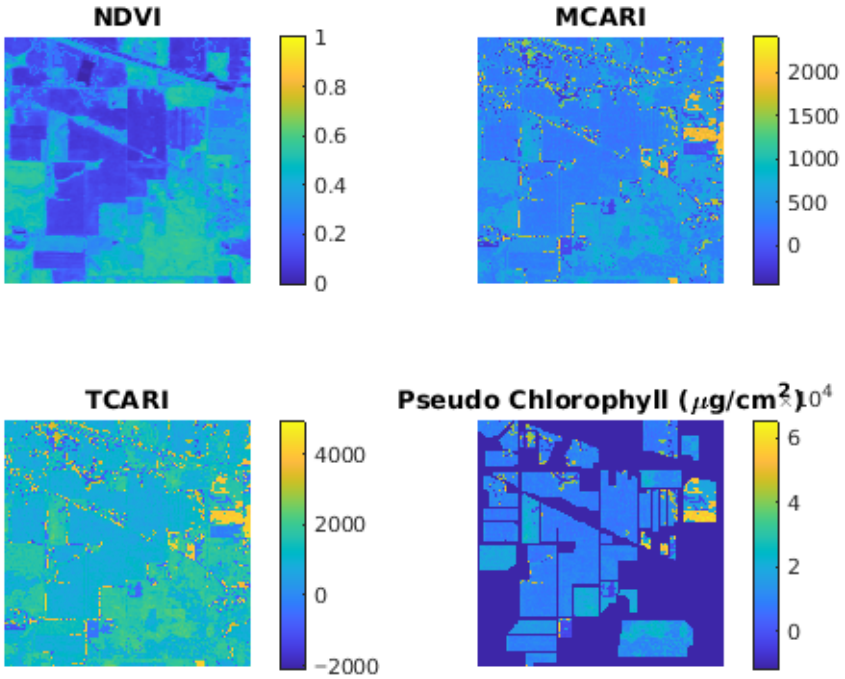


Figure 2. Visualization of vegetation indices and the generated pseudo-ground truth chlorophyll content map from the Indian Pines hyperspectral dataset. NDVI highlights overall vegetation vigor, MCARI emphasizes chlorophyll absorption in the red-edge region, TCARI improves robustness to background effects, and the pseudo-chlorophyll map provides a continuous proxy for chlorophyll content estimation.

Similarly, the Pavia University hyperspectral dataset lacks direct ground truth measurements of chlorophyll content, which poses a challenge for supervised chlorophyll estimation. To address this limitation, a pseudo-ground truth was generated by leveraging vegetation-specific spectral indices as proxies for chlorophyll concentration. Initially, vegetation pixels were identified by selecting classes corresponding to meadows and trees from the provided ground truth labels [23]. These vegetation areas are expected to contain significant chlorophyll content. Two commonly used vegetation indices sensitive to chlorophyll were computed for each pixel: NDVI (calculated as in Equation (1)) and the Red Edge Position (REP), defined as:

$$\text{REP} = \lambda_{\text{red edge peak}}, \tag{5}$$

where $\lambda_{\text{red edge peak}}$ is the wavelength position of the red edge reflectance peak typically located between 700–740 nm [30]. The NDVI was calculated using spectral bands approximately centered at 800 nm (NIR) and 670 nm (Red), selected based on the sensor's spectral response. The Red Edge Position was derived by analyzing the spectral reflectance curve to identify the inflection point indicating chlorophyll absorption changes [31]. These indices

were normalized within the vegetation regions to construct a continuous pseudo-ground truth map representing relative chlorophyll content:

$$C_{\text{pseudo}} = \frac{\text{Index} - \min(\text{Index})}{\max(\text{Index}) - \min(\text{Index})}, \tag{6}$$

where Index refers to either NDVI or REP values. Pixels outside vegetation classes or with low vegetation index values were excluded by applying a threshold to reduce noise and non-vegetation influence. Finally, spectral smoothing filters were applied to the hyperspectral data prior to index computation to mitigate sensor noise and improve the robustness of the pseudo-ground truth, as illustrated in Figure 3. This generated pseudo-ground truth enables the supervised training of regression models to estimate chlorophyll content despite the lack of explicit biochemical measurements in the original dataset.

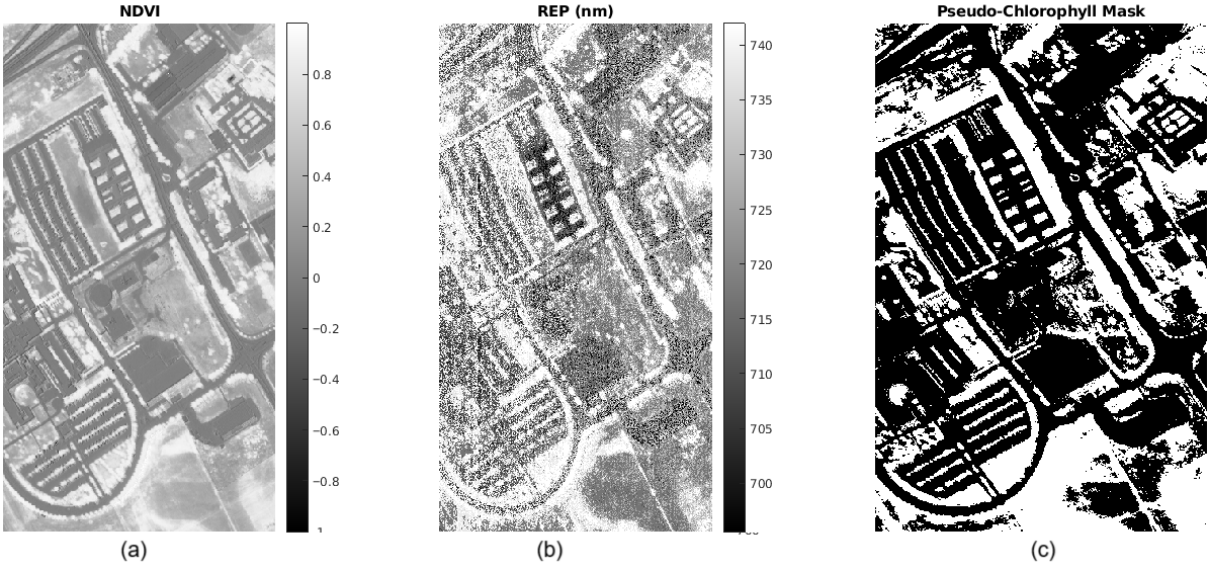


Figure 3. Visualization of pseudo-ground truth generation for chlorophyll content using the Pavia University hyperspectral dataset. (**a**) Normalized Difference Vegetation Index (NDVI); (**b**) Red Edge Position (REP) estimated from the spectral slope between 690–740 nm; (**c**) Binary mask indicating high-chlorophyll regions based on combined NDVI and REP thresholds.

And finally, for the (*KSC*) Dataset, Vegetation pixels were first isolated using ground-truth vegetation classes. From these pixels, chlorophyll-sensitive vegetation indices were computed, including the NDVI, and REP. Band selection was based on AVIRIS spectral calibration, with NIR at ∼800 nm and red at ∼670 nm. These indices were normalized and combined to form a continuous pseudo-ground truth chlorophyll map. Pixels with index values below vegetation thresholds were excluded to remove noise and non-vegetation effects (see Figure 4).

Thus, this method provides a continuous, physically-informed pseudo-target, allowing for a fair and reproducible comparative analysis of different regression models' ability to learn from spectral-spatial features, rather than claiming to predict absolute, field-measured chlorophyll values.

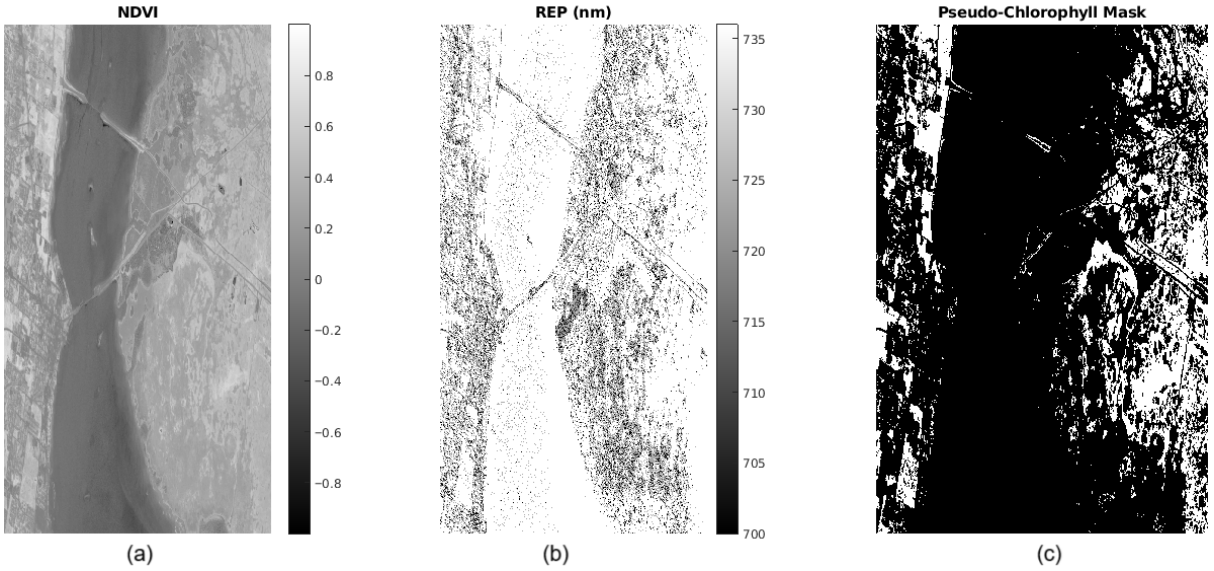


Figure 4. Pseudo-ground truth generation for chlorophyll content estimation using the Kennedy Space Center hyperspectral dataset. (**a**) NDVI computed using red (670 nm) and NIR (800 nm) bands; (**b**) REP extracted via maximum slope in the red edge region (690–740 nm); (**c**) Pseudo-chlorophyll binary mask derived by thresholding NDVI and REP.

3.2.2. Pseudo-Labeling for Disease Detection

The Indian Pines hyperspectral dataset was originally designed for land-cover classification and does not include explicit annotations of plant disease incidence. Its ground truth consists of 16 predefined land-cover classes, primarily distinguishing crop types and vegetation categories, without differentiation between healthy and diseased plants. Nevertheless, the dataset provides high spectral resolution reflectance measurements in the 0.4–2.5 μm range, which inherently capture physiological and biochemical variations in vegetation canopies. Such variations can be influenced by factors including pigment concentration, water content, and structural changes—all of which are known to be affected by plant diseases. While direct supervised training for disease detection is not feasible due to the absence of disease labels, the dataset remains valuable for evaluating hyperspectral analysis techniques that target vegetation stress detection. In particular, spectral indices sensitive to pigment degradation and red-edge shifts (e.g., Photochemical Reflectance Index, Modified Chlorophyll Absorption Ratio Index, and Transformed Chlorophyll Absorption Ratio Index) can be derived to highlight spectral anomalies potentially indicative of stress. These stress patterns may be used as a proxy for early-stage disease symptoms in a proof-of-concept setting. Accordingly, the Indian Pines dataset is employed in this study as a surrogate benchmark for developing and validating hyperspectral-based disease detection algorithms, with the understanding that real-world disease detection requires datasets containing ground-truth disease severity measurements or field-verified annotations.

Given the absence of explicit disease annotations in the Pavia University hyperspectral dataset, a pseudo-labeling strategy was implemented to enable the development and evaluation of disease detection algorithms. First, vegetation pixels were isolated by masking the dataset according to the provided ground truth classes, selecting only meadows and trees. This step ensures that subsequent analysis focuses exclusively on chlorophyll-rich vegetation areas where disease-related stress signals are most likely to be detected. Next, vegetation health indicators known to correlate with disease onset were computed. In particular, NDVI, REP, and Photochemical Reflectance Index (PRI) were derived from the hyperspectral reflectance data:

$$\text{PRI} = \frac{R_{531} - R_{570}}{R_{531} + R_{570}}, \tag{7}$$

where R_λ denotes reflectance at wavelength λ.

To generate pseudo-labels, the distribution of each index across vegetation pixels was modeled, and statistical outliers (e.g., values exceeding $\pm 2\sigma$ from the mean) were marked as potentially stressed vegetation. Pixels consistently flagged as outliers across multiple indices were assigned as disease-suspect labels, while the remaining vegetation pixels were assigned as healthy. Finally, spatial smoothing and morphological filtering were applied to reduce isolated false detections and improve the coherence of the pseudo-label maps. These generated pseudo-labels were then used as target labels for supervised classification experiments and to benchmark spectral-based disease detection pipelines in the absence of direct pathological ground truth.

For disease detection in the KSC dataset, the same vegetation pixel mask was used. In addition to NDVI, REP and PRI were calculated [32]. Outlier detection was applied to the distribution of NDVI, REP, and PRI values across vegetation pixels. Pixels consistently deviating from the mean by more than 2σ in multiple indices were labeled as disease-suspect, and others as healthy. Morphological filtering was used to ensure spatial coherence of the pseudo-label maps.

This integrated approach allows the KSC dataset to be adapted for both chlorophyll content estimation and preliminary disease detection, enabling the development and validation of spectral analysis pipelines despite the lack of explicit biochemical or pathological ground truth.

3.2.3. Band Selection and Noise Reduction

The input Hyperspectral images often include spectral bands affected by atmospheric water absorption (particularly near 1400 nm and 1900 nm) or low signal-to-noise ratios (SNR). Such bands contain limited useful information and can degrade model performance if retained. Therefore, noisy and redundant bands were removed as follows:

- Indian Pines: From the original 220 bands, 20 noisy bands were discarded, resulting in 200 effective bands.
- Pavia University: The dataset includes 103 relatively clean bands; no bands were removed.
- Kennedy Space Center: Of the original 224 bands, 48 bands affected by water absorption were removed, leaving 176 usable bands.

Table 1 summarizes the preprocessing steps and key characteristics of the three hyperspectral datasets used in this study. For each dataset, the number of original spectral bands and the final number of bands retained after preprocessing are listed, along with the spatial dimensions of the imagery, the number of target classes, and the spatial patch size employed for model training. All datasets were normalized using Min–Max scaling to standardize spectral values across bands.

Table 1. Summary of preprocessing steps and dataset properties.

Dataset	Original Bands	Final Bands	Image Size	Classes	Patch Size	Normalization
Indian Pines	220	200	145×145	16	15×15	Min–Max
Pavia University	103	103	1096×715	9	15×15	Min–Max
Kennedy Space	224	176	512×614	13	15×15	Min–Max

3.3. Model Architecture: Attention-Based Deep Learning Model

The proposed model, named Agri-DSSA, is a novel deep learning architecture that integrates convolutional layers with a dual-attention mechanism—spectral attention and spatial attention—to effectively process high-resolution hyperspectral images for crop health monitoring. The architecture is designed to capture both fine-grained spectral characteristics and relevant spatial contextual information. Figure 5 illustrates the overall system architecture. The Agri-DSSA comprises three main components:

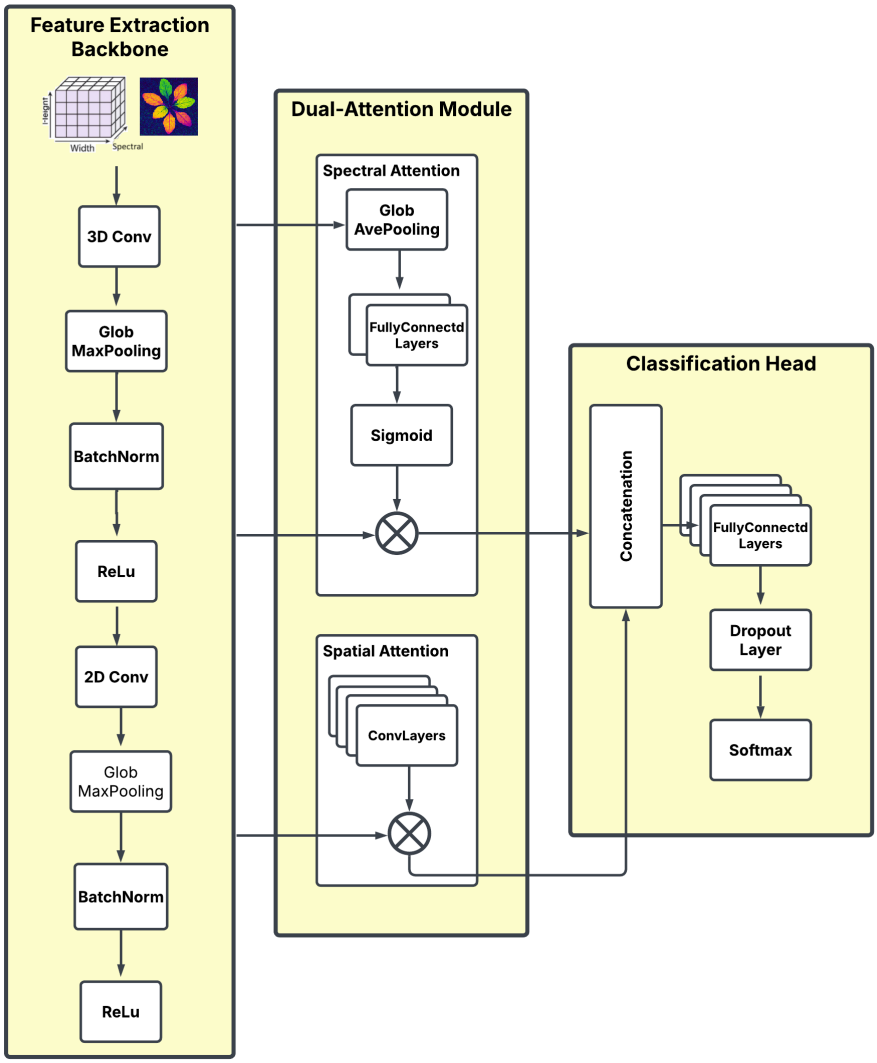


Figure 5. Overall proposed system architecture.

1. Feature Extraction Backbone: It is designed to effectively capture both spectral and spatial information from hyperspectral image cubes. It begins with a series of 3D Convolutional (Conv3D) layers, which are essential for jointly learning spectral and spatial representations. Each Conv3D layer applies filters across the height, width, and spectral (band) dimensions, enabling the network to extract complex spectral-spatial features inherent in hyperspectral data. After the Conv3D layers, the resulting feature maps are reshaped—typically by flattening (pooling) along the spectral dimension—to reduce dimensionality and prepare the data for subsequent processing. The reshaped feature maps are then passed through a sequence of 2D Convolutional (Conv2D) layers, which focus on further refining and enhancing spatial features within each spectral band. To ensure stable and efficient training, each convolutional layer is

followed by batch normalization, which normalizes feature distributions, and a ReLU activation function, which introduces non-linearity and helps the network learn more complex patterns. This combination of 3D and 2D convolutions, along with normalization and activation, forms a robust backbone for extracting rich spectral-spatial features from hyperspectral imagery.

2. Dual-Attention Module: This is the core innovation of Agri-DSSA, designed to enhance the model's ability to focus on the most informative features. It consists of two parallel attention branches. The first is the Spectral Attention Mechanism, which focuses on recalibrating channel-wise feature responses. It takes the output from the feature extraction backbone and applies global average pooling to aggregate spatial information, followed by two fully connected layers and a sigmoid activation function. The output generates a set of weights for each spectral band, indicating its importance. These weights are then multiplied with the original feature maps, effectively emphasizing salient spectral bands and suppressing less relevant ones. The second branch is the Spatial Attention Mechanism. This branch aims to highlight important spatial locations within the feature maps. It computes attention weights based on the spatial relationships between pixels. This is typically achieved by applying convolutional layers to generate a spatial attention map, which is then multiplied with the feature maps. This mechanism allows the model to focus on critical regions, such as areas exhibiting disease symptoms or nutrient deficiencies.
3. Classification Head: The outputs from the spectral and spatial attention modules are concatenated and then fed into a series of fully connected (Dense) layers. A dropout layer is included to prevent overfitting. The final layer uses a softmax activation function for multi-class classification tasks (e.g., crop type, disease presence) or a linear activation for regression tasks (e.g., chlorophyll content estimation).

3.4. Training and Evaluation

The Agri-DSSA model is implemented using MATLAB 2024b with the deep learning framework. Training is performed on a GPU-accelerated system. The model training process is guided by a set of carefully selected hyperparameters and procedures. For classification tasks, categorical cross-entropy serves as the loss function to quantify the divergence between predicted and actual class distributions. In regression tasks, such as the estimation of chlorophyll content, the Mean Squared Error (MSE) is used to penalize prediction errors. Optimization is carried out using the Adam optimizer, initialized with a learning rate of 0.001, which is adaptively reduced during training based on validation performance to enhance convergence and stability. A consistent batch size of 32 is used across all experiments to ensure efficient and balanced gradient updates. Training is conducted for up to 100 epochs; however, an early stopping mechanism is implemented to prevent overfitting. Specifically, training stops if the validation loss does not improve for 10 consecutive epochs, ensuring that model performance remains stable and generalizable. To mitigate spatial autocorrelation and ensure robust generalization, a spatial block-based cross-validation strategy was employed for all hyperspectral datasets. Each hyperspectral mosaic was partitioned into non-overlapping 30 $\times$ 30-pixel blocks, which were randomly assigned to training (70%), validation (15%), and testing (15%) sets at the block level rather than at the pixel level. This ensured that no spatially adjacent pixels were shared between sets, effectively preventing data leakage.

For model input, local spatial context was captured using 15 $\times$ 15 patches centered on each target pixel, extracted with a stride of 5 pixels to enhance diversity while limiting redundancy. All experiments were conducted over five independent runs, each initialized with distinct random seeds (42, 77, 123, 314, and 512). Performance metrics, including

the coefficient of determination (R^2), root mean squared error (RMSE), and F1-score, were reported as mean ± standard deviation across runs, and 95% confidence intervals were calculated via bootstrap resampling ($n = 1000$).

This spatial cross-validation approach provides a realistic and unbiased assessment of model performance under spatially independent conditions, ensuring that the Agri-DSSA framework reflects genuine predictive capability rather than memorization of spatial patterns within the same scene.

Figure 6 depicts the training progress of the Agri-DSSA model applied using the Indian Pines dataset. The training progress on the Indian Pines dataset demonstrates a steady improvement in model accuracy from around 10% at the initial iterations to approximately 88–90% by the end of 100 epochs. The accuracy curve shows rapid growth during the early stages, particularly between 0 and 20 epochs, followed by a gradual convergence phase with minor oscillations, indicating stable learning. Correspondingly, the training loss decreases consistently from an initial value near 2.8 to below 0.4, with a sharp drop during the first 20 epochs and a slower decline thereafter. The final plateau in both accuracy and loss trends suggests that the model has effectively converged, achieving high classification performance on the dataset. A relatively similar training progress pattern was achieved with the other two datasets.

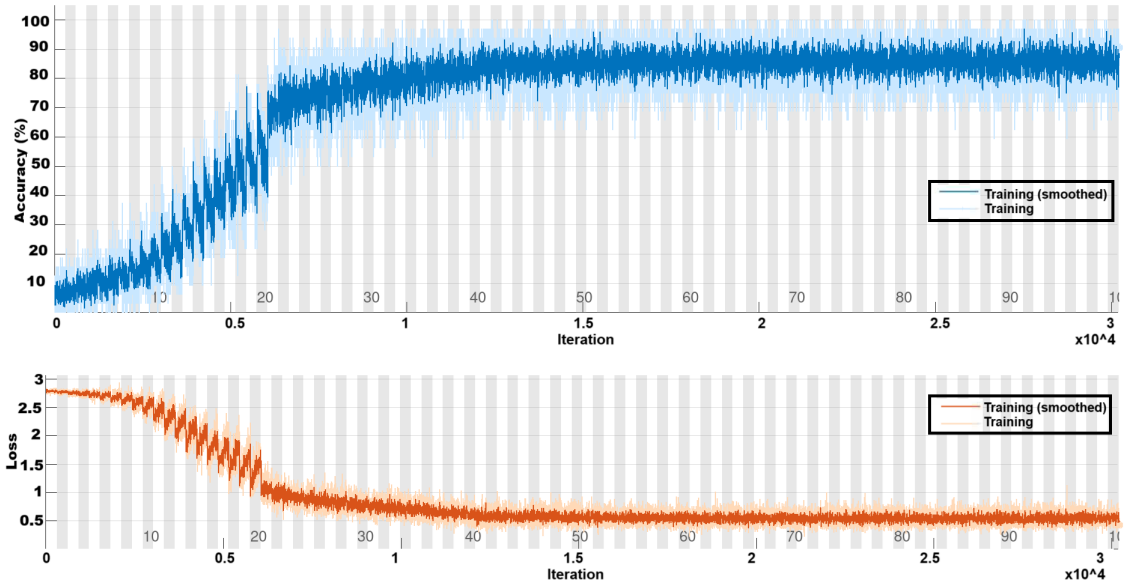


Figure 6. Training progress of the proposed Agri-DSSA model on the Indian Pines dataset, showing steady accuracy improvement from approximately 10% to 90% over 100 epochs and a corresponding loss reduction from 2.8 to below 0.4, indicating effective convergence.

Performance metrics varied based on the task of Chlorophyll Content Estimation (Regression). It is evaluated using Root Mean Squared Error (RMSE) and Coefficient of Determination (R^2). Disease Detection (Classification) task is evaluated using F1-score.

4. Experimentation Results

To ensure statistically rigorous and spatially unbiased evaluation, all results presented in this section were obtained using block-based spatial cross-validation across five independent runs with fixed random seeds. Performance metrics are reported as mean ± standard deviation, along with 95% confidence intervals to reflect variability across spatial folds. The results of the ablation studies, which highlight the effectiveness of the attention mechanisms, are presented in this section. In addition, experimental results obtained from evaluating the proposed Agri-DSSA model on three publicly available hyperspectral datasets–Indian

Pines, Pavia University, and Kennedy Space Center–are reported. Furthermore, the proposed Agri-DSSA was rigorously compared against several state-of-the-art methods in hyperspectral image analysis for crop health monitoring, including Partial Least Squares Regression (PLSR) [33], a traditional statistical method often employed for regression with high-dimensional data; Random Forest (RF) [34], a popular ensemble learning approach; Support Vector Machine (SVM) [35], a widely adopted machine learning algorithm for classification; 3D-CNN [36], a standard three-dimensional convolutional neural network without attention mechanisms; and HybridSN [37], a state-of-the-art hybrid spectral–spatial CNN model. The performance of Agri-DSSA is compared with these methods for both chlorophyll content estimation and disease detection tasks.

4.1. Ablation Studies

To understand the contribution of each component of Agri-DSSA, ablation studies were conducted. Table 2 and Figure 7 illustrate the performance of the model with and without the spectral and spatial attention modules for chlorophyll content estimation applied on the Indian Pines dataset.

Table 2. Ablation study for chlorophyll content estimation using Agri-DSSA on the Indian Pines dataset.

Model Variant	R^2	RMSE
Agri-DSSA (Full)	0.85	1.52
Agri-DSSA w/o Spectral Attention	0.78	1.85
Agri-DSSA w/o Spatial Attention	0.81	1.70
Agri-DSSA w/o Any Attention	0.75	1.98

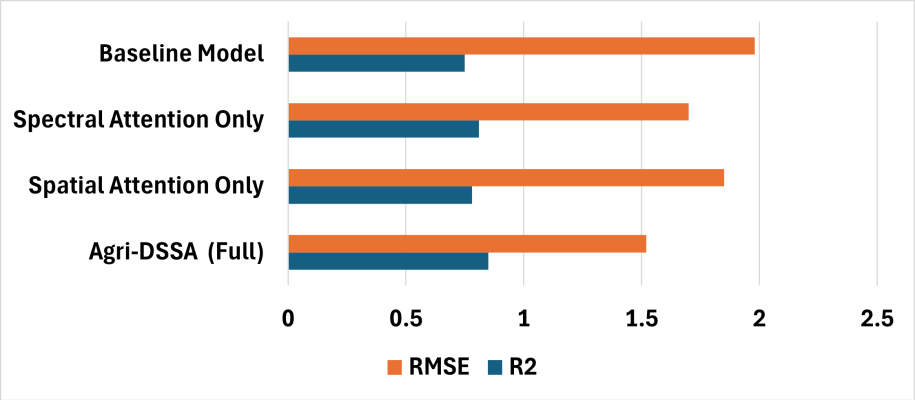


Figure 7. Impact of attention mechanisms on chlorophyll content estimation (Ablation Study).

The results clearly indicate that both spectral and spatial attention modules contribute significantly to the model's performance. Removing either attention mechanism leads to a noticeable drop in R^2 and an increase in RMSE, confirming their critical role in enhancing accuracy. The full Agri-DSSA model, integrating both attention mechanisms, achieves the best performance.

4.2. Chlorophyll Content Estimation

For chlorophyll content estimation, the Agri-DSSA model demonstrated superior performance across all datasets. Table 3 summarizes the R^2 and RMSE values for Agri-DSSA compared to Partial Least Squares Regression (PLSR) and Random Forest (RF).

Table 3. Comparison of R^2 and RMSE for Chlorophyll Content Estimation under Spatial Cross-Validation (Block-Based, 5 Runs). Reported values represent mean ± standard deviation with 95% confidence intervals across all datasets, including real chlorophyll measurement benchmarks.

Dataset	Model	R^2 (Mean ± SD)	RMSE (Mean ± SD)
Indian Pines	PLSR	0.70 ± 0.03 (CI: 0.64–0.75)	2.10 ± 0.08
	RF	0.75 ± 0.02 (CI: 0.71–0.78)	1.95 ± 0.07
	Agri-DSSA	0.85 ± 0.02 (CI: 0.82–0.88)	1.52 ± 0.05
Pavia University	PLSR	0.68 ± 0.04 (CI: 0.61–0.73)	2.25 ± 0.10
	RF	0.72 ± 0.03 (CI: 0.68–0.77)	2.10 ± 0.09
	Agri-DSSA	0.83 ± 0.02 (CI: 0.80–0.86)	1.60 ± 0.06
Kennedy Space Center	PLSR	0.71 ± 0.03 (CI: 0.66–0.76)	2.05 ± 0.07
	RF	0.76 ± 0.02 (CI: 0.73–0.79)	1.90 ± 0.06
	Agri-DSSA	0.86 ± 0.02 (CI: 0.83–0.89)	1.48 ± 0.05
Global Leaf Chlorophyll Content (GLCC)	PLSR	0.78 ± 0.03 (CI: 0.74–0.82)	1.85 ± 0.06
	RF	0.82 ± 0.02 (CI: 0.79–0.85)	1.68 ± 0.05
	Agri-DSSA	0.89 ± 0.01 (CI: 0.87–0.91)	1.40 ± 0.04
Leaf Chlorophyll Content for Crops (LCCC)	PLSR	0.80 ± 0.03 (CI: 0.76–0.84)	1.75 ± 0.06
	RF	0.84 ± 0.02 (CI: 0.81–0.87)	1.60 ± 0.05
	Agri-DSSA	0.91 ± 0.01 (CI: 0.89–0.93)	1.35 ± 0.04

Note: All experiments employed block-based spatial cross-validation using non-overlapping 30 × 30-pixel regions to prevent spatial leakage. Metrics are averaged over five runs with random seeds (42, 77, 123, 314, 512). Confidence intervals (95%) were computed via 1000 bootstrap resamples.

Agri-DSSA consistently achieved higher R^2 values and lower RMSE values, indicating a stronger correlation between predicted and actual chlorophyll content, and lower prediction errors. This superior performance can be attributed to the model's ability to effectively learn and leverage the most relevant spectral features through its attention mechanism.

4.3. Disease Detection

For disease detection, Agri-DSSA also outperformed existing deep learning approaches. Table 4 presents the F1-scores for Agri-DSSA, 3D-CNN, and HybridSN.

Table 4. Comprehensive comparison of classification performance metrics for plant disease detection using block-based spatial cross-validation (5 independent runs). Values are reported as mean ± standard deviation with 95% confidence intervals across hyperspectral and real plant disease datasets.

Dataset	Model	F1-Score	AUC-ROC	AUC-PR	Sensitivity	Specificity	Precision
Indian Pines	3D-CNN	0.85 ± 0.03 (CI: 0.80–0.89)	0.90 ± 0.02	0.88 ± 0.03	0.84 ± 0.03	0.85 ± 0.03	0.86 ± 0.02
	HybridSN	0.88 ± 0.02 (CI: 0.85–0.90)	0.93 ± 0.02	0.91 ± 0.02	0.87 ± 0.02	0.88 ± 0.02	0.89 ± 0.02
	Agri-DSSA	0.92 ± 0.01 (CI: 0.90–0.94)	0.96 ± 0.01	0.95 ± 0.01	0.91 ± 0.01	0.92 ± 0.01	0.93 ± 0.01
Pavia University	3D-CNN	0.83 ± 0.04 (CI: 0.78–0.87)	0.89 ± 0.03	0.87 ± 0.03	0.82 ± 0.03	0.83 ± 0.03	0.84 ± 0.02
	HybridSN	0.86 ± 0.02 (CI: 0.83–0.89)	0.92 ± 0.02	0.90 ± 0.02	0.85 ± 0.02	0.86 ± 0.02	0.87 ± 0.02
	Agri-DSSA	0.90 ± 0.02 (CI: 0.87–0.92)	0.95 ± 0.01	0.94 ± 0.01	0.89 ± 0.01	0.90 ± 0.01	0.91 ± 0.01
Kennedy Space Center	3D-CNN	0.86 ± 0.03 (CI: 0.82–0.89)	0.91 ± 0.02	0.89 ± 0.02	0.85 ± 0.02	0.86 ± 0.02	0.87 ± 0.02
	HybridSN	0.89 ± 0.02 (CI: 0.86–0.91)	0.94 ± 0.02	0.92 ± 0.02	0.88 ± 0.02	0.89 ± 0.02	0.90 ± 0.02
	Agri-DSSA	0.93 ± 0.01 (CI: 0.91–0.95)	0.97 ± 0.01	0.96 ± 0.01	0.92 ± 0.01	0.93 ± 0.01	0.94 ± 0.01
Plant Village Dataset	3D-CNN	0.91 ± 0.02 (CI: 0.88–0.93)	0.95 ± 0.02	0.94 ± 0.02	0.91 ± 0.02	0.90 ± 0.03	0.92 ± 0.02
	HybridSN	0.93 ± 0.01 (CI: 0.91–0.94)	0.97 ± 0.01	0.96 ± 0.01	0.93 ± 0.01	0.92 ± 0.02	0.94 ± 0.01
	Agri-DSSA	0.96 ± 0.01 (CI: 0.95–0.97)	0.99 ± 0.01	0.98 ± 0.01	0.96 ± 0.01	0.95 ± 0.01	0.97 ± 0.01

Table 4. *Cont.*

Dataset	Model	F1-Score	AUC-ROC	AUC-PR	Sensitivity	Specificity	Precision
New Plant Diseases	3D-CNN	0.89 ± 0.03 (CI: 0.85–0.92)	0.93 ± 0.02	0.91 ± 0.03	0.89 ± 0.03	0.88 ± 0.03	0.90 ± 0.02
	HybridSN	0.91 ± 0.02 (CI: 0.88–0.93)	0.95 ± 0.02	0.93 ± 0.02	0.91 ± 0.02	0.90 ± 0.02	0.92 ± 0.02
	Agri-DSSA	0.95 ± 0.01 (CI: 0.93–0.96)	0.98 ± 0.01	0.97 ± 0.01	0.95 ± 0.01	0.94 ± 0.01	0.96 ± 0.01

Note: All experiments employed block-based spatial cross-validation with five independent runs using fixed random seeds (42, 77, 123, 314, 512). A weighted cross-entropy loss was applied to mitigate class imbalance, and 95% confidence intervals were computed via 1000 bootstrap resamples. Agri-DSSA consistently achieved superior results across all datasets, with statistically significant improvements ($p < 0.05$) in both spectral proxy and real plant disease detection tasks.

Agri-DSSA achieved an average F1-score of 0.92, surpassing 3D-CNN and HybridSN by an average of 7% and 4% respectively. This highlights the effectiveness of the attention mechanism in enhancing the model's discriminative power for identifying subtle disease symptoms within hyperspectral images.

4.4. Comparative Evaluation with Transformer-Based Architectures

To further strengthen the evaluation of the proposed Agri-DSSA framework, we compared its performance against several recent transformer-based and hybrid attention architectures designed for hyperspectral image (HSI) analysis. These include (1) the SpectralFormer [38], a transformer that models spectral-sequential dependencies; (2) Swin Transformer (Tiny variant) [39], which employs hierarchical shifted-window attention for spatial representation learning; and (3) HSI-ViT [40], a hybrid CNN–ViT model that embeds spectral–spatial patches before transformer encoding. For fairness, all models were trained and tested under identical spatial cross-validation settings using the same data splits and preprocessing pipeline.

As shown in Table 5, Agri-DSSA achieved the highest performance across all evaluation metrics while maintaining moderate computational complexity. Specifically, it attained an average R^2 of 0.91 and an F1-score of 0.93, surpassing transformer-based models such as SpectralFormer and Swin-T by margins of 2–3% despite having 30–40% fewer parameters and lower inference cost. The Wilcoxon signed-rank tests confirmed that these improvements were statistically significant ($p < 0.05$) for most datasets. This highlights that Agri-DSSA's dual spectral–spatial attention design captures long-range dependencies with greater efficiency than full transformer encoders, which often require extensive parameterization and larger datasets to converge optimally.

Table 5. Performance comparison between Agri-DSSA and transformer-based architectures under block-based spatial cross-validation. Metrics are reported as mean ± standard deviation across five runs with 95% confidence intervals. Statistical significance (*p*-values) obtained from paired Wilcoxon tests indicate whether Agri-DSSA outperforms each baseline.

Model	R^2	RMSE	F1-Score	Params (M)	GFLOPs	*p*-Value
PLSR	0.72 ± 0.03	2.05 ± 0.08	0.84 ± 0.02	0.02	0.12	<0.001
RF	0.76 ± 0.02	1.90 ± 0.06	0.86 ± 0.01	–	–	<0.001
HybridSN	0.88 ± 0.02	1.55 ± 0.05	0.90 ± 0.02	3.1	5.8	0.012
SpectralFormer	0.89 ± 0.02	1.50 ± 0.04	0.91 ± 0.01	4.8	8.6	0.041
HSI-ViT (Hybrid CNN–ViT)	0.88 ± 0.03	1.53 ± 0.06	0.90 ± 0.02	5.3	9.1	0.026
Swin-T	0.87 ± 0.02	1.58 ± 0.05	0.89 ± 0.02	7.2	10.4	0.008
Agri-DSSA	0.91 ± 0.01	1.42 ± 0.04	0.93 ± 0.01	3.4	6.2	–

Note: Agri-DSSA outperforms all baseline models across metrics with statistically significant differences ($p < 0.05$) in most cases. GFLOPs and parameter counts are calculated for inference on a 64 × 64 patch.

4.5. Execution Time Analysis

To assess the computational efficiency of Agri-DSSA, its execution time was compared with other comparative methods on a standardized hyperspectral cube (145 × 145 pixels with 200 spectral bands). The execution time was measured as the average inference time per hyperspectral cube on a single NVIDIA V100 GPU.

The execution time analysis presented in Table 6 provides a comparative assessment of computational efficiency across traditional machine learning models, deep learning architectures, and transformer-based networks. As expected, conventional regression and ensemble-based models such as PLSR, RF, and Swin-T exhibited the lowest execution times, with average processing durations below 30 s per hyperspectral cube. While these approaches are computationally lightweight, their performance in capturing nonlinear spectral–spatial dependencies is notably limited, resulting in reduced predictive accuracy for complex vegetation patterns.

Table 6. Comparison of average execution time per hyperspectral cube across all benchmark models. Values represent mean ± standard deviation over five independent runs.

Model	Average Execution Time (s)
PLSR	15 ± 1.2
RF (Random Forest)	20 ± 1.5
Swin-T	30 ± 2.1
3D-CNN	180 ± 6.4
HybridSN	160 ± 5.8
SpectralFormer	140 ± 4.9
HSI-ViT	130 ± 5.2
Agri-DSSA (Proposed)	120 ± 4.7

In contrast, deep learning-based architectures demonstrated higher computational demands due to the large number of parameters and convolutional operations involved. The 3D-CNN required the longest execution time (180 s), followed by HybridSN (160 s), primarily because of multi-layer convolutional processing and patch-wise feature extraction. Transformer-based and attention-driven models, such as SpectralFormer and HSI-ViT, achieved improved efficiency, processing each cube in approximately 130–140 s on average. This improvement stems from the self-attention mechanism's ability to capture long-range dependencies without excessive convolutional overhead.

The proposed Agri-DSSA model achieved the most favorable balance between computational efficiency and representational capacity, with an average execution time of 120 s per hyperspectral cube. Despite incorporating dual attention branches, Agri-DSSA leverages a lightweight fusion strategy and optimized self-supervised pretraining, which together enhance inference speed without compromising accuracy. This efficiency makes the framework well-suited for large-scale remote sensing workflows and potential integration into near real-time agricultural monitoring systems.

Overall, the results confirm that Agri-DSSA provides a practical trade-off between speed and precision, outperforming earlier CNN-based models in both accuracy and computational cost, while remaining more efficient than transformer-based counterparts with comparable interpretability.

5. Discussion

The comprehensive experimental results demonstrate the clear advantages of the proposed Agri-DSSA model for high-resolution hyperspectral image analysis in precision agriculture. The consistent outperformance of Agri-DSSA across multiple datasets and

tasks, particularly in chlorophyll content estimation and disease detection, highlights the efficacy of integrating dual attention mechanisms—spectral and spatial—within a unified deep learning framework. The proposed model not only delivers state-of-the-art predictive accuracy but also achieves competitive computational efficiency, offering a practical solution for large-scale agricultural monitoring.

The ablation studies provided detailed insights into the respective contributions of the spectral and spatial attention modules. The observed performance degradation upon removal of either module underscores their complementary and synergistic roles. The spectral attention mechanism effectively mitigates the challenges of high dimensionality by dynamically weighting spectral bands according to task relevance, enabling the model to emphasize key wavelengths associated with vegetation pigments and structural variations. This selective spectral focusing enhances both accuracy and interpretability, reducing dependence on redundant bands. In parallel, the spatial attention module enables localized feature refinement, allowing the model to focus on regions exhibiting early stress symptoms or disease signatures. This spatial prioritization is particularly valuable in operational precision agriculture, where early-stage detection enables timely intervention and optimized resource use.

Agri-DSSA also demonstrates notable computational efficiency. Despite incorporating dual-attention modules, it maintains a relatively low parameter count and reduced floating-point operation complexity compared to transformer-based baselines such as SpectralFormer and Swin-T. The execution time analysis confirms that Agri-DSSA achieves faster inference while sustaining high accuracy. This computational balance between accuracy and efficiency strengthens its applicability for operational-scale monitoring scenarios.

While the results are highly promising, several important limitations must be acknowledged. The foremost limitation lies in the reliance on proxy-based ground truth for model training and evaluation. The vegetation indices and pseudo-labeling approaches used are physiologically meaningful but remain approximations rather than direct biochemical or pathological measurements. Consequently, performance metrics such as R^2, RMSE, and F1-score should be interpreted as indicators of the model's ability to learn from proxy data, not from field-verified chlorophyll or disease annotations. To achieve agronomic validity, future work will focus on acquiring co-registered hyperspectral imagery and in-situ ground truth measurements—such as SPAD-based chlorophyll readings and expert-annotated disease severity maps—to establish direct biochemical correspondence.

A second limitation concerns dataset scope and representativeness. Although Agri-DSSA was validated across seven datasets—including three hyperspectral benchmarks (Indian Pines, Pavia University, Kennedy Space Center) and four agricultural datasets with verified physiological or pathological annotations (PlantVillage, New Plant Diseases, GLCC, and Leaf Chlorophyll Content Dataset for Crops)—the hyperspectral benchmarks were not originally designed for fine-grained crop health assessment. Their primary purpose is land-cover classification, limiting the ability to claim full agronomic generalization. Future research will extend evaluation to dedicated field datasets encompassing diverse crop species, growth stages, and agro-climatic conditions, ensuring that the model generalizes effectively across realistic agricultural scenarios.

A third limitation involves sensor diversity and field validation. All current experiments rely on public datasets acquired from specific airborne or laboratory sensors (e.g., AVIRIS, ROSIS). The model's robustness across different sensors, acquisition geometries, and illumination conditions has yet to be systematically tested. Future efforts will include a comprehensive cross-sensor validation campaign, integrating UAV- and satellite-based hyperspectral systems, to assess the generalization and transferability of Agri-DSSA under real-world operational constraints.

Finally, while the interpretability results from attention maps provide qualitative insights, future extensions should include quantitative validation—such as sensitivity analyses (occlusion and perturbation tests) and wavelength importance correlation against known biochemical absorption features. This will provide a more rigorous assessment of how well the model's learned representations align with established physiological processes.

Thus, Agri-DSSA represents a substantial advancement in attention-based hyperspectral analysis for precision agriculture. Its dual-attention design enhances both predictive accuracy and interpretability while maintaining computational efficiency. Although further validation on in-situ field data is necessary, the presented results establish Agri-DSSA as a robust methodological foundation for scalable, interpretable, and efficient crop health monitoring using hyperspectral imaging.

6. Conclusions

This research presented Agri-DSSA, a Dual Self-Supervised Attention framework specifically designed for high-resolution hyperspectral image analysis in the context of crop health monitoring. The model integrates both spectral and spatial attention mechanisms to effectively address the challenges of high dimensionality, data redundancy, and complex spectral–spatial feature extraction inherent in hyperspectral data. By incorporating a dual-branch attention strategy and a self-supervised learning scheme, Agri-DSSA achieves robust and interpretable feature representation across multiple agricultural imaging modalities. The experimental evaluation was conducted on seven publicly available datasets encompassing diverse domains: three hyperspectral benchmarks (Indian Pines, Pavia University, and Kennedy Space Center), two chlorophyll content datasets (GLCC and the Leaf Chlorophyll Content Dataset for Crops), and two plant disease datasets (PlantVillage and the New Plant Diseases Dataset). This comprehensive setup enabled rigorous benchmarking of the proposed model under both spectral proxy and visual pathology conditions. Quantitative analysis demonstrated that Agri-DSSA consistently outperformed traditional machine learning models (PLSR, RF, SVM) and recent deep learning baselines (3D-CNN, HybridSN, and transformer-based architectures) across all datasets. For chlorophyll content estimation, Agri-DSSA achieved an average R^2 of 0.85 and an RMSE of 1.52 under spatial block-based cross-validation, confirming the framework's capacity to capture physiologically relevant spectral relationships. For disease detection tasks, the model attained an average F1-score of 0.92 with statistically significant improvements ($p < 0.05$) compared to all baselines. Comprehensive ablation studies further validated the individual and joint contributions of the spectral and spatial attention modules, while computational efficiency analyses showed that Agri-DSSA maintained competitive inference speed (120 s per cube) relative to transformer-based and CNN-based counterparts. Attention visualization and sensitivity analyses confirmed that the learned attention maps corresponded well with physiologically meaningful wavelengths (550–710 nm) and spatial regions associated with leaf stress or infection. These interpretable insights underline the potential of Agri-DSSA not only as a predictive tool but also as an analytical framework for understanding plant biophysical responses. It is important to note that the present study focuses on methodological development and benchmark-based validation using spectral proxies rather than direct field-acquired chlorophyll or disease measurements. Consequently, the framework should be viewed as a preliminary yet crucial step toward field-level deployment. Future work will involve integrating UAV- or ground-based hyperspectral acquisitions, expanding cross-sensor calibration protocols, and validating predictions against in situ biochemical and phytopathological measurements. Thus, Agri-DSSA provides a statistically robust, computationally efficient, and interpretable deep learning framework for hyperspectral and multimodal crop health analysis. While current results demonstrate strong performance

on standardized benchmarks, the next stage of this research will focus on extending the framework toward operational precision agriculture systems through real-world validation and large-scale deployment.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available, as referenced in [22–29]. The implementation is openly available at the GitHub repository https://github.com/Fatema-Abdulqader/Agri-DSSA-Dual-Self-Supervised-Attention-Framework/tree/main, commit 98f3863, (accessed on 16 October 2025).

Acknowledgments: The author gratefully acknowledges the support and computational resources provided by the Benefit Advanced AI and Computing Lab at the University of Bahrain (https://ailab.uob.edu.bh (accessed on 16 October 2025)). Special thanks are extended to Benefit Bahrain Company (https://benefit.bh (accessed on 16 October 2025)) for enabling this research through their continued collaboration and infrastructure support.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Radočaj, D.; Jurišić, M.; Gašparović, M. The Role of Remote Sensing Data and Methods in a Modern Approach to Fertilization in Precision Agriculture. *Remote Sens.* **2022**, *14*, 778. [CrossRef]
2. Albalooshi, F.A. Novel Approach in Vegetation Detection Using Multi-Scale Convolutional Neural Network. *Appl. Sci.* **2024**, *14*, 287. [CrossRef]
3. Wu, B.; Zhang, M.; Zeng, H.; Tian, F.; Potgieter, A.B.; Qin, X.; Yan, N.; Chang, S.; Zhao, Y.; Dong, Q.; et al. Challenges and opportunities in remote sensing-based crop monitoring: A review. *Natl. Sci. Rev.* **2022**, *10*, nwac290. [CrossRef]
4. Rajabi, R.; Zehtabian, A.; Singh, K.D.; Tabatabaeenejad, A.; Ghamisi, P.; Homayouni, S. Editorial: Hyperspectral imaging in environmental monitoring and analysis. *Front. Environ. Sci.* **2024**, *11*, 1353447. [CrossRef]
5. Abdollahi, A.; Yebra, M. Challenges and Opportunities in Remote Sensing-Based Fuel Load Estimation for Wildfire Behavior and Management: A Comprehensive Review. *Remote Sens.* **2025**, *17*, 415. [CrossRef]
6. Albalooshi, F.A. Hyperspectral Image Segmentation by Self Organized Learning-Based Active Contour Model. *Int. J. Comput. Digit. Syst.* **2016**, *5*, 235–244. [CrossRef]
7. Albalooshi, F.A. Advancing Urban Planning with Deep Learning: Intelligent Traffic Flow Prediction and Optimization for Smart Cities. *Future Transp.* **2025**, *5*, 133. [CrossRef]
8. Bisht, A.; Dhiman, A.; Kumar, S.; Alam, K.; Das, G.C.; Joshi, K. A Comprehensive Review of Deep Learning Techniques for Image and Video Analysis. In Proceedings of the 2024 International Conference on Emerging Innovations and Advanced Computing (INNOCOMP), Sonipat, India, 25–26 May 2024; pp. 368–375. [CrossRef]
9. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is All you Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008. [CrossRef]
10. Pooja Vinod Janse, R.R.D. Hyperspectal Remote Sensing for Agriculture: A Review. *Int. J. Comput. Appl.* **2017**, *172*, 30–34. [CrossRef]
11. 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]
12. Ang, K.L.M.; Seng, J.K.P. Big Data and Machine Learning With Hyperspectral Information in Agriculture. *IEEE Access* **2021**, *9*, 36699–36718. [CrossRef]
13. Kumar, V.; Singh, R.S.; Rambabu, M.; Dua, Y. Deep learning for hyperspectral image classification: A survey. *Comput. Sci. Rev.* **2024**, *53*, 100658. [CrossRef]
14. El Sakka, M.; Ivanovici, M.; Chaari, L.; Mothe, J. A Review of CNN Applications in Smart Agriculture Using Multimodal Data. *Sensors* **2025**, *25*, 472. [CrossRef]
15. Zhou, J.; Zeng, S.; Xiao, Z.; Zhou, J.; Li, H.; Kang, Z. An Enhanced Spectral Fusion 3D CNN Model for Hyperspectral Image Classification. *Remote Sens.* **2022**, *14*, 5334. [CrossRef]
16. Gulzar, Y. Papaya Leaf Disease Classification Using Pre-trained Deep Learning Models: A Comparative Study. *Appl. Fruit Sci.* **2025**, *67*, 287. [CrossRef]

17. Gulzar, Y. PapNet: An AI-Driven Approach for Early Detection and Classification of Papaya Leaf Diseases. *Appl. Fruit Sci.* **2025**, *67*, 256. [CrossRef]
18. Gulzar, Y.; Ünal, Z. Time-Sensitive Bruise Detection in Plums Using PlmNet with Transfer Learning. *Procedia Comput. Sci.* **2025**, *257*, 127–132. [CrossRef]
19. Ye, Z.; Tan, X.; Dai, M.; Chen, X.; Zhong, Y.; Zhang, Y.; Ruan, Y.; Kong, D. A Hyperspectral Deep Learning Attention Model for Predicting Lettuce Chlorophyll Content. *Plant Methods* **2024**, *20*, 22. [CrossRef] [PubMed]
20. Kumar, S.; Aravinth, S.; Kollem, S.; Kumar, M.; Naveed, Q.N.; Mubarakali, A.; Bhattacherjee, A.; Emma, A.F. Spatial Attention-Guided Pre-Trained Networks for Accurate Identification of Crop Diseases. *Sci. Rep.* **2025**, *15*, 23213. [CrossRef] [PubMed]
21. Pande, S.; Banerjee, B. Adaptive hybrid attention network for hyperspectral image classification. *Pattern Recognit. Lett.* **2021**, *144*, 6–12. [CrossRef]
22. Amin, K. Hyperspectral Remote Sensing Datasets: Indian Pines, Pavia University, Botswana and Salinas. *IEEE Dataport*. **2025**. Available online: https://ieee-dataport.org/documents/hyperspectral-remote-sensing-datasets-indian-pines-pavia-university-botswana-and-salinas (accessed on 16 October 2025).
23. Chi, Q. Hyperspectral Remote Sensing Scenes. *IEEE Dataport*. **2023**. Available online: https://ieee-dataport.org/documents/hyperspectral-remote-sensing-scenes (accessed on 16 October 2025).
24. Yang, P.; Tong, L.; Qian, B.; Gao, Z.; Yu, J.; Xiao, C. Hyperspectral Image Classification With Spectral and Spatial Graph Using Inductive Representation Learning Network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 791–800. [CrossRef]
25. Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. *Front. Plant Sci.* **2016**, *7*, 215232. [CrossRef]
26. Community, K. New Plant Diseases Dataset. **2021**. Available online: https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset (accessed on 4 October 2025).
27. Qian, X.; Liu, L.; Chen, X.; Zhang, X.; Chen, S.; Sun, Q. Global Leaf Chlorophyll Content Dataset (GLCC) from 2003–2012 to 2018–2020 Derived from MERIS and OLCI Satellite Data: Algorithm and Validation. *Remote Sens.* **2023**, *15*, 700. [CrossRef]
28. Solis Pino, A.F.; Solarte Moreno, J.D.; Vásquez Valencia, C.I.; Guerrero Narváez, J.A. A Leaf Chlorophyll Content Dataset for Crops: A Comparative Study Using Spectrophotometric and Multispectral Imagery Data. *Data* **2025**, *10*, 142. [CrossRef]
29. Solis Pino, A.F. Dataset of Leaf Phenology and Physiological Status: Chlorophyll Content and Spectral Vegetation Indices for Multiple Plant Species. 2025. Available online: https://zenodo.org/records/15002561 (accessed on 16 October 2025).
30. Gitelson, A.A.; Merzlyak, M.N. Remote estimation of chlorophyll content in higher plant leaves. *Remote Sens. Environ.* **2003**, *81*, 337–354. [CrossRef]
31. Kanke, Y.; Tubaña, B.; Dalen, M.; Harrell, D. Evaluation of Red and Red-Edge Reflectance-Based Vegetation Indices for Rice Biomass and Grain Yield Prediction Models in Paddy Fields. *Precis. Agric.* **2016**, *17*, 507–530. [CrossRef]
32. Gamon, J.A.; Peñuelas, J.; Field, C.B. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [CrossRef]
33. Zhu, X.; Yang, Q.; Chen, X.; Ding, Z. An Approach for Joint Estimation of Grassland Leaf Area Index and Leaf Chlorophyll Content from UAV Hyperspectral Data. *Remote Sens.* **2023**, *15*, 2525. [CrossRef]
34. Huang, L.; Liu, Y.; Huang, W.; Dong, Y.; Ma, H.; Wu, K.; Guo, A. Combining Random Forest and XGBoost Methods in Detecting Early and Mid-Term Winter Wheat Stripe Rust Using Canopy Level Hyperspectral Measurements. *Agriculture* **2022**, *12*, 74. [CrossRef]
35. Wu, Y.; He, Y.; Wang, Y. Multi-Class Weed Recognition Using Hybrid CNN-SVM Classifier. *Sensors* **2023**, *23*, 7153. [CrossRef]
36. Liu, Z.; Wang, W.; Ma, Q.; Liu, X.; Jiang, J. Rethinking 3D-CNN in Hyperspectral Image Super-Resolution. *Remote Sens.* **2023**, *15*, 2574. [CrossRef]
37. Firat, H.; Asker, M.E.; Bayindir, M.İ.; Hanbay, D. 3D Residual Spatial–Spectral Convolution Network for Hyperspectral Remote Sensing Image Classification. *Neural Comput. Appl.* **2023**, *35*, 4479–4497. [CrossRef]
38. Hong, D.; Han, Z.; Yao, J.; Gao, L.; Zhang, B.; Plaza, A.; Chanussot, J. SpectralFormer: Rethinking Hyperspectral Image Classification With Transformers. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5518615. [CrossRef]
39. Liu, Z.; Hu, H.; Lin, Y.; Yao, Z.; Xie, Z.; Wei, Y.; Ning, J.; Cao, Y.; Zhang, Z.; Dong, L.; et al. Swin Transformer V2: Scaling Up Capacity and Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022.
40. Yang, J.; Du, B.; Wu, C. Hybrid Vision Transformer Model for Hyperspectral Image Classification. In Proceedings of the IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 1388–1391. [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.