FMDB *Transactions on Sustainable Computing Systems*

Integrating Deep Learning and Spectral Analysis for Multi-Modal Data Fusion in Precision Agriculture for Enhancing Crop Health Monitoring and Yield Prediction

Sheradha Jauhari[1], Krishna Kant Agrawal[2,*], Satya Prakash Yadav[3], Angeles Quezada[4]

[1]School of Computing Science and Engineering, Galgotias University, Gautam Buddha Nagar, Uttar Pradesh, India.
[2]School of Computer Science and Engineering, Galgotias University, Gautam Buddha Nagar, Uttar Pradesh, India.
[3]Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, Uttar Pradesh, India.
[4]Department of Systems and Computing, Institute of Tijuana, Tijuana, Baja California, Mexico.
sheradha20@gmail.com[1], kkagrawal@outlook.com[2], prakashyadav.satya@gmail.com[3], angeles.quezada@tectijuana.edu.mx[4]

Abstract: The purpose of Precision Agriculture is to incorporate technology into various agricultural processes to increase efficiency and productivity. In fact, Precision Agriculture uses advanced technologies such as sensors and data analytics to improve crop yields. However, a significant challenge in this area is effectively integrating multiple data sources to accurately predict crop health and yield using all available information. This problem arises because traditional models typically use spectral analysis or deep learning techniques independently. Due to this separation, neither method generates the desired results. Researchers propose a solution to this issue by combining spectral analysis and deep learning for multimodal data fusion in precision agriculture. Our integrated approach begins with the collection of multispectral data from drone- or satellite-based sensors to characterise crop types. Spectral analysis will determine each crop type's chlorophyll and water content, which affect plant health. Deep learning will be used to analyse the intricate interconnections between crop yields and derived attributes to understand their relationships better. Integrated use of these two technologies will give us a broader range of data and knowledge about crop variety health and yield than single-use or standalone applications.

Keywords: Farming Practices; Yield Accurately; Precision Agriculture; Spectral Analysis; Deep Learning (DL); Chlorophyll Content; Multimodal Data Fusion; Intricate Interconnection.

Received on: 18/03/2025, **Revised on:** 13/05/2025, **Accepted on:** 21/08/2025, **Published on:** 11/01/2026

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSCS

DOI: https://doi.org/10.69888/FTSCS.2026.000610

Cite as: S. Jauhari, K. K. Agrawal, S. P. Yadav, and A. Quezada, "Integrating Deep Learning and Spectral Analysis for Multi-Modal Data Fusion in Precision Agriculture for Enhancing Crop Health Monitoring and Yield Prediction," *FMDB Transactions on Sustainable Computing Systems*, vol. 4, no. 1, pp. 63–77, 2026.

1. Introduction

The method of combining data from multiple sources using advanced algorithms to enhance the quality and presentation of the data is called Multimodal Data Fusion [1]. Multimodal data fusion enables accurate yield forecasting in precision agriculture

*Corresponding author.

and a more detailed assessment of crop condition. To collect data on the different parameters of the crop(s) (e.g., soil moisture, nutrients, plant health), the multimodal agricultural domain employs various technologies, such as sensors, drones, and satellite imagery [2]. However, each of the data sources mentioned has limitations, offering only a partial perspective on the crop and potentially providing inconsistent or inadequate information on the crop(s) [3]. By fusing data from multiple sources, multimodal data fusion removes these limitations and creates a comprehensive representation of the crop [4]. This is a significant advantage of using multimodal data fusion in precision agriculture, as it will improve crop health monitoring. Using multiple data sources enables the detection of patterns and relationships that would not be identifiable otherwise [5]. Accurate assessments of how water stress affects crop growth can be achieved by using drone images of a field, along with soil moisture data, to analyze this relationship [6].

Additionally, combining data from plant nutrient-sensing sensors with satellite imagery will enable rapid identification and response to nutrient deficiencies in crops [7]. When multiple sources provide different types of information about a crop's health, this allows for an accurate understanding of the crop's overall health, rather than relying on a single source [8]. Multi-modal data fusion has the potential to significantly improve the accuracy of yield prediction and crop health assessment in precision agriculture [9]. However, there are many technological barriers to realizing the full benefits of the multi-modal data fusion technique. One major obstacle is the diversity of data types and sources used in precision agriculture today. There is a variety of available data, including catapulted satellites and satellite-based sensors, as well as small IoT-type sensors located on fields and remote sensing from satellites and airborne devices [10]. Each sensor provides different levels of accuracy, resolution, and data formats depending on its type and technology. Due to the broad range of agricultural data types, many complex algorithms are required to integrate and fuse them, and to manage large volumes of agricultural data, given the overwhelming scale of precision agriculture's datasets. Furthermore, combining data from multiple modalities will create even larger volumes to manage [14]. Therefore, effective data processing and infrastructure are required to manage these massive datasets [15]. The key contribution of the research is as follows:

- **Enhanced Accuracy of Crop Health Monitoring:** Multimodal data fusion can combine data from multiple sources, including soil sensors, meteorological data, and satellite imagery, to provide a more thorough and precise evaluation of crop health. Through improved capacity to identify early indicators of illness or stress, you can take action more quickly to improve your crop yield. There is a new way to integrate deep learning and spectral analysis into multimodal data fusion that researchers believe will help you achieve this goal by providing an enhanced, much more reliable way to monitor crop health.
- **More Accurate Yield Prediction:** To forecast how many pounds of produce a crop will yield at harvest time, researchers have used the fusion of multiple data sources to build sophisticated machine learning and predictive models. These data-fusion techniques can enable farmers to make informed decisions on crop management practices, such as when to water, how to fertilize, and when to apply pesticides, to maximize crop yield. Thus, researchers believe that the developed method for deep-learning and spectral-analysis-based multi-modal data fusion provides enhanced predictive capabilities, significantly improving predicted yield and having a dramatic positive impact on the daily operations of adopting this technique.
- **Identification of Key Contributors to Crop Health:** By using multi-model data fusion, all data sources can be incorporated and a complete picture created; helping to identify the relationships among different variables, which lead to a deeper knowledge of the contributors to the health of the crop, areas to focus on to enhance crop yield, thus giving confidence in the thoroughness of this approach.

2. Related Words

Lin et al. [11] have written about tracking smallholder agriculture through Multimodal Earth Observation, which includes Optical, Radar, and Thermal imagery from various satellite sources, providing a more comprehensive and scalable means of monitoring smallholder farms, their crop health and yields, and land management practices. Strani et al. [12] have described a new method for predicting crop yield using multi-temporal data that combines historical and current data from a given field to forecast future yields. This is accomplished through Remote Sensing and Data Analysis, thereby providing farmers with important information about their crops. Qiao et al. [13] describe research on monitoring drought conditions in winter wheat during critical growth periods and how Artificial Intelligence applied to multiple types of data (Remote Sensing Imagery, Weather Data) can help identify and monitor drought conditions in winter wheat crops. These studies are designed to provide farmers with up-to-date information to help them make informed decisions about how drought affects crop production.

Zhang et al. [19], for instance, outline studies aimed at developing a multi-stage plant phenotyping method for soybeans using a variety of additional sensor technologies installed on UAVs. Agronomists will obtain precise, comprehensive evaluations of soybean growth and development by leveraging a variety of sensor data sources, enabling effective, focused plant resource management. Lu et al. [20] study is another illustration. Describe the creation of a model of the possible damage done by LDB based on IoT and hyperspectral forecasting data. This model will better leverage advanced data fusion techniques, enabling users to more accurately predict when and how severely an LDB will affect their crops.

Barbato et al. [21] describe a combined remote sensing data set with multiple modalities that can improve classification accuracy across many different types of imaging sensors, including optical (visible range), thermal, and LiDAR, by enabling feature extraction for analysis. The objective of the research is to train and evaluate algorithms for accurately locating and delineating humans, human-made structures (such as buildings and roads), and vegetation on Earth's surface. Fan and Li [22] discussed sensor fusion, which combines multiple sensor readings to produce a more accurate and comprehensive representation of an environment or system. The use of sensor fusion technologies for precision agriculture will enable users to improve crop management by integrating multiple data sources from sensors, including weather, soil, and drone imagery. Elsherbiny et al. [23] discussed a hybrid deep learning network that leverages both DL and IoT technologies to accurately assess the water status of wheat crops. This hybrid network uses multiple data sources, including camera images, soil sensors, and meteorological data, to accurately estimate a crop's water requirements. As a result, the crop will be irrigated more efficiently, leading to higher yields. Clamens et al. [24] discussed the application of YOLO. This real-time visual object detection model uses both RGB-D (image and depth) data to detect human activity on or near vineyards, as part of a computer vision application to analyze vineyards. This technology can immediately identify and classify items, providing precise, fast, and consistent monitoring of vineyard performance and growth. In addition, Lin et al. [25] describe a dataset for crop yield prediction based on multimodal climate change data. This dataset includes climate-related data (weather, soil, and historical crop yields) and is designed to help predict how climate change will impact crop production.

Researchers and policymakers in environmental science or agriculture who seek to understand and drive change in response to climate change may find the dataset useful. Garnot et al. [26] describe a multimodal temporal attention model that can generate high-accuracy crop maps using diverse data sources, including satellite imagery and time-series data. The model performs better than traditional crop mapping approaches because it uses temporal attention processes to identify which time points are more crucial and should consequently be given greater attention. Maillet et al. [27] describe a methodology for diagnosing downy mildew in plants that combines transformer-based deep learning algorithms with satellite imagery and meteorological data. Using a combination of current weather data and satellite imagery to train a neural network enables accurate and effective detection of plant health problems.

The Production Assessment Method Based on Multimodal and Temporal Networks, a technique for forecasting crop production for various heat-tolerant wheat genotypes, has been covered by Cheng et al. [28]. It combines multiple data modes, weather, and soil conditions, with a temporal network analysis to accurately estimate yield. Ultimately, this technology should help improve crop management and selection of optimal genotypes for maximum wheat production under variable climate conditions. Gené-Mola et al. [29] used multimodal deep learning to combine RGB-D camera data and radiometric data to improve object detection accuracy. This technique has been applied to detect Fuji apples, a popular fruit, to accurately identify and classify them under different lighting conditions, thereby enhancing automated sorting. Cai et al. [30] proposed a dual-branch spatiotemporal fusion network that leverages satellite data from multiple sources to accurately map agricultural field parcels. Integrating spatial and temporal information improves delineation accuracy, allowing for better monitoring and management of crops, as shown in Table 1. This can lead to increased productivity and better conservation practices in agriculture.

Table 1: Comprehensive analysis

Authors	Advantage	Limitation
Cecil [16]	Increased spatial and temporal resolution enables more accurate, more frequent monitoring of crop growth and land-use changes.	"Cloud cover, technology constraints, and the spatial resolution of individual satellite sensors limit accuracy."
Sagan et al. [17]	Improved accuracy of crop yield prediction compared to traditional methods by considering multiple time intervals to capture changes in vegetation growth.	Reliance on historical data may not accurately account for changes in weather patterns or pest/disease outbreaks.
Yao et al. [18]	Integrating multiple data sources can improve the accuracy and reliability of drought monitoring, enabling better decision-making in agriculture.	"Limited availability of high-quality satellite imagery or ground-based data for accurate model training and validation."
Zhang et al. [19]	Enables real-time collection of macro-level plant metrics with high spatial resolution, improving the efficiency and accuracy of crop phenotyping.	One limitation is that unpredictable weather conditions for data collection may limit the use of UAV sensor data.

Lu et al. [20]	Improved disease control by increasing precision in forecasting litchi downy blight damage through the integration of IoT and hyperspectral data.	One limitation of this study is that the prediction model may not accurately account for unforeseen environmental factors or data collection errors.
Barbato et al. [21]	It provides complementary information from different sources, such as spectral, spatial, and temporal, leading to potentially more accurate and detailed segmentation results.	One limitation is the potential for conflicting information arising from differences in sensor resolution, acquisition timing, and environmental conditions.
Fan and Li [22]	Improves accuracy and reliability by combining data from multiple sensors, enabling more precise decision-making in agriculture.	One limitation of Sensor Fusion is that it relies on the accuracy and reliability of the individual sensors used in the fusion process.
Elsherbiny et al. [23]	The achievement of higher accuracy in diagnosing water status in wheat crop due to the use of multiple modalities of data.	The model is only validated for the wheat crop and may not be generalizable to other crop types.
Clamens et al. [24]	Provides accurate, efficient analysis of vineyard health and canopy coverage to improve crop management and disease prevention.	Cross-seasonal variability in vineyards affects the accuracy of the segmentation and detection processes.
Lin et al. [25]	One benefit is that it enables the creation of more thorough and precise models for forecasting agricultural yields under climate change.	One limitation is that it may not capture local or regional factors that can strongly influence crop yields.
Garnot et al. [26]	Enable exploitation of multi-spectral and multi-scale data, resulting in more accurate and robust crop mapping for improved agricultural management.	Difficult to capture important information if remote sensing data does not sufficiently capture crop-specific features and phenology.
Maillet et al. [27]	Integration of multiple data sources enables more accurate, timely detection of downy mildew, improving early warning systems for farmers.	One limitation of the process is that it may not accurately detect the disease in regions where weather data is not consistently available.
Cheng et al. [28]	One advantage of this method is its ability to accurately assess wheat yield across different heat-tolerant genotypes, using multiple modes and over time.	One limitation could be reliance on network-based analysis, which may not fully capture the complexity of wheat yield determinants.
Gené-Mola et al. [29]	The fusion of data from RGB-D cameras and radiometric capabilities enables more robust, accurate detection of Fuji apples.	Limited applicability to other types of apples or fruits due to variability in colour and size.
Cai et al. [30]	"More accurate and detailed field parcel boundaries for improved precision in agricultural planning and management."	Difficulty accessing accurate, up-to-date satellite data can undermine the accuracy and effectiveness of the delineation process.

- **Lack of Standardization:** Using different sensors and data-collection methods requires greater standardization of the collected data. This makes it easier to integrate and compare data from other sources, leading to inaccurate fusion results.
- **Quality of Data:** The accuracy and precision of the data collected by different sensors can vary, leading to discrepancies and errors in the fusion process. Low-quality or noisy data can affect the final predictions and hinder the system's effectiveness.
- **Sensor Redundancy:** In some cases, multiple sensors may capture similar data, leading to redundancy and an unnecessary computational burden. This can impact the system's efficiency and delay results.
- **Data Fusion Methodologies:** Different data fusion techniques may be appropriate for different data types, such as spatial or spectral data. Choosing the wrong fusion method can lead to inaccurate data integration and degrade prediction accuracy.

One major area of research in artificial intelligence is integrating diverse data types to improve performance. The goal of the research was to develop a novel approach that integrates deep learning with spectral analysis to fuse multimodal data sources. The advantage of this integration is that it allows us to collect additional information from different data types and merge them into a single, coherent package. Researchers accomplish this by training several deep learning models using high-level features from each data type, then using spectral analysis methods to merge the resulting feature sets into a single, coherent result. The capabilities of multimodal fusion will significantly improve success rates for Classification, Prediction, and Anomaly Detection tasks by effectively combining information across datasets. In addition, our method is flexible enough to handle missing or incomplete datasets, making it applicable to real-world situations. This represents a significant step forward in AI and has the potential to improve AI system performance across industries at an unprecedented rate.

3. Proposed System

3.1. Construction Model

3.1.1. Backbone Network

The main function of the Backbone Network is to facilitate communication among the network's layers. It establishes a hierarchical structure that enables efficient data transmission between nodes. It is achieved through protocols and routing algorithms that ensure the efficient flow of data. Researchers used the Normalized Difference Vegetation Index (NDVI) derived from MODIS Aqua satellite data to analyze vegetation dynamics:

$$MCTB = \frac{MBL - Re\,d}{MBL + Re\,d} \tag{1}$$

To investigate relationships among crop development, soil nutrients, and associated variables across the research region, a spatial analysis of NDVI patterns relative to soil parameters was conducted. Determine the difference Di between potential evapotranspiration and daily precipitation:

$$C_b = F_b - FGV_b \tag{2}$$

Where Pi is the precipitation, measured in millimeters; Petti is the possible evapotranspiration on the first day, measured in millimeters. Determine the cumulative series of moisture surplus and deficit at various time intervals:

$$C_m^y = \sum_{b=0}^{y-1}(F_{m-b} - FGV_{m-b}), m \geq y \tag{3}$$

Where m is the total number of days and y is the time scale in days. The SPEI value is the normalised value obtained from the Dk n data series. Vicente-Serrano examined how the Generalized Extreme Value, Pearson, Log-logistic, and Log-normal fit the Dk n series:

$$P(h) = \left[1 + (\frac{\alpha}{x-\gamma})^{\beta}\right]^{-1} \tag{4}$$

The findings show that the Log-logistic distribution best fits the Dk n series, and the linear-moment technique is used to estimate the fitting parameters. Determine the SPEI index for each value by normalizing the Di data series utilizing the log-logistic probability distribution with three parameters:

$$\beta = \frac{2\omega_1 - \omega_0}{6\omega_1 - \omega_0 - 6\omega_2} \tag{5}$$

Symbolizes the probability density function; x is its independent variable; κ, β, and γ stand for the scale, shape, and origin parameters, respectively; Γ for the factorial function; and Di for the probability-weighted moment of the data series represented by s:

$$\gamma = \omega_0 - \alpha\Gamma(1 + 1/\beta)\Gamma(1 - 1/\beta) \tag{6}$$

The ordinal number of the probability-weighted moments (s = 0, 1, & 2) is used for N (No. of times calculation was done). The study uses multiple metrics to assess the performance of models developed to identify and grade winter wheat under drought conditions, including the precision of causal classifications (F1), the accuracy of actual drought identification (A1), and an overall assessment score based on the model's total F1:

$$J_1 = \frac{VF + PM}{VF + VM + PF + PM} \tag{7}$$

Precision (P1) is used to evaluate the agreement between actual and predicted causal classification outcomes. Accuracy (A1) is for evaluating the ability of predicting drought, and F1 is the harmonic mean of P1 and 1-A1. This section describes a unique methodology (i.e., adaptive weighted fusion) for creating combined probability vectors for drought using both SPEI and DenseNet-121, thereby improving the overall accuracy and robustness of the models for predicting drought levels:

$$F + \omega_1 F_q + \omega_2 F_C \quad (8)$$

Where ω1 + ω2 = 1 and ω1 and ω2 represent, respectively, the DenseNet-121 and SPEI index techniques' weights, the probability vector obtained by fusing probabilities is also normalised to ensure that the fusion results are consistent and interpretable. The wheat canopy height model was created for each sample period by computing pixel differences between samples and using the same image-based mask to exclude pixels representing the dirt background:

$$TP = \frac{m}{M} \quad (9)$$

With the mean height of the canopy developed using zonal counting from the total number of wheat height measurements collected and using the ratio of wheat pixels identified from the sampled study area compared to the total number of wheat pixels in that sampled study area for each plot, the results were evaluated for the accuracy of the derived canopy height model. An area for each unit image section was calculated by multiplying the number of plants present on the unit image section (plant height) by the total plant heights in the wheat plant images to produce a crop volume model for the wheat crops:

$$CTN = J \times \sum_{b=1}^{m} FX \quad (10)$$

A = Area Per Pixel, n = Wheat Pixel Count in Statistical Area, N = Total Pixel Count in Statistical Area, Phi = Height of Crop on Pixel:

- **Neck Network:** An interface between the Backbone Network and lower levels of the network that converts data from the Backbone Network into a format that can be processed by lower-level networks, ensuring data transfer occurs in a timely and accurate manner throughout the network.
- **Detector:** A class that defines the detection and identification of failures/errors in the network. It continuously monitors data flow and checks for anomalies. The following Figure 1 displays the building diagram.

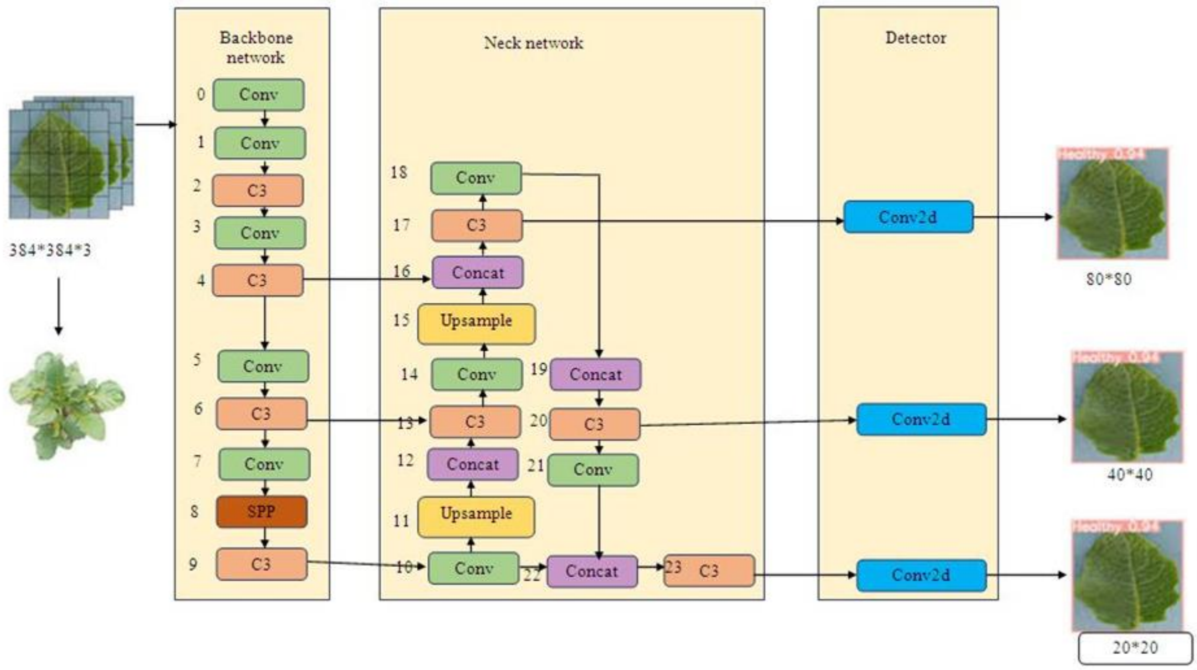

Figure 1: Construction diagram

3.1.2. Classes

A network's backbone consists of the Backbone Network, Neck Network, and Detector class. They perform crucial tasks, such as data conversion, error detection, and communication, that are necessary for the network to run smoothly. Although each class has distinct functions and duties, they all work together to ensure the network operates as well as possible. Canopy temperature information for biomass was predicted by statistically deriving ROI partitioning, which was used to standardize the canopy-air clearance difference and relative canopy temperature from UAV thermal infrared images:

$$MLDV = \frac{V_b - V_{min}}{V_{max} - V_{min}} \tag{11}$$

The field experiment's highest and lowest temperatures were recorded during the experiment. Tax is the highest temperature, and I stands for the INd canopy temperature at the sample point, or sample-point data:

$$DVC = V_b - V_j \tag{12}$$

Ta is the field trial's ambient temperature, and Tin is its lowest recorded temperature. Given the n-set image f, the high-pass image Xu is obtained from a specified number of source images via Laplace filtering:

$$X_m = x_m(p) * R(p) \tag{13}$$

Where is a matrix of mm that depicts the image's log-spectrum. The significance map's computation is frequently not significantly affected by changes in m = 3 and m. The local average of the absolute values of is used to construct the significance map n_Sn of Nh:

$$Q_m = |G_m| * e_{le,\delta e} \tag{14}$$

Let g be the low-pass filter, g is the low-pass filter, and m−1 is the rig size. Lastly, the source picture's initial weights are calculated using the saliency map obtained in the preceding stage:

$$F_m(b,a) = \begin{cases} 1 if\, Q_m(b,a) = max\{\, Q_1(b,a), Q_2(b,a), \ldots Q_m(b,a)\} \\ 0 otherwise \end{cases} \tag{15}$$

Where (,) n S I j indicates the important value in the n Nd picture at pixel (,)I j. Where T is the maximum number of repetitions, t is the current number, and rand is a random number between 0 and 1, and (,) Ij x, and () Best t j are classified as initial and candidate solutions:

$$GQ(v) = 2 \times l_3 \times \left(1 - \frac{1}{V}\right) \tag{16}$$

η(,) I_j indicates the j-position operator in the solution. is a sensitivity parameter set to 0.1 that regulates the individual's detection accuracy during the iteration process. It is employed to train on meteorological data in the MIS-ME architecture. The extraction of meteorological features for MSME:

$$NQNG(H_{met}): N(H_{met}) \to \mathbb{R}^n \tag{17}$$

Here, met ∈ R k M denotes the input vector of meteorological data, and M denotes the sequential neural network operations. The number of extracted characteristics is denoted by m. To predict VWC, these characteristics are then fed into a linear regression layer:

$$Meteo\ Re\ g\, ressionLayer(\mathbb{R}^m): e(\mathbb{R}^n) \to \mathbb{R} \tag{18}$$

Researchers suggest using the MIS-ME framework to improve knowledge-encoded VWC prediction by combining picture and meteorological data. By using this technique, the model may adaptively give the modality with the highest predictive power greater weight:

$$\hat{k} = \alpha . F_{meteo}(nU) + B . f_{image}(bU) \tag{19}$$

Maximizing the integration of several data sources to estimate soil moisture. Two learnable parameters are used in the model prediction formulation. While EVI can reduce atmospheric effects and enhance sensitivity to biomass areas, NDVI is regarded as a classic vegetation index:

$$GTB = 2.5 \times \frac{MBL - L}{MBL + D_1(L) - C_2(I) + R} \tag{20}$$

Where L is the canopy backdrop adjustment, C1 and C2 are coefficients, typically set at 6 and 7.5, respectively, R is red reflectance, B is blue reflectance, and NIR is near-infrared reflectance. The process of upsampling involves two main steps: interpolation and anti-aliasing. Interpolation consists of creating new pixels by analysing the neighbouring pixels in the original image. Different interpolation methods, such as bilinear, bicubic, and Lanczos, can be used. Each method has its own algorithm for generating new pixels based on the surrounding pixels.

3.2. Functional Working System

- **Upsample:** Upsampling increases the resolution and size of an image, typically in a convolutional neural network (CNN) architecture. It is achieved by multiplying each pixel by a factor and inserting zeros to create space between the pixels.
- **Concat:** Concatenation is one of the methods that allows layers in Convolutional Neural Networks (CNNs) to leverage information from previous layers. When combined, layers provide a different view of the same inputs than when kept separate. As a result, the network can produce improved results.
- **MaxPool:** MaxPooling, a common operation in CNNs, allows feature maps to be downsampled. MaxPooling splits feature maps into smaller subregions on a 2D grid. Each subregion will then return only the maximum value from that subregion, as depicted in Figure 2.

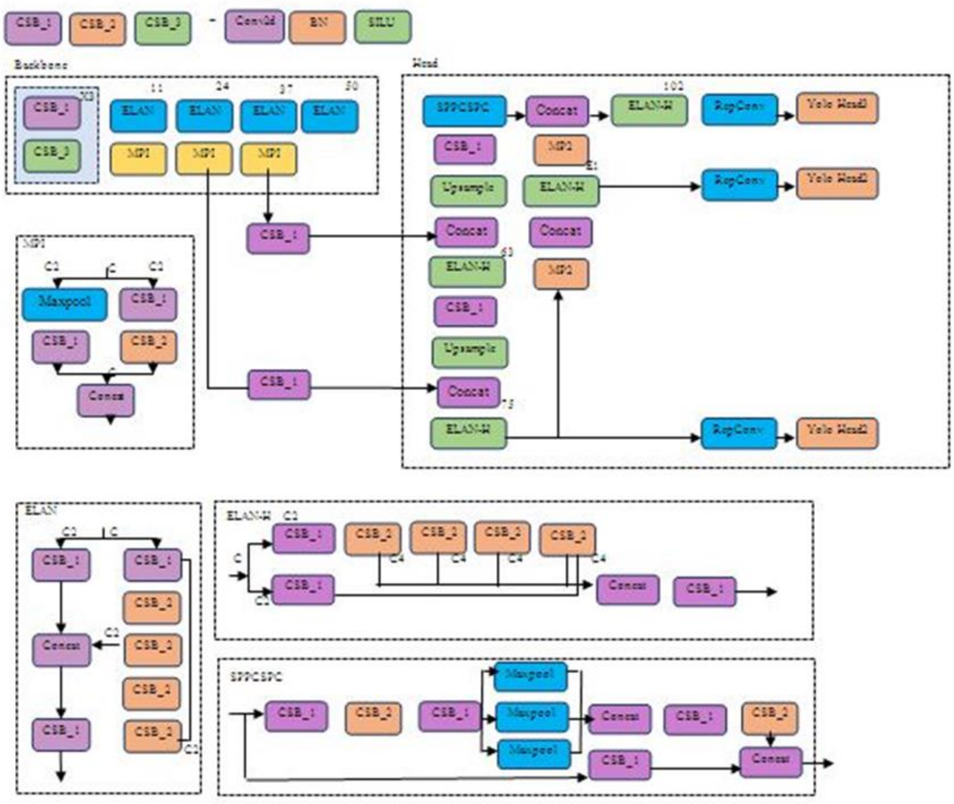

Figure 2: Functional block diagram

- **ELAN:** The ELAN (Efficient Light Attention Network) is a new architecture for object detection, featuring an efficient, lightweight backbone and an attention module that processes feature maps to generate accurate predictions.
- **Yolo Head:** The Yolo Head is the final layer of the YOLO (You Only Look Once) architecture, which is used to detect and localise objects in an image.
- **ELAN-H:** An ELAN is a type of ELAN Architecture (that is, an ELAN with an ELAN-H as opposed to simply ELAN) with an ELAN-H as a Module. ELAN-H will provide enhanced prediction capabilities by using an Hourglass Module to improve feature extraction from the data. The Hourglass Module contains numerous up- and down-samples, which improve the accuracy of multiple-scale feature fusion and data predictions.

The percentage of photosynthetically active radiation (FPAR) and leaf area index (LAI), the other two products, were also collected:

$$PFJL = 1 - e^{-0.5(RJB)} \quad (21)$$

The plant's evapotranspiration capacity is directly related to LAI, the amount of leaf area per unit of ground area. However, the fraction of photosynthetically active radiation that a green plant receives is known as FPAR. where Pi is the value that the model predicts, Ri is the reported or observed value, mean is the mean of all reported values, and n is the total number of samples:

$$NJG = \frac{\sum_{i=1}^{m} |L_b - F_b|}{m} \tag{22}$$

Ranges from 0 to 1, where 1 denotes optimal model performance and full agreement between reported and predicted values. The model prediction equation represents the probability density function of a normal distribution:

$$p_b(h) = \frac{1}{\sqrt{2\pi\sigma_b}} \tag{23}$$

In this case, fi(x) represents the probability density function of the itch model for the variable x. prediction formula following the combination of two models. The fused prediction is represented by the function f(x), which is the outcome of fusing two normal distributions:

$$p(h) = \frac{1}{2\pi\sigma_1\sigma_2} \tag{24}$$

By combining the means and variances from both models into a single exponent, the combined effect of both distributions on the variable x is suggested. The function's inclusion center, which represents the prediction model's average value, is where f'(x) = 0:

$$\mu' = \mu_1 + \frac{\sigma_1^2(\mu_2 - \mu_1)}{\sigma_1^2 + \sigma_2^2} \tag{25}$$

When two models are fused, this formula yields the new mean μ ′. It is a weighted average of the individual models' averages (μ1 and μ2), with weights determined by the variances. The function's degree of dispersion is zero if f ″(x) = 0:

$$\sigma'^2 = \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2 + \sigma_2^2} \tag{26}$$

The prediction model's variance, or the function's level of dispersion, may be found as follows if f ″(x) = 0. A new prediction model is created when several models are combined. By doing this, the idea of combining two models is expanded to include I models:

$$f(x) = \frac{1}{2\pi\sigma_1\sigma_2 \dots \sigma_b} \tag{27}$$

The effect of I distinct normal distributions is incorporated into the final fused prediction model, f(x). The multi-model integration approach, which fuses a dynamic variance KF, was developed to address the need for vegetation index monitoring models with varying LAI to account for inversion errors within specific ranges:

$$LNQG = \sqrt{\frac{\sum_{b=1}^{n} (U_b - F_b)^2}{m}} \tag{28}$$

Four experimental groups were created from the nine previously described experiments. The association between the computed vegetation index and the rice LAI value was investigated using correlation analysis:

$$L^2 = \left(\frac{\sum_{b=1}^{m} (U_b - \overline{U})(F_b - \overline{F}}{\sqrt{\sum_{b=1}^{m} (U_b - \overline{U})^2}} \right)^2 \tag{29}$$

The input variables for the previously described data fusion procedure were then selected from a collection of vegetation indicators that showed significant associations. Coefficients of determination (R2), root mean square error (RMSE), and mean absolute percentage error are used to evaluate the precision of data fusion:

$$NJFG = \frac{100\%}{m} \sum_{b=1}^{m} \left| \frac{F_b - U_b}{U_b} \right| \quad (30)$$

In this case, n is the number of samples, and O and P denote the measured and projected values, respectively:

- **CSB 1:** Conditional Scale Block 1 (CSB 1) is a scale-adaptive module used in the ELAN architecture. It adapts the scale and resolution of the feature map to the sizes of objects in the image, enabling better representation of objects at different scales.

3.3. Operating Principles

- **Concatenation:** The operation Concatenation joins different input vectors into one output. In the case of the above texts, it could be seen as merging or joining together different inputs to create an output value. In ML systems, Concat is most commonly used to combine feature maps or layers from NNs, enabling the capture of more complex patterns.
- **Max Pooling:** A common method you also use in your Convolutional neural networks to reduce the size of your feature maps while preserving relevant information. It involves dividing the feature map into non-overlapping regions and only keeping the maximum value from each region.
- **Up-Sampling:** The inverse of pooling, up-sampling enlarges feature maps. Zeros are inserted in between pixels, and new values are then interpolated using a convolutional layer. The operating flow diagram is shown in Figure 3.

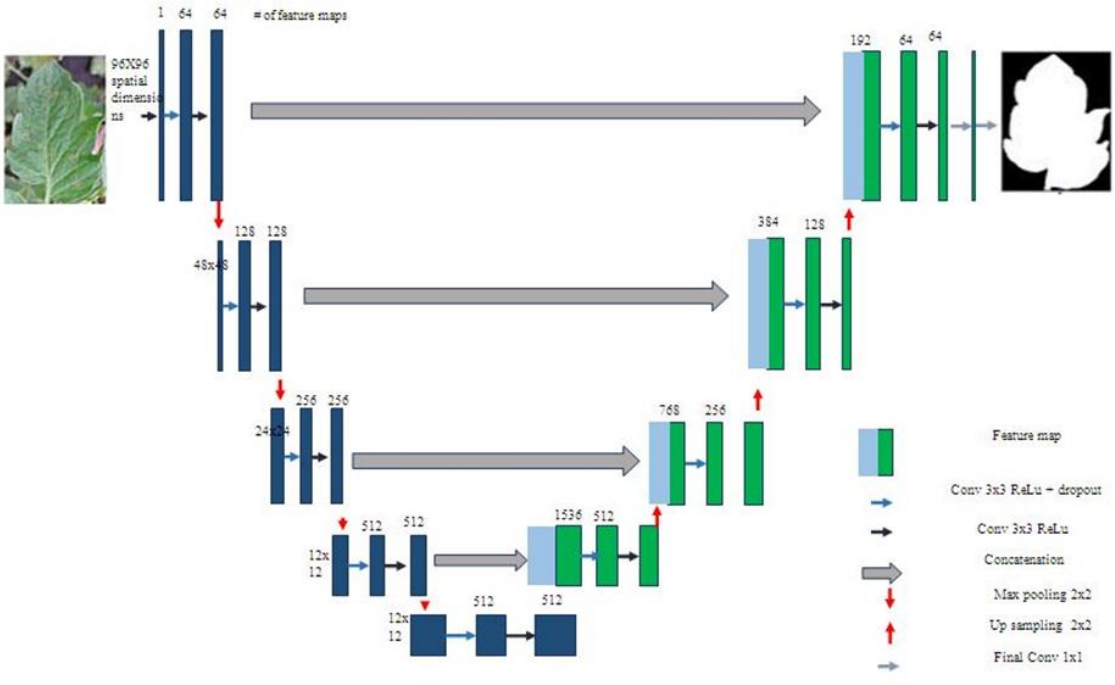

Figure 3: Operational flow diagram

- **InceptionNet:** The InceptionNet is a popular deep convolutional neural network architecture called GoogLeNet. It uses the concept of 'Inception modules,' which are blocks of layers (such as convolutional, pooling, and concatenation layers) run in parallel, allowing for more diverse and effective feature extraction.
- **Healthy and Unhealthy:** These terms often refer to the output of a classification task in medical imaging. In this context, 'healthy' typically means that the image is normal or shows no signs of disease, while 'unhealthy' indicates the presence of abnormalities or potential pathologies.

4. Result and Discussion

The proposed DL-PCA: Deep Learning for Principal Component Analysis has been compared with the existing SA-DLP (Spectral Analysis Deep Learning Precision Agriculture), ML-DDP (Multi-Layered Deep Learning for Data Fusion and Prediction), and IS-CHP (Integrative Spectral Analysis for Crop Health Prediction).

4.1. Absorption Bands

The absorption bands to be measured for integrating deep learning and spectral analysis in precision agriculture include near-infrared, visible, and thermal bands. These bands may offer useful data on soil characteristics, water content, and plant health, supporting crop health monitoring and yield forecasting.

Table 2: Comparison of absorption bands (in %)

No. of Inputs	SA-DLP	ML-DDP	IS-CHP	DL-PCA
100	70.56	64.53	74.70	77.20
200	72.23	66.27	76.12	78.80
300	72.66	68.61	77.37	81.00
400	73.91	69.43	79.36	82.60
500	76.06	71.70	81.81	83.74

Table 2 and Figure 4 below show the comparative absorption spectra of the proposed model and the current ELAN-H model.

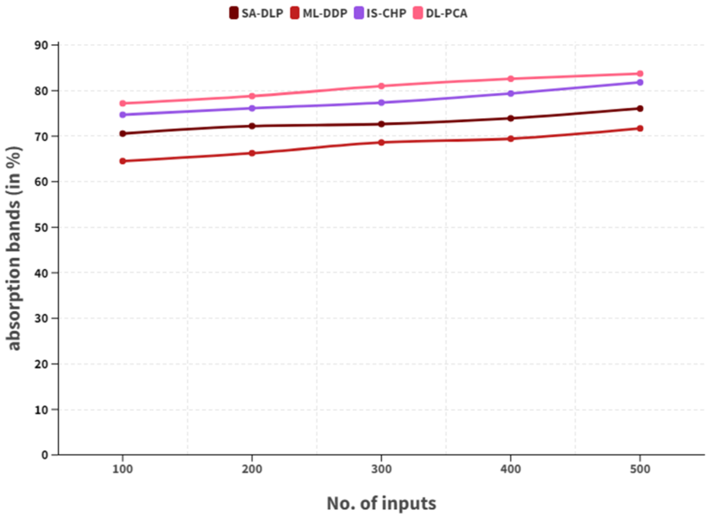

Figure 4: Absorption bands

4.2. Drop-Out Rate

Table 3 and Figure 5 show the comparative dropout rates of the proposed and current ELAN models, as they relate to the correlation between deep learning and spectral analysis, including, but not limited to, precision agriculture, multi-modal data fusion, and accurate crop health and yield predictions.

Table 3: Comparison of drop-out rate (in %)

No. of Inputs	SA-DLP	ML-DDP	IS-CHP	DL-PCA
100	74.56	68.53	79.20	78.70
200	76.23	70.26	80.10	80.10
300	76.65	72.61	80.00	81.36
400	77.94	73.41	82.63	83.35
500	80.01	75.72	85.76	85.84

Figure 5 shows how the number of inputs affects the dropout rate (%) for four methods: SA-DLP, ML-DDP, IS-CHP, and DL-PCA. As the number of inputs increases from 100 to 500, the dropout rate rises slowly across all techniques, indicating a steady upward trend. DL-PCA and IS-CHP are the best methods, as they maintain drop-out rates above 80% across most input levels. SA-DLP shows some improvement, but not as much as expected. ML-DDP, on the other hand, always has the lowest drop-out rates, but it also follows the same pattern of growing. Overall, Figure 5 shows that DL-PCA and IS-CHP perform better than the other methods at higher input sizes. This suggests that they are more robust and efficient.

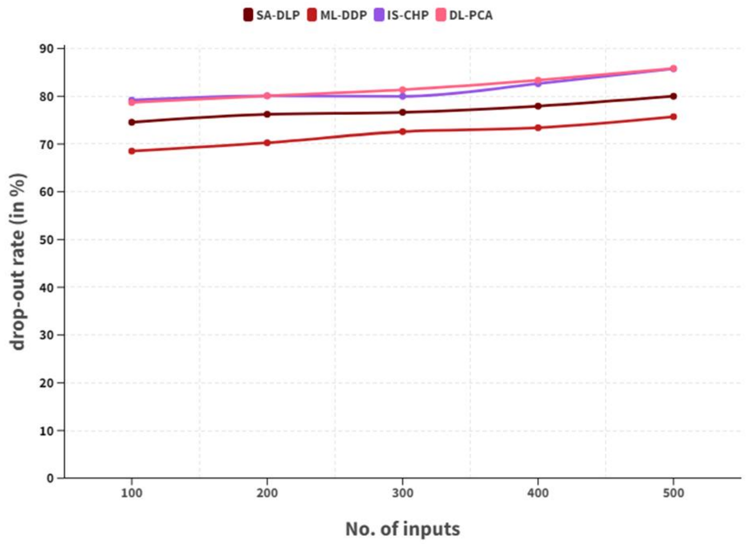

Figure 5: Drop-out rate

4.3. Field Topography

Topographic data for the fields should be quantified as a primary metric when integrating deep learning (DL) with spectral analysis (SA) in the field of precision agriculture; it will supply useful information concerning terrain, soil properties, and the movement of water, providing better monitoring of crop health and predicting yield through accurate and targeted interventions.

Table 4: Comparison of field topography (in %)

No. of Inputs	SA-DLP	ML-DDP	IS-CHP	DL-PCA
100	77.52	72.53	80.22	81.70
200	79.24	74.27	82.86	83.15
300	79.66	76.61	82.00	84.33
400	80.90	77.45	85.61	86.34
500	83.01	79.79	88.76	88.87

Table 4 and Figure 6 present the topographic data for both the current and suggested models.

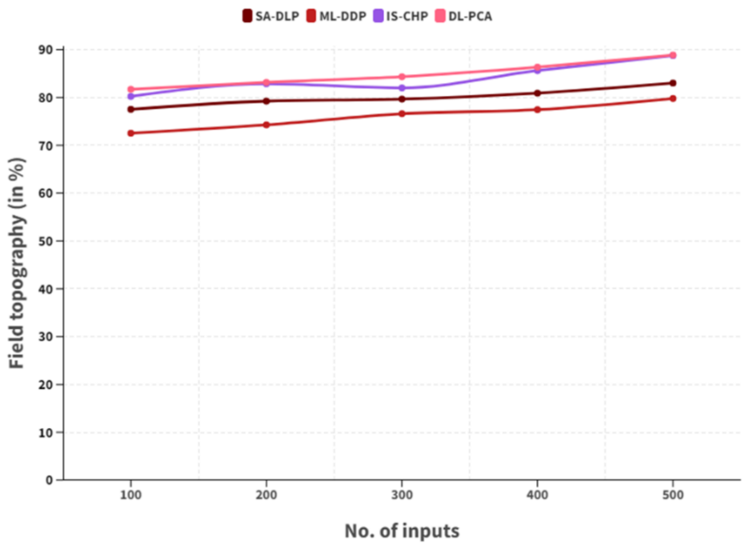

Figure 6: Field topography

4.4. Data Preprocessing

To enable accurate and reliable analysis, DL/Spectral Analysis integration requires extensive data pre-processing (cleaning, normalization, and transformation). The Fourier transform of the pre-processed data is also necessary to enhance the accuracy of prediction models and the validity of hypothesis testing.

Table 5: Comparison of data preprocessing (in %)

No. of Inputs	SA-DLP	ML-DDP	IS-CHP	DL-PCA
100	82.50	76.52	83.21	84.75
200	84.20	78.24	85.86	86.14
300	84.61	80.64	86.03	87.38
400	85.96	81.43	88.67	89.31
500	88.04	83.75	91.72	91.84

The data preprocessing for both the suggested and currently implemented models is presented in Table 5 and Figure 7.

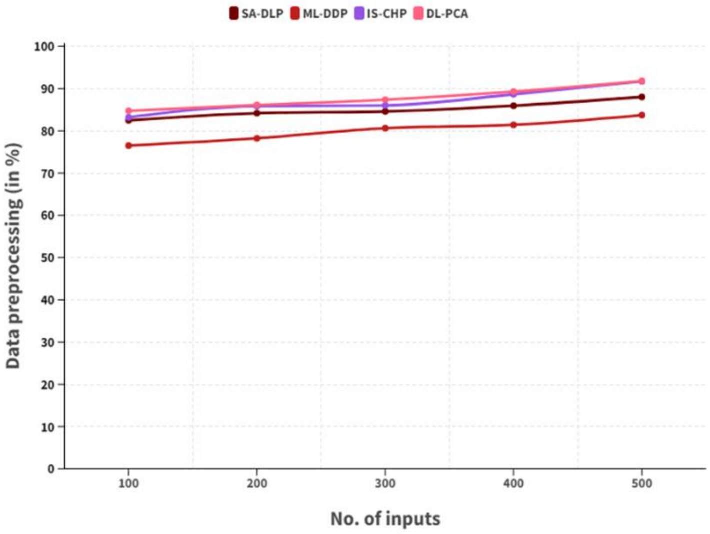

Figure 7: Data preprocessing

5. Conclusion

In conclusion, combining the two technologies could greatly enhance crop health and yield predictions by making it easier to obtain accurate measurements of crop health and production from diverse data sources. This combination strategy uses a wide range of information, including satellite images, sensor-based field data, weather patterns, and historical agricultural records, to provide a more complete and accurate picture of crop conditions. This not only makes it easier to make predictions but also reduces the uncertainty that farmers and agricultural planners typically face. So, changes in these areas will continue to affect macroeconomic management and decision-making regarding crops and yields, helping policymakers and other stakeholders develop smart, data-driven plans. In addition, using both of these modern technologies will open up many new possibilities for precision agriculture by improving monitoring and forecasting. This will enable the adoption of long-term, sustainable farming methods. Farmers may make better use of resources such as water, fertilisers, and pesticides, which helps the environment and makes work more efficient. Also, real-time insights and predictive analytics enable early detection of diseases, pests, and stress factors, allowing interventions to be implemented quickly and losses to be kept to a minimum. In general, this integration opens up exciting possibilities for the future of precision agriculture and the agricultural industry as a whole. It will lead to higher productivity, resilience, and sustainability in global food systems.

Acknowledgement: The authors collectively extend their sincere gratitude to Galgotias University, Madan Mohan Malaviya University of Technology, and the Institute of Tijuana for their continuous support, guidance, and academic environment, which made this research possible.

Data Availability Statement: Relevant data used in this study are available upon justified request to the corresponding author, with restrictions applied to protect participant confidentiality.

Funding Statement: The authors conducted this research without any external financial support or sponsorship.

Conflicts of Interest Statement: The authors declare that there are no conflicts of interest that may have influenced the results or interpretation of this research.

Ethics and Consent Statement: The research was conducted ethically, with all participants providing informed consent and their rights and privacy being strictly maintained.

References

1. U. Shafi, R. Mumtaz, N. Iqbal, S. M. H. Zaidi, S. A. R. Zaidi, I. Hussain, and Z. Mahmood, "A multi-modal approach for crop health mapping using low altitude remote sensing, internet of things (IoT) and machine learning," *IEEE Access*, vol. 8, no. 6, pp. 112708–112724, 2020.
2. M. Maimaitijiang, "Multimodal Remote Sensing Data Fusion and Machine Learning for Crop Monitoring and Food Security," PhD dissertation, Saint Louis Univ., Missouri, United States of America, 2020.
3. S. Fei, M. A. Hassan, Y. Xiao, X. Su, Z. Chen, Q. Cheng, F. Duan, R. Chen, and Y. Ma, "UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat," *Precision Agriculture*, vol. 24, no. 1, pp. 187–212, 2023.
4. A. D. Yewle, L. Mirzayeva, and O. Karakus, "Multi-modal Data Fusion and Deep Ensemble Learning for Accurate Crop Yield Prediction," *Remote Sensing Applications: Society and Environment*, vol. 38, no. 4, p. 101613, 2024.
5. K. Xu, Y. Zhu, W. Cao, X. Jiang, Z. Jiang, S. Li, and J. Ni, "Multi-modal deep learning for weeds detection in wheat field based on RGB-D images," *Frontiers in Plant Science*, vol. 12, no. 11, p. 732968, 2021.
6. A. D. Yewle and O. Karakus, "Enhanced Crop Yield Prediction: Leveraging Multi-Modal Remote Sensing Imagery and Ensemble Machine Learning Models," *SSRN*, 2024. [Accessed by 06/02/2025].
7. R. R. Patil and S. Kumar, "Rice-fusion: A multimodality data fusion framework for rice disease diagnosis," *IEEE Access*, vol. 10, no. 1, pp. 5207–5222, 2022.
8. M. M. Ali, N. Hashim, S. A. Aziz, and O. Lasekan, "Utilisation of deep learning with multimodal data fusion for determination of pineapple quality using thermal imaging," *Agronomy*, vol. 13, no. 2, p. 401, 2023.
9. M. F. Danilevicz, P. E. Bayer, F. Boussaid, M. Bennamoun, and D. Edwards, "Maize yield prediction at an early developmental stage using multispectral images and genotype data for preliminary hybrid selection," *Remote Sensing*, vol. 13, no. 19, p. 3976, 2021.
10. X. Lu, W. Li, J. Xiao, H. Zhu, D. Yang, J. Yang, X. Xu, Y. Lan, and Y. Zhang, "Inversion of Leaf Area Index in Citrus Trees Based on Multi-Modal Data Fusion from UAV Platform," *Remote Sensing*, vol. 15, no. 14, p. 3523, 2023.
11. Z. Lin, R. Fu, G. Ren, R. Zhong, Y. Ying, and T. Lin, "Automatic monitoring of lettuce fresh weight by multi-modal fusion based deep learning," *Frontiers in Plant Science*, vol. 13, no. 8, p. 980581, 2022.
12. L. Strani, C. Durante, M. Cocchi, F. Marini, I. Måge, and A. Biancolillo, "Data fusion strategies for the integration of diverse non-destructive spectral sensors (NDSS) in food analysis," *TrAC Trends in Analytical Chemistry*, vol. 180, no. 11, p. 117957, 2024.
13. M. Qiao, X. He, X. Cheng, P. Li, H. Luo, L. Zhang, and Z. Tian, "Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks," *Int. J. Applied Earth Observation and Geoinformation*, vol. 102, no. 10, p. 102436, 2021.
14. J. Deng, D. Hong, C. Li, J. Yao, Z. Yang, Z. Zhang, and J. Chanussot, "RustQNet: Multimodal deep learning for quantitative inversion of wheat stripe rust disease index," *Computers and Electronics in Agriculture*, vol. 225, no. 10, p. 109245, 2024.
15. S. Sarkar, A. Dey, R. Pradhan, U. M. Sarkar, C. Chatterjee, A. Mondal, and P. Mitra, "Crop Yield Prediction Using Multimodal Meta-Transformer and Temporal Graph Neural Networks," *IEEE Trans. AgriFood Electronics*, vol. 2, no. 2, pp. 545-553, 2024.
16. M. Cecil, "Multi-Modal Earth Observation for Scalable Monitoring of Smallholder Agriculture in Data Sparse Settings," PhD dissertation, Clark Univ., Massachusetts, United States of America, 2024.
17. V. Sagan, M. Maimaitijiang, S. Bhadra, M. Maimaitiyiming, D. R. Brown, P. Sidike, and F. B. Fritschi, "Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning," *ISPRS J. Photogrammetry and Remote Sensing*, vol. 174, no. 4, pp. 265–281, 2021.
18. J. Yao, Y. Wu, J. Liu, and H. Wang, "Multimodal deep learning-based drought monitoring research for winter wheat during critical growth stages," *PLOS ONE*, vol. 19, no. 5, p. e0300746, 2024.

19. Y. Zhang, Y. Yang, Q. Zhang, R. Duan, J. Liu, Y. Qin, and X. Wang, "Toward multi-stage phenotyping of soybean with multimodal UAV sensor data: A Comparison of Machine Learning Approaches for Leaf Area Index Estimation," *Remote Sensing*, vol. 15, no. 1, p. 7, 2023.
20. J. Lu, Z. Wu, Y. Lan, X. Deng, and J. Huang, "Study on the Prediction Model of Litchi Downy Blight Damage Based on IoT and Hyperspectral Data Fusion," *IEEE Internet of Things Journal*, vol. 11, no. 16, pp. 27184-27200, 2024.
21. M. P. Barbato, F. Piccoli, and P. Napoletano, "Ticino: A multi-modal remote sensing dataset for semantic segmentation," *Expert Systems with Applications*, vol. 249, no. 9, p. 123600, 2024.
22. S. Fan and C. Li, "Sensor Fusion," in Encyclopedia of Digital Agricultural Technologies, *Springer*, Cham, Switzerland, 2023.
23. O. Elsherbiny, L. Zhou, Y. He, and Z. Qiu, "A novel hybrid deep network for diagnosing water status in wheat crop using IoT-based multimodal data," *Computers and Electronics in Agriculture*, vol. 203, no. 12, p. 107453, 2022.
24. T. Clamens, J. Rodriguez, M. Delamare, L. F. C. L. Y. Voon, E. Fauvet, and D. Fofi, "YOLO-based Multi-Modal Analysis of Vineyards using RGB-D Detections," in *5th International Conference on Advances in Signal Processing and Artificial Intelligence (ASPAI' 2023)*, Tenerife, Spain, 2023.
25. F. Lin, K. Guillot, S. Crawford, Y. Zhang, X. Yuan, and T. Nian-Feng, "An Open and Large-Scale Dataset for Multi-Modal Climate Change-aware Crop Yield Predictions," *arXiv preprint*, 2024. [Accessed by 17/02/2025].
26. V. S. F. Garnot, L. Landrieu, and N. Chehata, "Multi-modal temporal attention models for crop mapping from satellite time series," *ISPRS J. Photogrammetry and Remote Sensing*, vol. 187, no. 5, pp. 294–305, 2022.
27. W. Maillet, M. Ouhami, and A. Hafiane, "Fusion of satellite images and weather data with transformer networks for downy mildew disease detection," *IEEE Access*, vol. 11, no. 1, pp. 5406–5416, 2023.
28. T. Cheng, M. Li, L. Quan, Y. Song, Z. Lou, H. Li, and X. Du, "A Multimodal and Temporal Network-Based Yield Assessment Method for Different Heat-Tolerant Genotypes of Wheat," *Agronomy*, vol. 14, no. 8, p. 1694, 2024.
29. J. Gené-Mola, V. Vilaplana, J. R. Rosell-Polo, M. Josep-Ramon, J. Ruiz-Hidalgo, and E. Gregorio, "Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities," *Computers and Electronics in Agriculture*, vol. 162, no. 7, pp. 689–698, 2019.
30. Z. Cai, Q. Hu, X. Zhang, J. Yang, H. Wei, J. Wang, Y. Zeng, G. Yin, W. Li, L. You, B. Xu, and Z. Shi, "Improving agricultural field parcel delineation with a dual branch spatiotemporal fusion network by integrating multimodal satellite data," *ISPRS J. Photogrammetry and Remote Sensing*, vol. 205, no. 11, pp. 34–49, 2023.