

Article

MangiSpectra: A Multivariate Phenological Analysis Framework Leveraging UAV Imagery and LSTM for Tree Health and Yield Estimation in Mango Orchards

Muhammad Munir Afsar ^{1,*} , Muhammad Shahid Iqbal ² , Asim Dilawar Bakhshi ³ , Ejaz Hussain ¹  and Javed Iqbal ¹ 

¹ Institute of Geographical Information Systems, National University of Sciences and Technology, Islamabad 44000, Pakistan; ejaz@igis.nust.edu.pk (E.H.); javed@igis.nust.edu.pk (J.I.)

² Land Information Management System, Rawalpindi 46000, Pakistan; shahid.iqbal@limspakistan.org

³ Department of Electrical Engineering, Military College of Signals, National University of Science and Technology, Rawalpindi 46000, Pakistan; asim.dilawar@mcs.edu.pk

* Correspondence: mafsar.phd18igis@igis.nust.edu.pk

Abstract: Mango (*Mangifera Indica* L.), a key horticultural crop, particularly in Pakistan, has been primarily studied locally using low- to medium-resolution satellite imagery, usually focusing on a particular phenological stage. The large canopy size, complex tree structure, and unique phenology of mango trees further accentuate intrinsic challenges posed by low-spatiotemporal-resolution data. The absence of mango-specific vegetation indices compounds the problem of accurate health classification and yield estimation at the tree level. To overcome these issues, this study utilizes high-resolution multi-spectral UAV imagery collected from two mango orchards in Multan, Pakistan, throughout the annual phenological cycle. It introduces MangiSpectra, an integrated two-staged framework based on Long Short-Term Memory (LSTM) networks. In the first stage, nine conventional and three mango-specific vegetation indices derived from UAV imagery were processed through fine-tuned LSTM networks to classify the health of individual mango trees. In the second stage, associated data such as the trees' age, variety, canopy volume, height, and weather data were combined with predicted health classes for yield estimation through a decision tree algorithm. Three mango-specific indices, namely the Mango Tree Yellowness Index (MTYI), Weighted Yellowness Index (WYI), and Normalized Automatic Flowering Detection Index (NAFDI), were developed to measure the degree of canopy covered by flowers to enhance the robustness of the framework. In addition, a Cumulative Health Index (CHI) derived from imagery analysis after every flight is also proposed for proactive orchard management. MangiSpectra outperformed the comparative benchmarks of AdaBoost and Random Forest in health classification by achieving 93% accuracy and AUC scores of 0.85, 0.96, and 0.92 for the healthy, moderate and weak classes, respectively. Yield estimation accuracy was reasonable with $R^2 = 0.21$, and RMSE = 50.18. Results underscore MangiSpectra's potential as a scalable precision agriculture tool for sustainable mango orchard management, which can be improved further by fine-tuning algorithms using ground-based spectrometry, IoT-based orchard monitoring systems, computer vision-based counting of fruit on control trees, and smartphone-based data collection and insight dissemination applications.



Academic Editor: Justin F Moat

Received: 4 December 2024

Revised: 13 February 2025

Accepted: 17 February 2025

Published: 19 February 2025

Citation: Afsar, M.M.; Iqbal, M.S.; Bakhshi, A.D.; Hussain, E.; Iqbal, J. MangiSpectra: A Multivariate Phenological Analysis Framework Leveraging UAV Imagery and LSTM for Tree Health and Yield Estimation in Mango Orchards. *Remote Sens.* **2025**, *17*, 703. <https://doi.org/10.3390/rs17040703>

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

Keywords: precision agriculture; mango orchard management; UAV multi-spectral imagery; vegetation indices; plant health classification; yield prediction; LSTM

1. Introduction

Precision agriculture is increasingly adopted for resource optimization, productivity enhancement, and minimizing environmental impact [1]. Recent technological advancements in remote sensing and data analytics have accelerated this adoption [2]. Conventional vegetation indices, such as the Normalized Difference Vegetation Index (NDVI) and the Green Normalized Difference Vegetation Index (GNDVI), derived from remote sensing satellite imagery, have proven applicability in assessing crops and plant vigor by detecting stress conditions before these become visually apparent [3,4]. While applying this satellite imaging-based analytical paradigm, perennial fruit crops like mango present unique challenges [5–7]. Mango trees have complex canopy structures with dense and heterogeneous foliage leading to occlusion [8]. In addition space between the trees often has crops planted or there is other vegetative growth [9,10]. This makes it difficult to fully discriminate underlying tree health conditions in low- to medium-resolution satellite imagery due to spectral mixing [11]. Mango orchards also exhibit intra-orchard phenological variability due to tree age, cultivar, or micro-climatic conditions [12,13]. To study and cater for this variability, remote sensing data at higher temporal frequency and spatial resolution is required to effectively discriminate spectral responses in various phenological stages [14]. Additional factors like planting distance, tree age, structure and canopy area, orchard management practices, and environmental stress also impact tree health and yield estimates [15]. This scenario warrants accessible and affordable higher-spatiotemporal-resolution data, which are often not available through satellite remote sensing, especially to the practitioners and orchard managers in the field.

Unmanned aerial vehicles (UAVs) with multi-spectral and RGB cameras and other sensors have demonstrated their utility in precision agriculture by addressing a few of the above issues. Vegetation indices derived from unmanned aerial vehicles (UAVs) have higher spatiotemporal resolution as compared to satellite imagery [16]. Thus, UAV-based imagery analysis enables near real-time monitoring and dynamic crop management during various growth stages [17]. UAV-based imagery has been applied effectively for the study of a range of crops, including cereals, oil seeds, and fruit trees [18]. Mango yield at orchard scale has been mapped using RGB UAV data by applying geographic object-based image analysis (GEOBIA) and a canopy height model (CHM) [19]. While higher-resolution UAV imagery presents a promising solution, its large data volume presents new challenges due to the limitations of traditional imagery analysis techniques such as GEOBIA. This requires the adoption of new data processing pipelines based on artificial intelligence algorithms.

Machine learning and deep learning approaches have been extensively utilized within precision agriculture from disease detection to yield prediction [20,21]. Normally, the studies based on these approaches have used data of one snapshot from the crop's phenological cycle due to the construct of the algorithms, availability of the data, or objective of the study [22,23]. In such scenarios, normal machine learning and deep learning models may ignore the sequential dependencies in the crop growth cycle [24]. Recurrent neural networks (RNNs), specifically Long Short-Term Memory (LSTM) networks, are well designed to effectively model temporal patterns [25]. LSTM networks incorporate memory cells to retain information over extended periods and across stages, making them suitable for time-series analysis where stage-wise context is crucial for classification and estimation [26].

A review of diverse LSTM applications in the context of precision agriculture (Table 1) demonstrates their effectiveness both for normal crops, as well as fruit trees. Studies on maize, wheat, and soybean have successfully used LSTM to estimate crop yield by incorporating time-series vegetation indices and meteorological conditions [27–29]. In irrigation scheduling, LSTM has been deployed with Sentinel-2 LAI and soil moisture data to forecast irrigation needs in maize fields, outperforming traditional regression-

based models in predictive accuracy [30]. Similarly, an LSTM model trained on IoT-based matric potential and evapotranspiration data has been used in mango orchards to improve irrigation forecasting and facilitate more efficient water use [31].

For fruit orchards, hybrid models combining LSTM with convolutional neural networks (CNNs) have been applied to predict cocoa yields by integrating climatic variables with past production trends, thereby reducing forecasting errors as compared to statistical methods [23]. In citrus yield estimation, UAV-based fruit counting was combined with time-series analysis to predict orchard productivity, achieving minimal deviation from actual harvest data [32]. An LSTM-integrated CNN model significantly improved mango disease classification with higher accuracy than conventional classifiers in different infection stages [33].

Table 1. LSTM applications for disease classification, resource optimization, and yield estimation in precision agriculture.

Crop Type	LSTM Application	Data and Study Area	Results
Maize, Rice, Soybean	Yield estimation [27]	MODIS indices and meteorological data (China)	$R^2 = 0.76\text{--}0.80$ RMSE = 375–519 kg/ha
Paddy, Maize, Ragi, Sugarcane, Cotton	Country-level yield forecasting [29]	Historical rainfall and yield data (India)	$R^2 = 0.99$ RMSE = 1.14–2.31
Wheat	Yield estimation [28]	MODIS LAI time-series (US)	$R^2 = 0.88$ RMSE = 532 kg/ha
Maize	Optimizing irrigation schedule [30]	Sentinel-2 LAI, soil moisture, and weather data (Turkey)	$R^2 = 0.92$ RMSE = 1.03
Mango	Optimizing irrigation schedule [31]	IoT-based matric potential and evapotranspiration (Colombia)	$R^2 = 0.99$ RMSE = 1.56
Cocoa	Yield estimation [23]	Climatic and production data (Nigeria)	RMSE = 28.41 MAPE = 9.77%
Citrus	Orchard yield estimation [32]	UAV-based fruit counting with time-series analysis (Brazil)	SE = 4.53%
Mango	Multi-disease classification of mango leaves and fruit [33]	UAV imagery, optimized CNN-LSTM hybrid (Benchmark dataset)	MCC = 84%

In the case of mango, satellite and UAV imagery data have been employed through traditional and machine learning methods [5,34,35]. The incorporation of environmental factors into the modeling process has allowed better insights regarding tree growth and fruit development at different phenological stages [36]. Despite research advancements, several gaps exist particularly related to mango orchards. There have been few studies that have employed high-resolution satellite imagery or UAV RGB imagery. However, general reliance is still on low- to medium-resolution satellite imagery, thereby lacking the required precision for tree-level health assessment. While LSTM models have proven effective in analyzing temporal sequences of environmental variables, their application

in phenology-based mango yield prediction remains largely unexplored. Existing models fail to integrate stage-specific variations in flowering, fruit development, and canopy conditions, thus limiting their ability to track progression over time. By predominantly relying on historical trends and canopy structure, yield prediction studies have also not leveraged the higher-resolution vegetation indices from UAV imagery which can afford better stage-wise insights.

Detection of vegetation colors during the phenological cycle is also an important technique for the identification of critical stages like flowering, senescence, and disease impacts [37]. For mango crops, color threshold techniques, Support Vector Machines (SVMs), and convolutional neural network (CNN)-based Faster R-CNN have been employed to assess flowering intensity, for automated panicle detection, and to classify growth stages [38]. A study has used GNDVI derived from UAV multi-spectral imagery to assess thrips damage by identifying color changes in mango leaves [14]. However, normal vegetation indices relying on chlorophyll levels often underperform when used to detect flowers in the case of mangoes because both new leaves and flowers exhibit lower chlorophyll content [39]. With the amount of flowering a good indicator of tree vigor, developing a dedicated index to measure the mango tree canopy area covered by flowers can potentially enhance the accuracy of health and yield estimation models.

To address the identified research gaps, this study introduces MangiSpectra, an innovative Multivariate Phenological Analysis Framework that leverages UAV imagery integrated with LSTM for mango orchards and tree monitoring. The primary objectives are (1) to explore and develop a phenological stage-based LSTM model for health classification; (2) to introduce three novel vegetation indices for flowering detection, namely the Mango Tree Yellowness Index (MTYI), Weighted Yellowness Index (WYI), and Normalized Adjusted Flowering Detection Index (NAFDI); (3) to propose a Cumulative Health Index (CHI) to monitor tree health for prompt in-season interventions; and (4) to enhance the accuracy of yield estimation using decision trees.

2. Data Acquisition and Processing

2.1. Study Area

This research was conducted on two adjacently located mango orchards in the Multan region of Pakistan (30.1575°N, 71.5249°E), which is a key area for commercial mango production in the country (Figure 1). Each orchard has over 1300 mature, fruiting mango trees of local varieties primarily *Chaunsa Samar Bahisht* (CSB), *Sindhri* (Si), *Samar Chaunsa* (SC), and *Langra*, with sporadic presence of other varieties. The age of fruiting trees ranges from 5 to over 40 years, with a median age of 27 years. Trees have a median height of 25 feet with several trees over 50 feet tall [11]. Management practices combine both traditional and modern agricultural techniques. The soil in the orchard is calcareous with a silt loam texture, providing moderate water-holding capacity, but requires careful irrigation to maintain optimal moisture. The calcium deposits challenge nutrient availability, demanding due soil and nutrient management for tree health and productivity.

The region's subtropical climate, with hot summers and mild winters, supports mango growth but also poses challenges across phenological stages [40]. Cooler temperatures in February and early March aid bud formation. Flowering occurs as temperatures rise in March. April temperatures (up to 38 °C) promote fruit set, while May to July bring intense heat, frequently exceeding 45 °C. This accelerates fruit growth but increases the risk of heat stress. During this period, rainfall is minimal and hence irrigation is crucial. The monsoon season in July introduces rainfall spikes, sometimes exceeding 70 mm, beneficial but potentially leading to water-logging if excessive. By August, during the ripening stage,

there are intermittent rains. Cooler temperatures and drier conditions in October mark the post-harvest period and prepare the trees for dormancy.

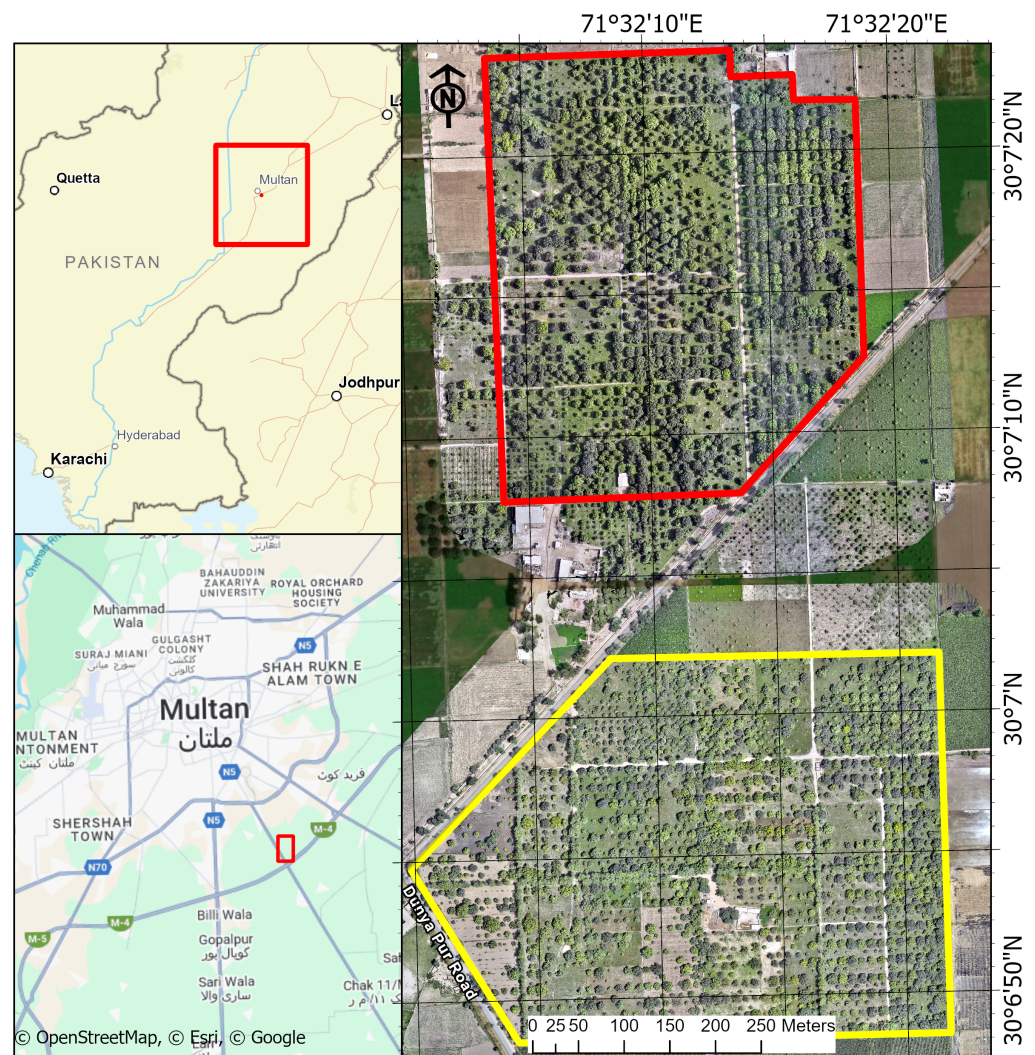


Figure 1. The study area is located in Multan, Punjab, Pakistan, as shown in the inset maps. The main experimental site, Orchard 1 (outlined in red), covers an area of 45 acres and contains 1305 trees. The validation site, Orchard 2 (outlined in yellow), spans 55 acres with 1833 trees. The location of both orchards is indicated in the high-resolution satellite image. The grid overlay provides geospatial reference points, with latitude and longitude markers.

2.2. Tree Health Assessment and Orchard Management Data

Extensive data were collected to support health classification and yield prediction models as shown in Table 2. Each tree was assigned a unique identification number, recording its variety and age. Based on visual inspections in February, trees were categorized into *healthy*, *moderate*, or *weak*. Yield ranges were tabulated according to variety, age, height, and health category based on the knowledge of local experts.

Table 2. Summary of the data types, collection and processing methods, and spatiotemporal resolutions for experimental study.

Data Type and Description	Collection and Processing Methods	Spatiotemporal Resolution
UAV and Related Data Unprocessed: Imagery with embedded DLS data (RGB 20 MP, 5 multi-spectral 5 MP), GCPs, reflectance panel data. Processed: Elevation (point cloud, DSM, DTM, CHM), RGB ortho-mosaics, multi-spectral reflectance mosaics processed into vegetation indices and zonal statistics.	UAV flights at 100 m AGL with 80% overlap. Processed using Pix4D, ArcGIS Pro, and Python scripts.	Flight schedule aligned with phenological stages.
Tree Attributes Variety, age, canopy size (m ²), height (m), health status.	Field measurements and inspections.	Annual and periodic health updates.
Orchard Management Irrigation, pruning, pesticide, fertilizer applications, and field observations.	Farm management records and visual inspections.	Post-activity and aligned with phenological stages. Control trees and orchard level.
Yield Data (kg/tree) Actual for control trees; yield estimates for remaining trees.	Physical measurements and estimate tables.	Tree-level for control trees. Orchard-level (total yield).
Weather Data Temperature (°C), precipitation (mm), wind speed (m/s).	Data sourced from Pakistan Meteorological Department.	Daily.

A randomized selection method was used for the selection of 200 trees in the main experimental orchard to represent a diverse range of varieties, ages, and health statuses. Their height and canopy volume were measured in March, and their health was monitored through visual inspection after each UAV flight for reference and model validation. Orchard management practices were documented, including dates and the amount of pruning, trenching, irrigation, pesticide, and fertilizer applications. Pruned trees exhibit altered reflectance responses, which change the CHM in-season, and recording this is an important data point [12]. This may potentially affect health and yield classification accuracy. From a nearby weather station managed by the Pakistan Meteorological Department, weather data (temperature, precipitation, and wind data) were obtained.

2.3. Observed Yield Data

The mango yield dataset used to prime and train the models is based on two components. For 200 control trees, yield (kg/tree) was physically measured post-harvest to provide ground truth data for model validation. For the remaining trees in the first orchard and the entire validation orchard, yield estimates were obtained using attribute-based estimation methods. Yield tables were prepared by experts based on historical local yield trends as observed by the Mango Research Institute (MRI) Multan and correlated with the tree attributes (variety, age, canopy size, and health status) [41–43]. Yield estimate for each variety was further categorized as per three tree health classes, i.e., healthy, moderate, and weak, with typical planting distances of about 40–45 feet in fertile soils and sound management practices. As per the local expert knowledge, healthy trees produce at full capacity, moderate trees around 60%, and weak trees about 30%. Yields tend to increase steadily from 10 to 20 years, growing at roughly 20% each year until 30–40 years. After 40 years, there is a decline in the yield which is rapid in the case of poor canopy management. Proper

canopy management helps in sustaining yields after 20 years, while its absence may cause a 25–50% drop past 40 years. CSB variety peaks at about 350 kg per tree in the 20–40-year range, Si and Langra reach around 300 kg, while SC achieves a yield of about 275 kg per tree. Even within the same age and health category, yields can vary by $\pm 25\%$ depending on canopy size, management practices, and the yearly asynchronous variations that are typical of local mango varieties [43]. Figure 2 shows these experts prepared estimates graphically. These estimations were then validated using the control tree dataset to ensure accuracy and were used to prime the models. As noted by other researchers, collecting accurate yield data is one of the most difficult tasks [35]. Thus, cumulative commercial tonnage sold from the orchards was also noted to provide an extra reference data point.

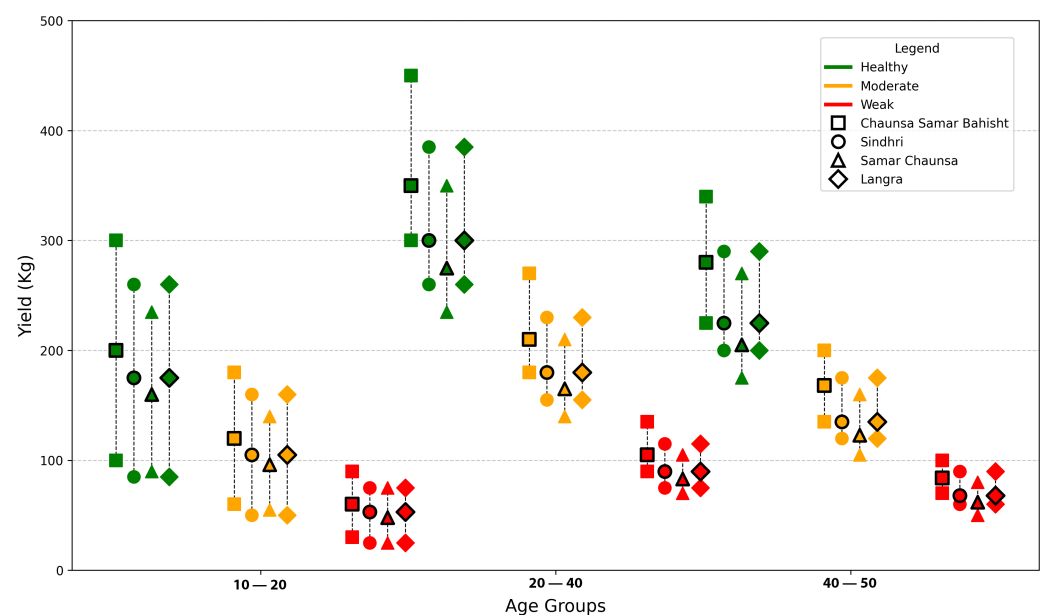


Figure 2. Mango yield estimates across different varieties, age groups, and health conditions, showing variations in productivity based on cultivar type and overall mango tree health.

2.4. Imagery Data Acquisition and Processing

A DJI Mavic 3M quadcopter was used for imaging. It is equipped with a multi-spectral camera with five spectral bands, i.e., Green ($560 \text{ nm} \pm 16 \text{ nm}$), Red ($650 \text{ nm} \pm 16 \text{ nm}$), Red Edge ($730 \text{ nm} \pm 16 \text{ nm}$), Near-Infrared (NIR; $840 \text{ nm} \pm 26 \text{ nm}$), and an RGB camera with 20-megapixel resolution [44]. Its Downwelling Light Sensor (DLS) was used for radiometric calibration in combination with the MicaSense Calibration Panel. A DJI D-RTK 2 Mobile Station was used for high-precision GNSS correction during flights and to record the ground control points for precise geo-referencing of the imagery across flights [45].

Best practices for imagery acquisition for agricultural applications as prescribed by the quadcopter manufacturer were adhered to [11,44]. Using DJI Pilot 2 software, flights were always conducted under clear sky conditions between 09:00 to 12:00 local time to maintain consistent lighting and reduce the effects of shadows. Flights were undertaken at altitudes of 100 m above ground level with 70% front and 80% side overlaps. A uniform flight pattern was used for all the flights. The flight schedule was aligned with six critical phenological stages of mango development for temporal analysis (Table 3).

Table 3. Flight dates and corresponding phenological stages of Mango crop.

Phenological Stage	Flight Dates	Description
Pre-Flowering	27 February 2024	Dormancy stage with minimal vegetative activity and early bud development.
Flowering and Full Bloom	10 March 2024, 24 March 2024	Emergence of panicles followed by full bloom and initial fruit set.
Fruit Set	21 April 2024	Formation of fruitlets.
Early Fruit Development	19 May 2024, 9 June 2024	Rapid fruit growth and increased reproductive activity.
Mid to Late Fruit Development	28 June 2024, 26 July 2024	Fruits approaching full size, with some showing signs of ripening.
Ripening and Post-Harvest	24 August 2024, 24 September 2024	Fruits ripening, harvest period, followed by transition into dormancy.

2.5. Data Processing

Imagery from UAV flights was processed using Pix4D Mapper 4.9.0. For this purpose, prescribed workflows in the software documentation were followed [46]. The Pix4D processing pipeline included tie-point matching, bundle block adjustment, point cloud generation, and ortho-rectification. A ground sampling distance (GSD) of 2 cm/pixel for RGB imagery and 5 cm/pixel for multi-spectral imagery was achieved. Photogrammetric point clouds, digital surface models (DSMs), digital terrain models (DTMs), band reflectance mosaics, and vegetation indices were derived. These were processed to calculate structural parameters such as tree height, canopy volume, and canopy, which are basic constituents of CHM and derive zonal indices from band vegetation indices using a combination of Pix4D Mapper 4.9.0, ArcGIS Pro 10.3, and Python 3.12 scripts [11].

3. MangiSpectra: Preliminary Formulations and Pre-Processing Modalities

MangiSpectra's data-driven pipeline is shown in Figure 3. The framework integrates UAV-based remote sensing data, carefully collected ground truth data, and feature engineering in each phenological stage to enable tree-level health classification and yield estimation. By leveraging spatiotemporal features for LSTM training, the output of the MangiSpectra includes predicted health classes and yield estimates of individual trees.

3.1. Spectral Analysis: Canopy Segmentation and Vegetation Indices

3.1.1. Canopy Detection and Segmentation

Accurate localization and segmentation of mango canopies are necessary for reliable vegetation index-based spectral analysis. For canopy detection and segmentation of mango tree canopies, this study builds upon the recent foundational work on the same orchards by the authors of [11]. Following the procedure described therein, a pre-trained You Only Look Once (YOLO) localized the tree canopies to bounding boxes, and the Segment Anything Model (SAM) precisely delineated the canopies. The resulting geo-tagged polygons were then used for the computation of the trees' physiological parameters, such as canopy area, height, and volume, as well as for the calculation of zonal statistics for various indices.

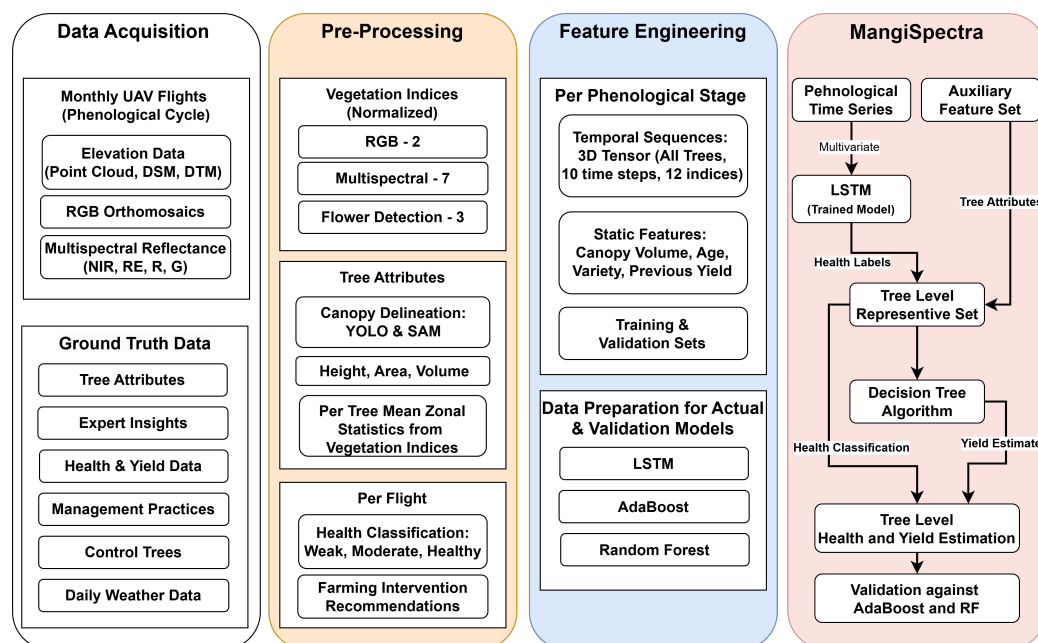


Figure 3. Overview of the four-staged integrated MangiSpectra framework for tree-level health and yield estimation.

3.1.2. Conventional Vegetation Indices

From each flight data, a total of nine vegetation indices were calculated. RGB-based indices included MCARI (Modified Chlorophyll Absorption Ratio Index) and VARI (Visible Atmospherically Resistant Index), while multi-spectral indices, namely RENDVI (Red Edge NDVI), NDVI (Normalized Difference Vegetation Index), NDRE (Normalized Difference Red Edge), GNDVI (Green NDVI), EVI (Enhanced Vegetation Index), SAVI (Soil-Adjusted Vegetation Index), and LAI (Leaf Area Index) were derived. Theoretical aspects, mathematical formulae, and practical applications of these indices in the context of precision agriculture in general and in the study of mango are already well documented and hence are not being reproduced [47].

These indices were normalized to the range of -1 to 1 using min-max normalization. The normalization process is achieved as follows:

$$v_{\text{normalized}} = 2 \cdot \frac{v - v_{\min}}{v_{\max} - v_{\min}} - 1 \quad (1)$$

where v represents the original value, $v_{\min}$ is the minimum value in the dataset, and $v_{\max}$ is the maximum value. The formulation in (1) ensures that all values are scaled proportionally within the range $[-1, 1]$.

Figure 4 shows a visual representation of 12 mango trees in RGB imagery and GNDVI on different dates. During data pre-processing, vegetation indices for tree canopies through zonal statistics were calculated for different phenological stages (Figure 5). A consistently increasing trend was observed from pre-flowering, reaching a peak during early fruit development followed by a decline in the values of these indices, which is generally aligned with the findings in similar studies on mango phenology [6,48,49]. Using the mean of normalized zonal statistics for various indices for every tree provided an insightful visual comparison during various phenological stages.

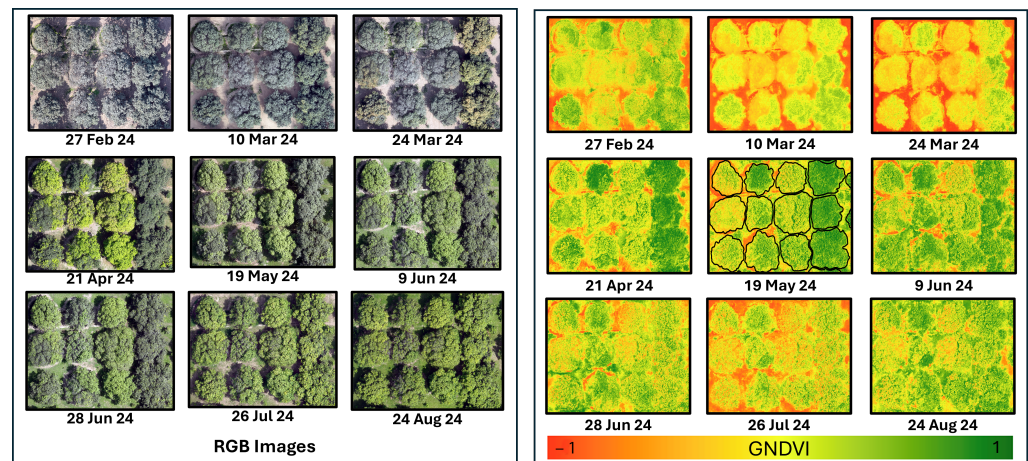


Figure 4. Unsegmented tree canopies of the same group of 12 mango trees over different flight dates in Orchard 1 showing the effect of underlying vegetation. (Left): RGB image; (Right): Normalized GNDVI.

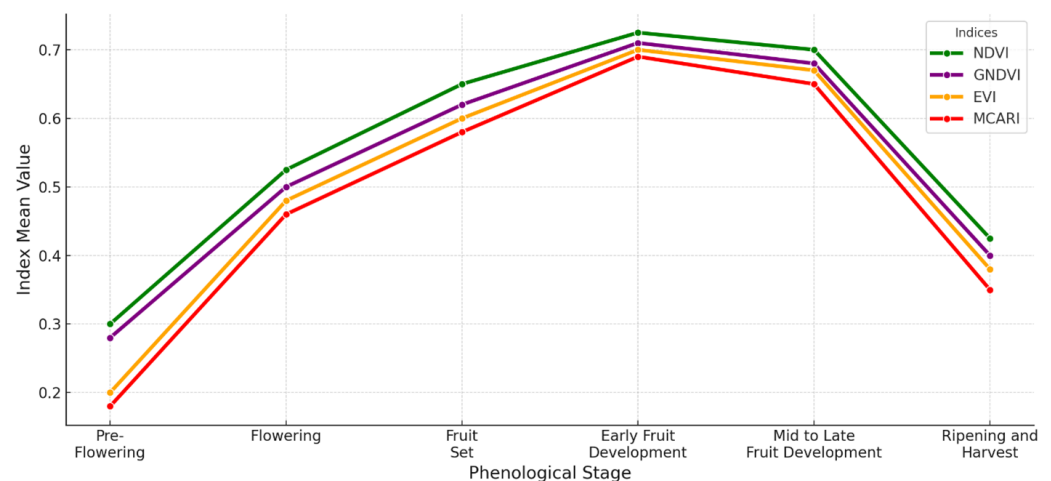


Figure 5. Cumulative trend of key vegetation indices across phenological stages at Orchard 1.

3.2. Development of Mango Flowering Spectral Indices

As mango trees emerge from dormancy, the emergence of new vegetative growth and flower panicles signals the tree's vitality and reproductive readiness. Hence, monitoring mango flowering is pivotal for health and yield estimation. The distinctive cream-yellow hue of panicles affords an interesting opportunity to measure the degree to which flowers cover tree canopy [39,40]. As normal vegetation indices lack the required sensitivity to measure flowering intensity, there have been several studies to develop flowering-specific indices. Faster R-CNN-based algorithms have been utilized to detect white flower clusters in the case of apple orchards [37]. To detect yellow flowers, the utility of canola-specific indices has been demonstrated [50,51]. For the mango itself, ground-based RGB imaging for the measurement of mango flowering intensity has been utilized [38]. Leaf damage due to changes in leaf color caused by thrips in mango orchards has been measured using a modified Maximum-Likelihood Classifier (MLC) and multi-spectral drone imagery [14]. However, there are not many studies specifically focused on measuring the degree of mango tree canopy covered by flowers during the flowering stage.

To improve the robustness of MangiSpectra by addressing this gap, three novel indices have been developed, i.e., the Mango Tree Yellowness Index (MTYI), the Weighted Yellowness Index (WYI), and the Normalized Adjusted Flowering Detection Index (NAFDI), to capture the unique spectral characteristics of mango canopies during flowering and leveraging the high spatiotemporal resolution afforded by the UAV imagery.

3.2.1. MTYI

It is a multi-spectral normalized difference index developed to capture the cream/pale yellowish hue of mango flowers during the flowering season. MTYI emphasizes the Green (G) and Red (R) channels while reducing the impact of the Red Edge (RE) and Near-Infrared (NIR) bands. The MTYI is calculated as follows:

$$MTYI = \frac{(G + R) - (RE + NIR)}{(G + R) + (RE + NIR)} \quad (2)$$

where:

- G = Reflectance in the G band (helps in distinguishing yellowish hues of flowers from the surrounding green foliage);
- R = Reflectance in the R band (sensitive to chlorophyll absorption of the green foliage);
- RE = Reflectance in the RE band (indicative of plant stress and leaf area);
- NIR = Reflectance in the NIR band (correlated with biomass and leaf structure).

The band grouping in MTYI is performed to leverage the reflectance properties of mango flowers and leaves during anthesis. Mango leaves strongly absorb red light because of their chlorophyll content while also moderately reflecting green light [14]. The combination of G and R bands ensures that the floral spectral response is less affected by chlorophyll absorption and enhances the discrimination of flowers from leaves [39]. Healthy green leaves show strong reflectance in the RE and NIR bands due to leaf structure, and hence grouping these effectively isolates the leaves. Subtracting this from the $(G + R)$ combination minimizes the influence of dense leaf cover, thus enhancing flower detection. Being a normalized index, MTYI's output remains consistent under varying lighting conditions, shadows, or sensor calibration differences. It enhances comparability and aids analysis by providing clear thresholds for strong floral presence. High MTYI values indicate dense flowering, while low or negative values correspond to regions with minimal flowering.

3.2.2. WYI

WYI is a weighted multi-spectral index developed to enhance flowering detection by emphasizing flowering-related spectral bands (G and R) while reducing the contributions of NIR and RE , which are more associated with vegetation and chlorophyll. The WYI is computed as follows:

$$WYI = \frac{0.5 \cdot G + 0.3 \cdot R - 0.1 \cdot NIR - 0.1 \cdot RE}{0.5 \cdot G + 0.3 \cdot R + 0.1 \cdot NIR + 0.1 \cdot RE} \quad (3)$$

While the WYI follows the same spectral logic as MTYI, it uses custom weighting factors for each band to further enhance flower detection. Unlike MTYI, it assigns distinct emphasis to each band based on empirical findings of its contribution towards mango flower and foliage reflectance. Higher weights assigned to G (0.5) and R (0.3) emphasize the mango flowering response, while lower and negative weights assigned to NIR (−0.1) and RE (−0.1) suppress leaf and chlorophyll reflectance. This weighted approach was validated through preliminary field measurements. The coefficients were subsequently chosen to maximize the contrast between flowering canopies and surrounding vegetation.

3.2.3. NAFDI

NAFDI is an RGB-based index that utilizes high-resolution RGB imagery from the Mavic 3M to detect the spectral hue of mango flowers. The index is computed as follows:

$$\text{NAFDI} = \frac{\left(\frac{R}{180}\right) \times \left(\frac{G}{150}\right) \times \left(1 - \frac{Blue}{255}\right)}{\left(\frac{R}{180}\right) \times \left(\frac{G}{150}\right) \times \left(1 - \frac{B}{255}\right) + 1} \quad (4)$$

where the constants 180, 150, and 255 normalize the respective RGB components to balance their influence in the computation. These were empirically determined to normalize the dynamic range based on observed flowering intensity in calibration datasets.

NAFDI utilizes the higher sensitivity of the *R* channel [0,255] range to carotenoid pigments in the mango flowers coupled with the *G* to detect flowers. The blue channel (*B*) is used to reduce interference from non-flowering elements. The constants, which are empirically derived normalizing factors, ensure that each RGB channel contributes effectively to the NAFDI to emphasize floral reflectance while minimizing variability arising from lighting and background vegetation. This approach was validated through preliminary field trials under typical orchard conditions. Higher NAFDI values indicate regions with intense flowering, whereas values approaching 0 signify non-flowering areas.

Figure 6 demonstrates the temporal dynamics of mango flowering as detected by the three indices on three dates (9 March 2024, 24 March 2024, and 21 April 2024) using the above indices. The RGB images (first row) provide a visual representation of canopy changes, while the index maps (subsequent rows) quantify the distribution of floral abundance. The flowering is sparse on the 9th of March being the early season, whereas by the 24th of March, the yellowness increases, and peak floral bloom is reached. By the 21st of April, flowering declines. As tree canopies have been precisely segmented, mean values of these indices for each tree on three flight dates were calculated using the zonal statistics feature of ArcGIS and then exported as a CSV file for use in LSTM.

3.2.4. Per-Flight Cumulative Health Index (CHI) for In-Season Recommendations

One of the unique features of UAVs is their ability to afford very high frequency of data during a phenological cycle. The availability of vegetation indices soon after every flight presented an excellent opportunity to undertake per-flight health classification of trees for in-season intervention recommendations to the orchard managers. Thus, a Cumulative Health Index (CHI) was developed for stage-specific health assessments using weighted combinations of normalized vegetation indices. These weights were assigned based on the earlier research and expert knowledge as per each phenological stage [6,52,53]. Inferences were shared with orchard managers and helped them to focus on *weak* trees for prompt remedial actions.

For each phenological stage, the stage-specific CHI is computed as a weighted sum of normalized indices. The CHI is calculated as follows:

$$\text{CHI}_{\text{Stage}} = \frac{\sum_{i=1}^n w_i \cdot I_i}{\sum_{i=1}^n w_i} \quad (5)$$

where:

- w_i = Weight assigned to the i th index, reflecting its relative importance in the given phenological stage.
- I_i = Normalized value of the i th index, scaled within the range of -1 to 1 .
- n = Total number of spectral indices considered in the computation.

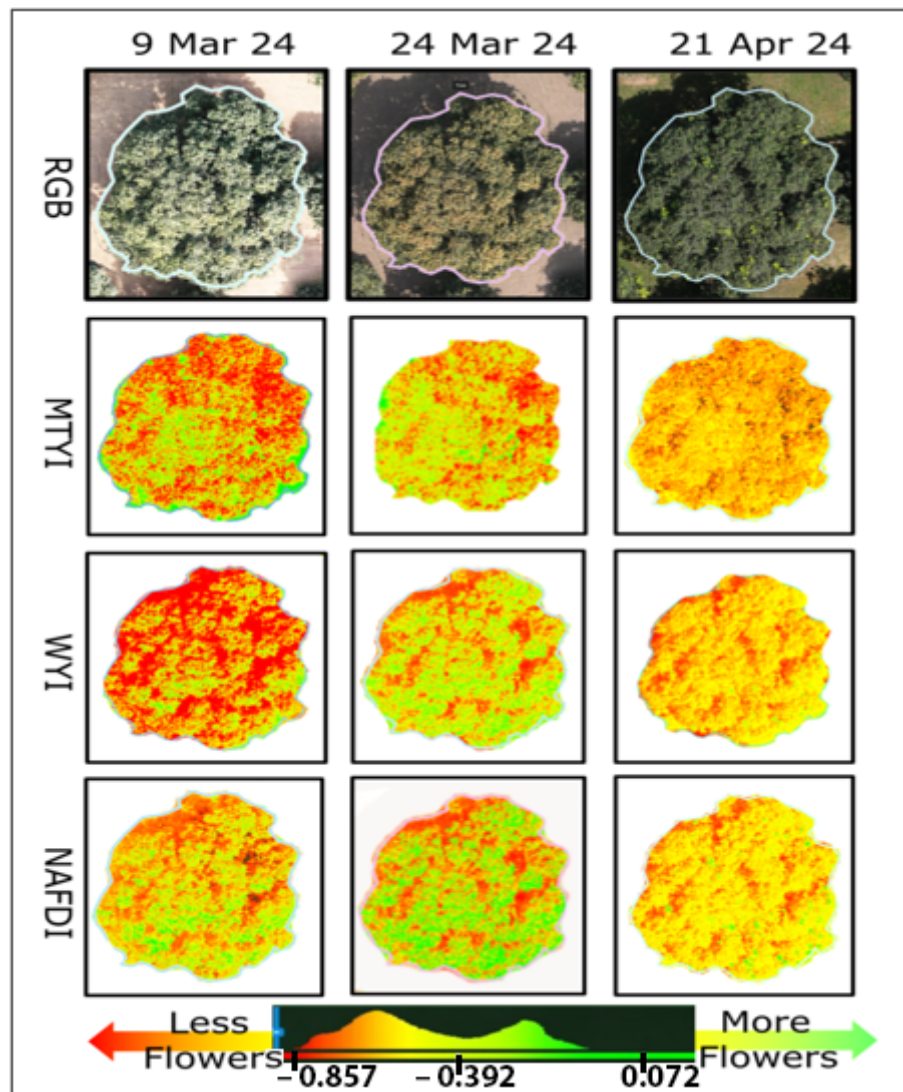


Figure 6. Progression of flowering from March to April 2024 as detected by Mango Tree Yellowness Index (MTYI), Weighted Yellowness Index (WYI), and Normalized Automatic Flowering Detection Index (NAFDI) on the canopy of the same tree. RGB UAV imagery in the top row is for visual reference. The color scale represents the relative density of flowers, with red indicating a lower degree of flowering and green indicating higher.

A sample outcome of the health classification based on CHI from the 24 March flight is presented in Figure 7. This example illustrates the application of the methodology, demonstrating how CHI-based classification maps were generated and subsequently handed over to orchard managers for on-ground verification and intervention. This step was crucial in validating the approach before its broader application in the study.

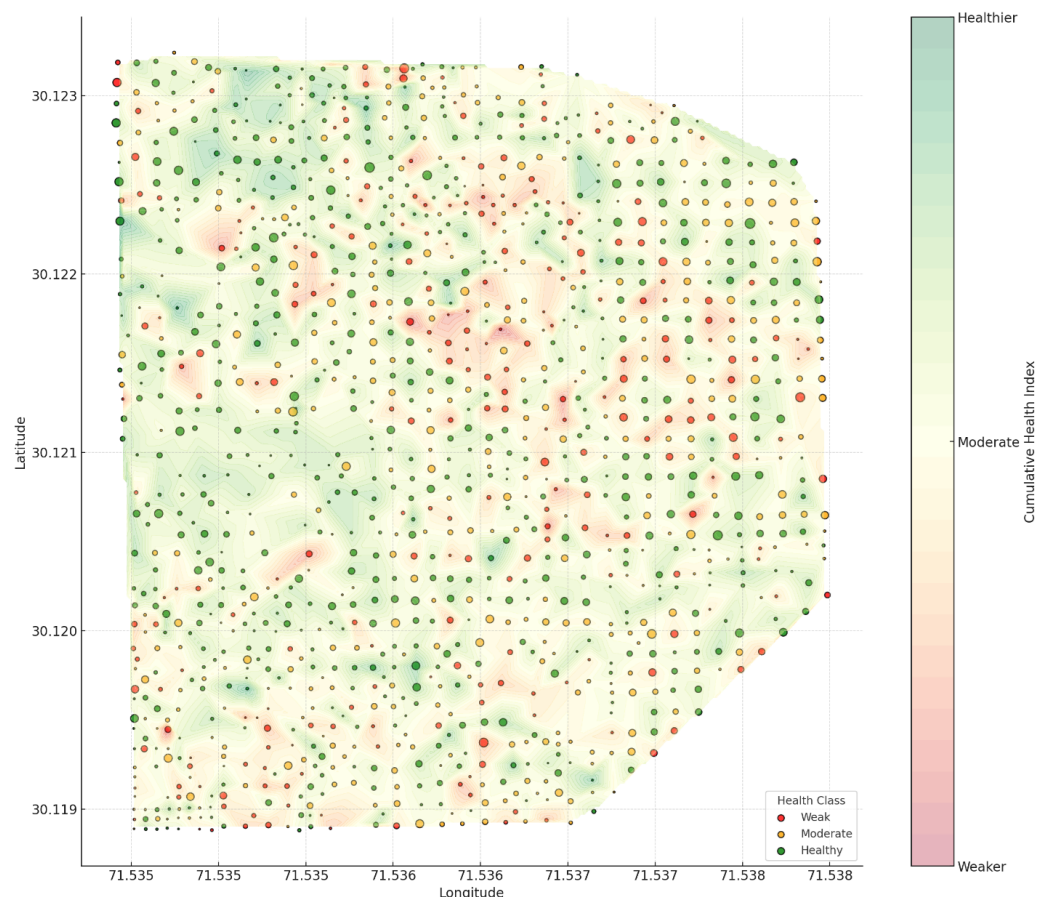


Figure 7. Sample of per-flight health classification and in-season farming intervention recommendations. The map shows the health of categories of trees in Orchard 1 on 24 March 24 during the flowering stage.

4. MangiSpectra: LSTM-Based Health Classification and Yield Estimation

4.1. Health Prediction: Multivariate Time Series Classification

MangiSpectra employs multivariate time-series data for each tree in the study area based on vegetation indices covering discrete temporal intervals to classify tree health during the phenological cycle.

4.1.1. Data Structure for Multivariate Time Series

Each tree in the orchard is represented by a multivariate time series of vegetation indices collected over N UAV flights during the phenological cycle. For X trees and J vegetation indices per flight, the dataset T is structured as a 3D tensor of shape (X, N, J) , where each element $P_x(n, j)$ represents the j -th vegetation index at the n -th time step for tree x :

$$T = \{P_x(n, j) \mid x = 1, \dots, X; n = 1, \dots, N; j = 1, \dots, J\}$$

Each tree is assigned a health label from three classes: *weak* (0), *moderate* (1), and *healthy* (2). The complete label set is represented as $Y = \{y_x \mid x = 1, 2, \dots, X\}$, where $y_x \in \{0, 1, 2\}$.

4.1.2. LSTM Architecture for Health Classification and Yield Estimation

Figure 8 abstracts the working of the LSTM component within the MangiSpectra framework. For input to the LSTM, each tree's time series tensor T_x is structured as (N, J) , where $N = 9$ represents the number of UAV flights (time steps), and $J = 12$ denotes the

number of vegetation indices. The LSTM processes each tree's time series sequentially to capture temporal dependencies within the phenological data.

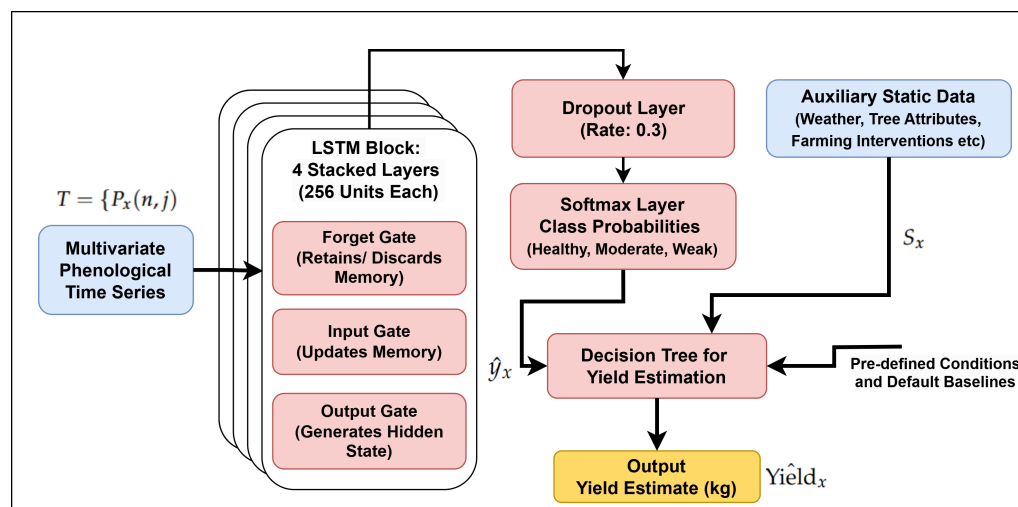


Figure 8. Utilization of LSTM component within the MangiSpectra framework for health classification and yield estimation.

The hidden and cell states are initialized as zero and updated at each time step n using standard LSTM operations for forget, input, and output gates and candidate cell states. These operations enable the model to selectively retain or discard information across time steps. After processing all N time steps, the final hidden state $h_x(N)$ is passed through a softmax layer to predict the health class probabilities for tree x :

$$\hat{y}_x = \text{Softmax}(Wh_x(N) + b)$$

The model is trained using the cross-entropy loss function:

$$L = - \sum_{x=1}^X \sum_{c=0}^2 y_x(c) \log(\hat{y}_x(c))$$

where $y_x(c)$ is the true label for class c , and $\hat{y}_x(c)$ is the predicted probability for tree x .

4.1.3. LSTM Training for MangiSpectra

The LSTM network is implemented in PyTorch 2.6 and comprises four stacked LSTM layers, each with 256 cells, followed by dropout layers to prevent over-fitting and a final softmax output layer. Each tree's time series tensor T_x , structured as (N, J) , is processed sequentially by the LSTM to capture temporal dependencies across the phenological stages.

To address the class imbalance, inverse frequency weighting is applied to the labels during training. The model's hyperparameters were tuned through grid search final parameters and are summarized in Table 4. The dataset was split into 80% training and 20% testing, ensuring an even distribution across health categories. To evaluate the model's generalization ability, a cross-validation approach was used.

The training was performed as per Algorithm 1, and the model's performance was evaluated at each epoch using accuracy, precision, recall, and F1 score metrics. Validation and testing are conducted with fixed random seeds to ensure reproducibility.

Algorithm 1: MangiSpectra Training Cycle for LSTM Model

Input: Dataset T with shape (X, N, J) , Labels Y
Output: Trained LSTM model

```

1 Initialize parameters, optimizer, and random seed;
2 for  $epoch = 1$  to  $num\_epochs$  do
3   Shuffle  $T$ , split into batches;
4   foreach batch  $B_i$  do
5     Process each tree  $x$  sequentially, updating hidden and cell states via LSTM;
6     Compute cross-entropy loss  $L$ ;
7     Back-propagate loss and update parameters;
8   end
9 end

```

Table 4. Hyper-parameters for LSTM training in MangiSpectra.

Hyper-Parameter	Value
Dropout Rate	0.3
Number of Epochs	2000
Learning Rate	10^{-6}
Batch Size	1024
Optimizer	Adagrad
LSTM Layers	4
Cells per Layer	256

4.2. Yield Estimation: Leveraging Auxiliary Static Data with Decision Tree Approach

The predicted health classes $\hat{y}_x$ are augmented with auxiliary features S_x for each tree x , including static data such as weather:

$$S_x = [\text{Weather}, \text{Age}_x, \text{Variety}_x, \text{CanopyVolume}_x, \text{TreeHeight}_x]$$

The complete input for each tree x is represented as $(T_x, S_x, \hat{y}_x)$, where T_x is the time-series tensor, and S_x is concatenated with the LSTM's output $\hat{y}_x$ to estimate yield using a conditional decision tree approach incorporating prior knowledge about how factors such as variety, age, and health status influence yield [54].

Algorithm 2 outlines the generalized process for yield estimation based on S_x . This pseudo-algorithm mimics the nested conditional structure used in the implementation. The whole process ensures adaptability across varying tree characteristics and orchard conditions.

4.3. Evaluation Metrics

To comprehensively evaluate the performance of both health classification and yield estimation, several metrics are calculated based on the model's predictions. The primary evaluation metrics for health classification are accuracy, precision, recall, F1 score, and confusion matrix, while for yield estimation, mean absolute error (MAE), mean squared error (MSE), and R^2 are utilized to assess prediction accuracy.

4.3.1. Health Classification Metrics

Accuracy measures the proportion of correct classifications:

$$\text{Accuracy} = \frac{1}{X} \sum_{x=1}^X \mathbf{1}\{\arg \max(\hat{y}_x) = y_x\} \quad (6)$$

where $\mathbf{1}\{\cdot\}$ is the indicator function.

Precision, recall, F1 score, and a confusion matrix were used to evaluate the performance for each health class c .

Algorithm 2: Yield Estimation for Mango Trees

Input: Input (T_x, S_x) , where
 $S_x = [\text{Weather}, \text{Age}_x, \text{Variety}_x, \text{CanopyVolume}_x, \text{TreeHeight}_x]$
Output: Estimated yield Yield_x

```

1 foreach tree  $x$  do
2   Extract  $\hat{y}_x, S_x$ ;
3   if  $\text{Variety}_x$  matches predefined conditions then
4     if  $\text{Weather}, \text{Age}_x, \text{CanopyVolume}_x, \text{TreeHeight}_x$  fall within specific ranges then
5       Estimate yield based on  $\hat{y}_x$  and matching attribute ranges;
6     end
7   else
8     Assign yield based on default baselines;
9   end
10 end
11 else
12   Use general yield estimation for unmatched varieties;
13 end
14 Save  $\text{Yield}_x$ ;
15 end
  
```

4.3.2. Yield Estimation Metrics

To evaluate the performance of yield predictions, error-based metrics quantify the deviation between predicted yield $\hat{\text{Yield}}_x$ and actual yield Yield_x :

$$\text{MAE} = \frac{1}{X} \sum_{x=1}^X |\hat{\text{Yield}}_x - \text{Yield}_x| \quad (7)$$

$$\text{MSE} = \frac{1}{X} \sum_{x=1}^X (\hat{\text{Yield}}_x - \text{Yield}_x)^2 \quad (8)$$

The R^2 metric measures the proportion of variance in actual yield explained by the predictions:

$$R^2 = 1 - \frac{\sum_{x=1}^X (\text{Yield}_x - \hat{\text{Yield}}_x)^2}{\sum_{x=1}^X (\text{Yield}_x - \overline{\text{Yield}})^2} \quad (9)$$

where $\overline{\text{Yield}}$ is the mean of the actual yield values.

5. Results and Discussion

5.1. Health Classification

5.1.1. LSTM Model Training and Accuracy

The LSTM model was trained on a tri-class classification task (*healthy, moderate, weak*) using the multivariate phenological time series as the input. This input is a combination of standard indices and the three newly developed yellowness indices. Figure 9 shows the key performance metrics across the training and testing phases.

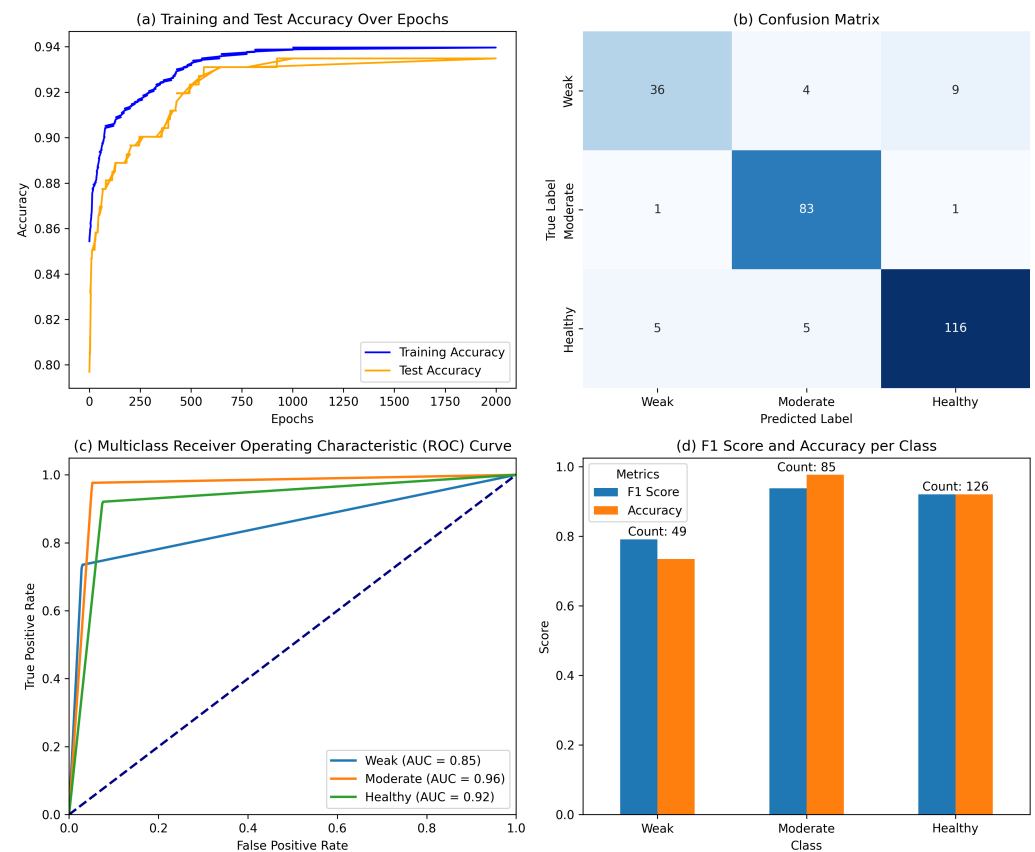


Figure 9. Key performance metrics of the LSTM model for tree health classification: (a) training and test accuracy over epochs, (b) confusion matrix, (c) ROC curves for each class, and (d) F1 score, accuracy, and class distribution.

The accuracy curves (Figure 9a) indicate that the model convergence is around 90% training accuracy, with test accuracy stabilizing around 93%, beginning at approximately epoch 2000. The gap between training and test accuracy indicates that the model generalizes well to unseen data, with minimal over-fitting.

The majority of predictions fall along the diagonal (Figure 9b), showing good classification performance of the model. The model correctly predicts 36 *weak* samples, 83 *moderate* samples, and 116 *healthy* samples. However, there are some classification errors as well. Nine *weak* samples are wrongly classified as *healthy* and four as *moderate*. Similarly, five *healthy* samples are misclassified as *weak* and five as *moderate*. Overall, the *moderate* class shows the best outcome, with only a few classification errors.

The model is quite effective at differentiating between the classes (Figure 9c). An AUC of 0.85 for the *weak* class indicates an average performance. Here, too, the *moderate* class reflects excellent separability with an AUC of 0.96. The *healthy* class also performs well, with an AUC of 0.92.

The model performs well for the classes with large samples (Figure 9d). An F1 score and accuracy of approximately 0.85 for the *weak* class is slightly lower than the other classes. This is likely due to a smaller sample size (49 instances), a fact which is attributed to the reason that mango is a cash crop. Moreover, with the help of CHI, orchard managers undertook a very prompt intervention against any indication of plant stress. Thus, the sample size is apt to remain small. For its 85 instances, the *moderate* class achieves an F1 score and accuracy near 0.9. The *healthy* class too demonstrates an equally good performance for both metrics. Overall, the LSTM model effectively leverages the phenological time-series combination of vegetation and flowering indices.

5.1.2. Analytical and Experimental Validation

The health analysis of trees in the orchard using the proposed MangiSpectra framework is presented in Figure 10. A majority of trees are classified as moderately healthy, particularly in the age range of 30 to 40 years (Figure 10a). *Weak* and *healthy* classes are less common but spread across various age ranges. A slight concentration of *weak* trees in the younger (5–15 years) and older (40–45 years) age groups was observed. This pattern suggests that middle-aged trees (20–40 years) tend to be in better health, while younger and older trees are more prone to health challenges.

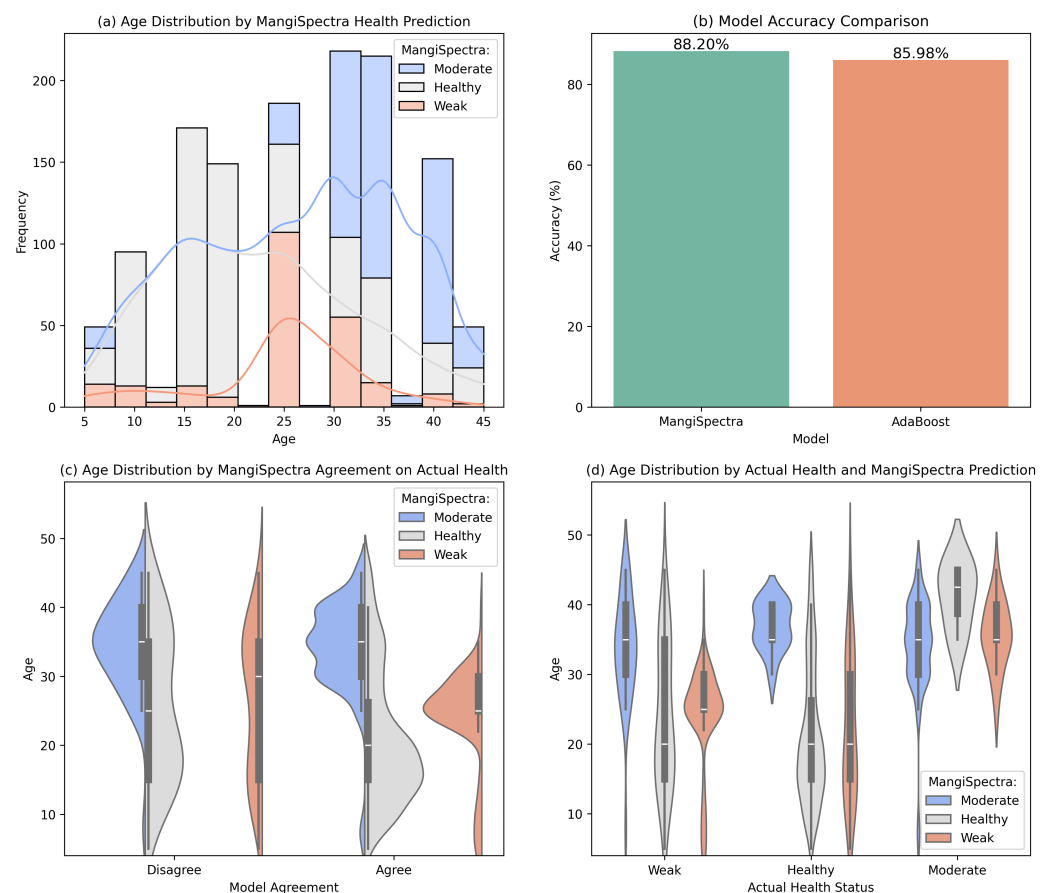


Figure 10. Analysis of tree health in the orchard using the MangiSpectra framework: (a) age distribution by tree health status, (b) model accuracy comparison, (c) age distribution by model agreement on tree health status, (d) age distribution by tree health status and MangiSpectra prediction.

The health classification component of the framework was validated against the AdaBoost classifier (Figure 10b). Both methods performed effectively, with MangiSpectra achieving a slightly higher accuracy of approximately 93% compared to AdaBoost’s 91%. MangiSpectra’s marginally better performance presents it as a reliable alternative for tree health assessments.

Both models also tend to agree on health status predictions (Figure 10c). The age distribution shows a concentrated range, especially for moderately healthy trees in the age group of 20–35 years. However, disagreements between models are more frequent among younger (5–15 years) and older (40–45 years) trees. This indicates limitations in the models’ predictive consistency for these age groups. Additional training data or model adjustments may be needed to improve prediction consistency across all age ranges and make the framework more generalizable.

Figure 10d shows that MangiSpectra predictions for moderately healthy trees are fairly consistent across all ages, indicating no significant age bias. For *weak* and *healthy* trees, the predictions show more variance, particularly among younger and older trees. This suggests that MangiSpectra may struggle to predict extreme health conditions consistently for trees at the age extremes.

5.1.3. Spatial Distribution of Classification Clusters

The spatial distribution of tree health zones is shown in Figure 11. The visualization highlights variations in tree health across the orchard. The size of the points is normalized as per the canopy volume of each tree. The map shows classification clusters helpful for orchard managers to identify areas requiring interventions. These interventions may include pest control, soil improvement, or nutrient application in risk zones.

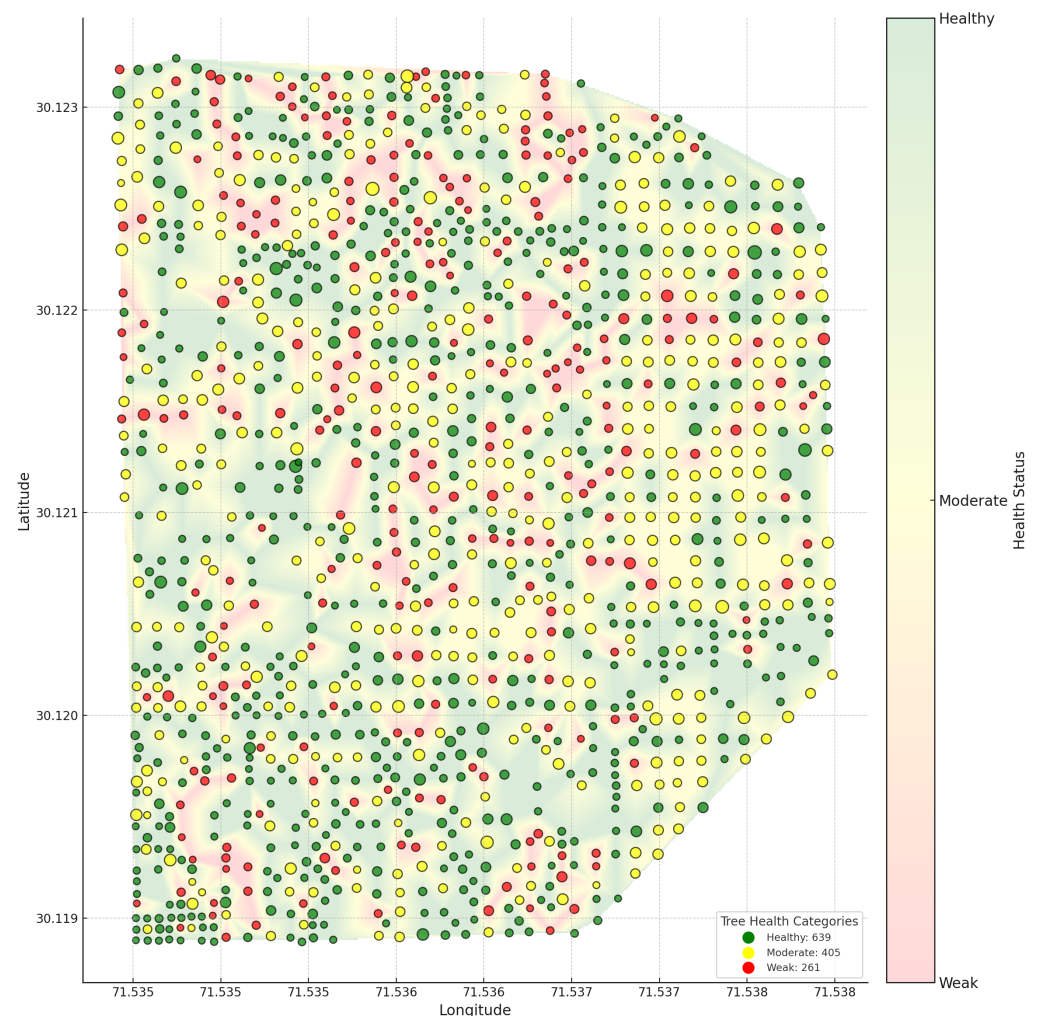


Figure 11. Spatial distribution of tree health as estimated by MangiSpectra for Orchard 1. Each dot represents an individual tree, categorized into three health classes: healthy (green, 639 trees), moderate (yellow, 405 trees), and weak (red, 261 trees). The background heat map provides an interpolated health estimate, highlighting areas of varying tree conditions with greenish arcs indicating prevalence of healthier trees, yellowish indicating moderate, and reddish areas showing clustering of weak trees.

5.2. Yield Estimation

5.2.1. Experimental Validation

The decision tree component of the MangiSpectra framework was compared with the Random Forest model. The choice was considered logical as Random Forest is a meta-estimator that fits various decision tree regressors over snapshot samples of data.

Figure 12 illustrates the relationship between actual yields and the predicted values from the original and validation models. The MangiSpectra shows an $R^2 = 0.21$, indicating a modest correlation with actual yields (Figure 12a). An MAE = 36.99, and RMSE = 50.18 suggest a generally reasonable estimation error. The Random Forest model, with an $R^2 = 0.18$, MAE = 31.93, and an RMSE = 50.88 indicates (Figure 12b) that both models capture variability in yields with comparable accuracy.

The comparison also shows that the models produce close predictions (Figure 12d). Slight variations may indicate different model sensitivities to underlying factors influencing yield.

Lastly, the scatter of yield versus plant age shows no pronounced trend (Figure 12c). To gain better insight into this, additional data points may be required to establish age-related yield-influencing factors to improve predictive accuracy.

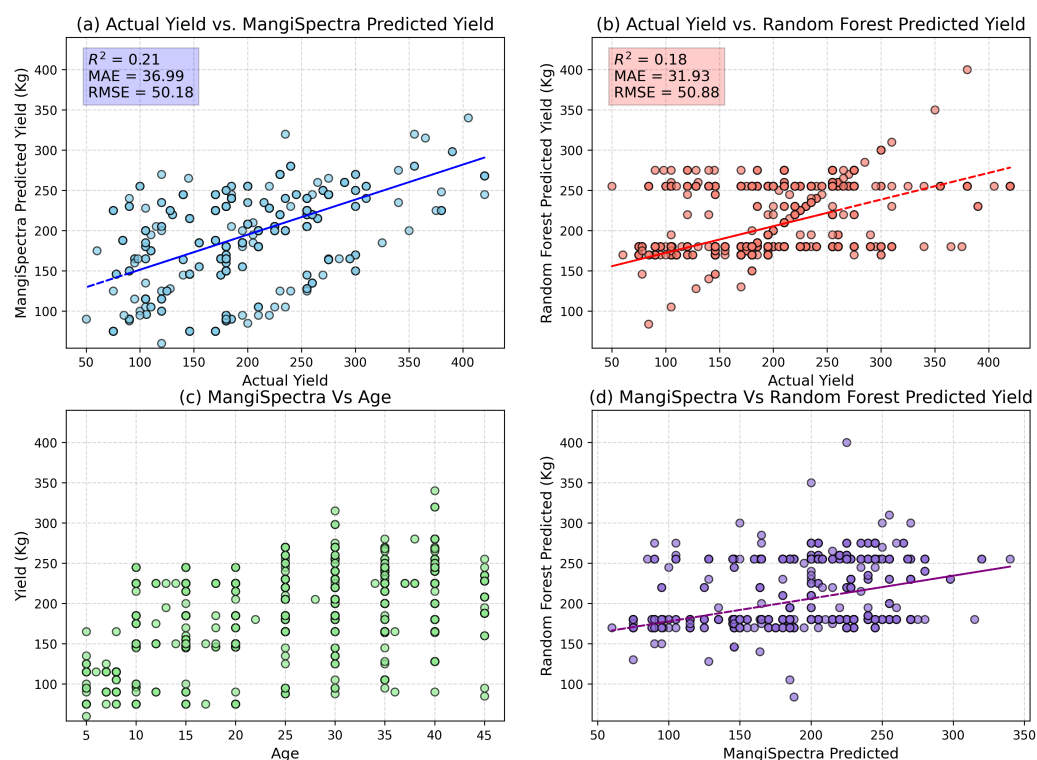


Figure 12. Comparison of actual yield with model predictions: (a) actual yield compared with MangiSpectra estimate, (b) actual yield compared with Random Forest estimate, (c) estimated yield of MangiSpectra correlated with age, and (d) predicted yields of MangiSpectra and Random Forest.

5.2.2. Spatial Distribution of Yield Estimates

Figure 13 illustrates the spatial distribution of yield estimates across Orchard 1 demonstrating the ability to provide orchard-scale yield trends. The dominantly green areas indicate optimal yield performance zones. Ground visits confirmed that these zones had favorable growing conditions such as quality of soil and adequate water availability. Conversely, yellowish-red clusters indicate suboptimal yield zones linked to localized issues like pest infestations, poor soil health, insufficient irrigation, and suboptimal nutrient management.

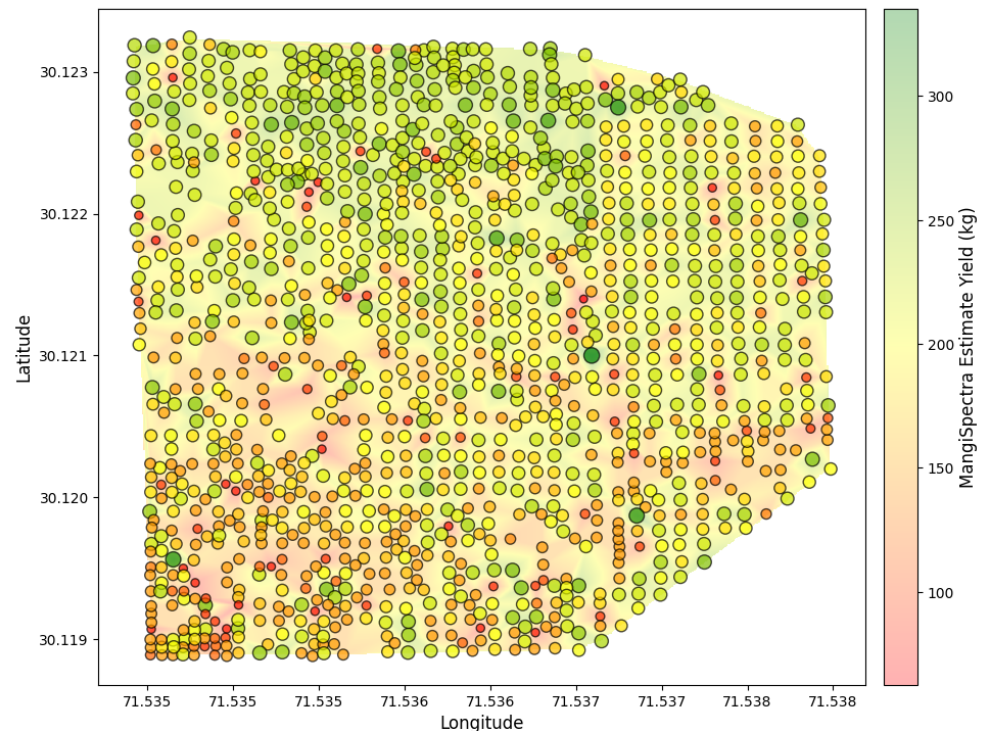


Figure 13. Spatial distribution of yield estimates in Orchard 1. Individual trees are represented by circles. Circle sizes correspond to the normalized yield, while colors indicate yield estimation: low (red), moderate (yellow), and healthy (green). The gradient background depicts yield estimate zones from low (red) to high (green).

5.3. Application of MangiSpectra on Validation Site Orchard

MangiSpectra is finally applied on the validation orchard for health assessment as shown in Figure 14. A pre-trained LSTM classifier on the original experimental orchard was used for this purpose with input time-series related to the validation orchard. It can be seen that the majority of trees in the central and northern regions of the orchard exhibit better health scores, a fact that was confirmed on the ground. On the other hand, areas towards the southern and southwestern edges demonstrated a noticeable decline in health. On-ground inspection revealed that there is a higher density of *weak* trees in the area.

5.4. Addressing Limitations of Traditional Remote Sensing Methods

UAV imagery offers significant advantages in terms of spatial and temporal resolution, but it is still not immune to canopy occlusion and limited canopy penetration.

MangiSpectra employs several key strategies to address these limitations. First and foremost, it utilizes high-resolution UAV imagery combined with object detection and segmentation to reduce spectral mixing. Secondly, the LSTM-based two-stage approach captures the temporality of indices over multiple phenological stages. This mitigates the impact of occlusion at a single time point. Thirdly, the indices are specifically designed to be sensitive to flowering intensity. This ultimately allows for a more accurate health and yield assessment compared to chlorophyll-based conventional indices.

The performance of MangiSpectra is validated against traditional machine learning techniques. In health classification, it achieved a slightly higher accuracy (93%) compared to AdaBoost (91%). For yield estimation, it performed comparably to the Random Forest model ($R^2 = 0.21$ and $R^2 = 0.18$, respectively). The framework's novelty lies in the fact that previous research has not fully leveraged UAV-derived spectral indices or phenological data for orchard health mapping and yield estimation in mango crops.

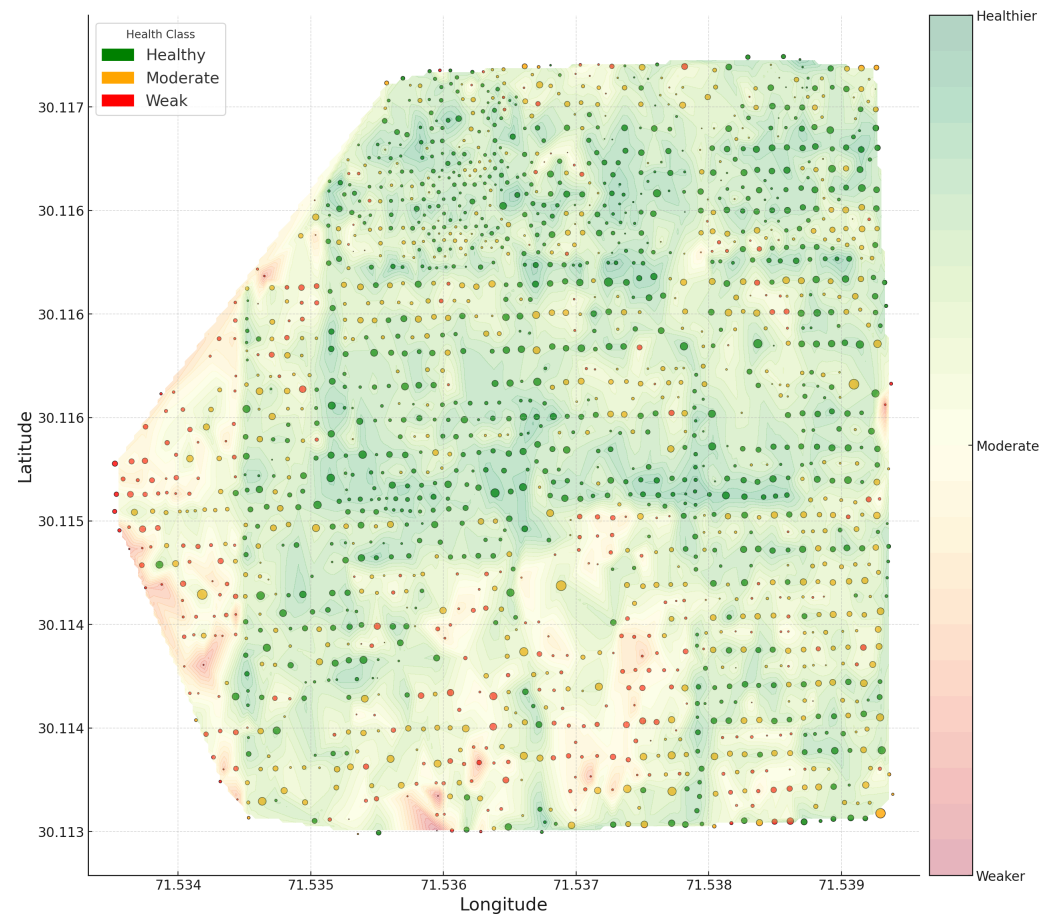


Figure 14. Spatial distribution of health classification over Orchard 2.

6. Conclusions

This study introduced a practical and effective framework for near-real-time health and yield monitoring of mango orchards, as well as end-of-season cumulative health and yield estimation. UAV-based multi-spectral and RGB imagery are utilized with object detection and segmentation deep learning models to build a rich dataset containing multivariate phenological time series, environmental constraints, expert farm management variables, and tree canopy properties. An LSTM classifier is trained on a temporal subset to classify tree-level health. The health classification labels are subsequently concatenated with the static spatial and auxiliary data. This is fed to a decision tree with pre-set baseline conditions. The decision tree fits a regressor to estimate tree-level yield. The health classifier and yield estimator are validated against the Adaboost and Random Forest algorithms, respectively.

By capitalizing on LSTM's capabilities, MangiSpectra accurately captures complex temporal patterns in tree growth and environmental interactions, resulting in reliable health classification and yield predictions. The results of this framework were compared and validated against AdaBoost and Random Forest, which have proven applications for health classification and yield estimation [19,54]. MangiSpectra therefore has the potential to be a robust tool for near-real-time mango orchard management and yield optimization. It can be modified and extended to other perennial fruit crops such as avocado after crop-specific retraining.

7. Future Work

Future work could explore the potential of expanding the dataset to cover a wider range of orchard conditions and incorporating additional environmental factors. Moreover,

the algorithms used may be made more computationally efficient and fine-tuned. In addition, the potential of attention-based classifiers can be explored with more advanced yield estimators. Low-cost LiDAR and hyper-spectral cameras equipped with UAVs can further enhance the accuracy of research, especially for variety-based insights during the phenological cycle and improving the CHM. An extension of the framework through a smartphone and web-based application can be evaluated to ensure that orchard managers in the field can easily benefit. Integrating output from additional sensors, such as IoT-based soil sensors and handheld spectrometers and radiometers, can improve the more accurate recording of ground data, especially during vegetative growth, flowering, measuring distress, and harvesting.

It is also important to note that variability in orchard management practices, such as selective pruning, impacts canopy reflectance properties. It is recommended that records of pruning related to date, amount, type, and tree be kept and included. Future studies may account for pruning history as an additional variable in the predictive models.

Furthermore, UAV-based data can be further integrated with high-resolution satellite imagery for spectral response comparison. This will enhance scalability and long-term monitoring. A broader spatial coverage through satellite is also possible with fine-grained insights provided by UAVs.

Data collection from the field for MangiSpectra and the dissemination of inferences to the orchard manager could benefit from a smartphone-based application, thus enhancing the practical application of the framework.

With these advancements, MangiSpectra has the potential to evolve into a more adaptable and reliable tool, empowering researchers and farmers with actionable insights to boost orchard productivity and sustainability.

Author Contributions: Conceptualization, M.M.A.; Data Curation, M.S.I.; Investigation, M.M.A.; Methodology, M.M.A., M.S.I. and A.D.B.; Project administration, E.H. and J.I.; Resources, E.H. and J.I.; Software, M.M.A. and M.S.I.; Supervision, E.H. and J.I.; Validation, M.M.A., M.S.I. and A.D.B.; Writing—original draft, M.M.A., M.S.I. and A.D.B.; Writing—review and editing, E.H. and J.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors wish to convey special gratitude to Abdul Ghaffar Grewal and Maqbool Ahmad and the farm owner and the manager of Khakwani Mango Farms for their guidance through expert knowledge and administrative support during this research. We are also grateful to the Pakistan Meteorological Department for the provision of weather data for this research.

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

References

1. Bongiovanni, R.; Lowenberg-Deboer, J. Precision Agriculture and Sustainability. *Precis. Agric.* **2004**, *5*, 359–387. [\[CrossRef\]](#)
2. Zaman, Q. *Precision Agriculture: Evolution, Insights and Emerging Trends*; Elsevier: Amsterdam, The Netherlands, 2023.
3. Pourazar, H.; Samadzadegan, F.; Javan, F.D. Aerial Multispectral Imagery for Plant Disease Detection: Radiometric Calibration Necessity Assessment. *Eur. J. Remote Sens.* **2019**, *52*, 17–31. [\[CrossRef\]](#)
4. Ampatzidis, Y.; Partel, V. UAV-Based High Throughput Phenotyping in Citrus Utilizing Multispectral Imaging and Artificial Intelligence. *Remote Sens.* **2019**, *11*, 410. [\[CrossRef\]](#)
5. Rahman, M.M.; Robson, A.; Bristow, M. Exploring the Potential of High Resolution WorldView-3 Imagery for Estimating Yield of Mango. *Remote Sens.* **2018**, *10*, 1866. [\[CrossRef\]](#)
6. Torgbor, B.A.; Rahman, M.M.; Robson, A.; Brinkhoff, J.; Khan, A. Assessing the Potential of Sentinel-2 Derived Vegetation Indices to Retrieve Phenological Stages of Mango in Ghana. *Horticulturae* **2022**, *8*, 11. [\[CrossRef\]](#)

7. Wu, D.; Johansen, K.; Phinn, S.; Robson, A.; Tu, Y.H. Inter-Comparison of Remote Sensing Platforms for Height Estimation of Mango and Avocado Tree Crowns. *J. Appl. Geophys.* **2020**, *89*, 102091. [\[CrossRef\]](#)
8. Rajan, S.; Kumar, R.; Negi, S.S. Variation in Canopy Characteristics of Mango (*Mangifera indica* L.) Cultivars from Diverse Eco-Geographical Regions. *JAH* **2001**, *03*, 95–97. [\[CrossRef\]](#)
9. Maillard, P.; Gomes, M.F. Detection and Counting of Orchard Trees from VHR Images Using a Geometrical-Optical Model and Marked Template Matching. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* **2016**, *III-7*, 75–82. [\[CrossRef\]](#)
10. Mohan, M.; Silva, C.A.; Klauber, C.; Jat, P.; Catts, G.; Cardil, A.; Hudak, A.T.; Dia, M. Individual Tree Detection from Unmanned Aerial Vehicle (UAV) Derived Canopy Height Model in an Open Canopy Mixed Conifer Forest. *Forests* **2017**, *8*, 340. [\[CrossRef\]](#)
11. Afsar, M.M.; Bakhshi, A.D.; Iqbal, M.S.; Hussain, E.; Iqbal, J. High-Precision Mango Orchard Mapping Using a Deep Learning Pipeline Leveraging Object Detection and Segmentation. *Remote Sens.* **2024**, *16*, 3207. [\[CrossRef\]](#)
12. Johansen, K.; Raharjo, T.; McCabe, M.F. Using Multi-Spectral UAV Imagery to Extract Tree Crop Structural Properties and Assess Pruning Effects. *Remote Sens.* **2018**, *10*, 854. [\[CrossRef\]](#)
13. Nagaraja, A.; Sahoo, R.; Kalidindi, U.; Gupta, V.; Singh, S. Weather Based Mapping of Susceptible Areas of Mango Malformation through GIS. *Progress. Hortic.* **2012**, *44*, 201–2015.
14. Wang, L.; Tang, Y.; Liu, Z.; Zheng, M.; Shi, W.; Li, J.; He, X. Prediction of Thrips Damage Distribution in Mango Orchards Using a Novel Maximum Likelihood Classifier. *Agronomy* **2024**, *14*, 795. [\[CrossRef\]](#)
15. Khan, M.; Ahmed, N. Sustainable Management of Mango Nutrition for Better Yield and Quality. *Cercet. Agron. Mold.* **2020**, *53*, 473–501. [\[CrossRef\]](#)
16. Colomina, I.; Molina, P. Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [\[CrossRef\]](#)
17. Mogili, U.R.; Deepak, B.B.V.L. Review on Application of Drone Systems in Precision Agriculture. *Procedia Comput. Sci.* **2018**, *133*, 502–509. [\[CrossRef\]](#)
18. Berzéki, M.; Kozma-Bognár, V.; Berke, J. Examination of Vegetation Indices Based on Multitemporal Drone Images. *GRADUS* **2023**, *10*, 1–6. [\[CrossRef\]](#)
19. Sarron, J.; Malezieux, E.; Sane, C.; Faye, E. Mango Yield Mapping at the Orchard Scale Based on Tree Structure and Land Cover Assessed by UAV. *Remote Sens.* **2018**, *10*, 1900. [\[CrossRef\]](#)
20. Narmilan, A.; Gonzalez, F.; Salgado, A.S.A.; Powell, K. Detection of White Leaf Disease in Sugarcane Using Machine Learning Techniques over UAV Multispectral Images. *Drones* **2022**, *6*, 230. [\[CrossRef\]](#)
21. Osco, L.P.; Ramos, A.P.M.; Pereira, D.R.; Moriya, E.A.S.; Imai, N.N.; Matsubara, E.T.; Estrabis, N.; de Souza, M.; Junior, J.M.; Eduardo Creste, J.; et al. Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote Sens.* **2019**, *11*, 2925. [\[CrossRef\]](#)
22. Datt, R.M.; Kukreja, V. Neural Network Model for Predicting Apple Yield Based on Arrival of Phenological Stage in Conjunction with Leaf Disease, Soil and Weather Parameters. *SN Comput. Sci.* **2024**, *5*, 141. [\[CrossRef\]](#)
23. Olofintuyi, S.S.; Olajubu, E.A.; Olanike, D. An Ensemble Deep Learning Approach for Predicting Cocoa Yield. *Heliyon* **2023**, *9*, e15245. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Magalhães, I.A.L.; de Carvalho Júnior, O.A.; de Carvalho, O.L.F.; de Albuquerque, A.O.; Hermuche, P.M.; Merino, E.R.; Gomes, R.A.T.; Guimaraes, R.F. Comparing Machine and Deep Learning Methods for the Phenology-Based Classification of Land Cover Types in the Amazon Biome Using Sentinel-1 Time Series. *Remote Sens.* **2022**, *14*, 4858. [\[CrossRef\]](#)
25. Salehinejad, H. Learning Over Long Time Lags. *arXiv* **2016**, arXiv:1602.04335.
26. Karim, F.; Majumdar, S.; Darabi, H. Insights Into LSTM Fully Convolutional Networks for Time Series Classification. *IEEE Access* **2019**, *7*, 67718–67725. [\[CrossRef\]](#)
27. Lu, J.; Li, J.; Fu, H.; Tang, X.; Liu, Z.; Chen, H.; Sun, Y.; Ning, X. Deep Learning for Multi-Source Data-Driven Crop Yield Prediction in Northeast China. *Agriculture* **2024**, *14*, 794. [\[CrossRef\]](#)
28. Wang, J.; Si, H.; Gao, Z.; Shi, L. Winter Wheat Yield Prediction Using an LSTM Model from MODIS LAI Products. *Agriculture* **2022**, *12*, 1707. [\[CrossRef\]](#)
29. Bhimavarapu, U.; Battineni, G.; Chintalapudi, N. Improved Optimization Algorithm in LSTM to Predict Crop Yield. *Computers* **2023**, *12*, 10. [\[CrossRef\]](#)
30. Dolaptsis, K.; Pantazi, X.E.; Paraskevas, C.; Arslan, S.; Tekin, Y.; Bantchina, B.B.; Ulusoy, Y.; Gündoğdu, K.S.; Qaswar, M.; Bustan, D.; et al. A Hybrid LSTM Approach for Irrigation Scheduling in Maize Crop. *Agriculture* **2024**, *14*, 210. [\[CrossRef\]](#)
31. Noguera-Polania, J.; Daconte-Blanco, A.; Moreu-Ceballos, J.; Linero-Ospino, C.; Munera-Luque, R.; Guevara-Barbosa, P. Preliminary Results of Irrigation Management for Mango Using LSTM Neural Networks and IoT. *Rev. Fac. Ing. Univ. Antioq.* **2024**. [\[CrossRef\]](#)
32. Apolo-Apolo, O.E.; Martínez-Guanter, J.; Egea, G.; Raja, P.; Pérez-Ruiz, M. Deep Learning Techniques for Estimation of the Yield and Size of Citrus Fruits Using a UAV. *Eur. J. Agron.* **2020**, *115*, 126030. [\[CrossRef\]](#)

33. Veling, S.S.; Mohite-Patil, T.B. Multi-Disease Classification of Mango Tree Using Meta-Heuristic-Based Weighted Feature Selection and LSTM Model. *Int. J. Image Graph.* **2023**, *24*, 2450039. [CrossRef]
34. Borianne, P.; Sarron, J.; Borne, F.; Faye, E. Deep mango cultivars: Cultivar detection by classification method with maximum misidentification rate estimation. *Precis. Agric.* **2023**, *24*, 1619–1637. [CrossRef]
35. Torgbor, A.B.; Brinkhoff, J.; Rahman, M.M.; Robson, A.; Sinha, P. Integrating Remote Sensing and Weather Variables for Yield Forecasting of Horticultural Tree Crops—A Case Study of Mango in Ghana and Australia. Ph.D. Dissertation, University of New England, Sydney, Australia, 2024.
36. Torgbor, B.A.; Rahman, M.M.; Brinkhoff, J.; Sinha, P.; Robson, A. Integrating Remote Sensing and Weather Variables for Mango Yield Prediction Using a Machine Learning Approach. *Remote Sens.* **2023**, *15*, 3075. [CrossRef]
37. Lee, J.; Gadsden, S.A.; Biglarbegian, M.; Cline, J.A. Smart Agriculture: A Fruit Flower Cluster Detection Strategy in Apple Orchards Using Machine Vision and Learning. *Appl. Sci.* **2022**, *12*, 11420. [CrossRef]
38. Koirala, A.; Walsh, K.B.; Wang, Z.; Anderson, N. Deep Learning for Mango (*Mangifera Indica*) Panicle Stage Classification. *Agronomy* **2020**, *10*, 143. [CrossRef]
39. Nagaraja, A.; Sahoo, R.; Kalidindi, U.; Singh, S.K.; Sivaramanae, N.; Gupta, V.K. Spectral Discrimination of Healthy and Malformed Mango Panicles Using Spectroradiometer. *Indian J. Hort.* **2014**, *71*, 40–44.
40. Ullah, H.; Malik, M.T.; Khan, A.H. *A Guide for Precise Mango Cultivation*; Mango Research Institute Multan: Multan, Pakistan, 2020; p. 146.
41. Bakhsh, K.; Hassan, I.; Akhter, M. Profitability and Cost in Growing Mango Orchards. 2008. Available online: <https://api.semanticscholar.org/CorpusID:13906961> (accessed on 16 February 2025).
42. Iqbal, M.; Niamatullah, M.; Hussain, A.; Munir, M.; Khan, I.; Khan, M.Q. Performance of selected parameters of mango cultivars in Muzaffargarh district (Punjab), Pakistan. *Sarhad J. Agric.* **2012**, 395–398. Available online: <https://www.semanticscholar.org/paper/PERFORMANCE-OF-SELECTED-PARAMETERS-OF-MANGO-IN-%2C-Iqbal-Niamatullah/0d692e56ae8406fd4e1908be8b5a2fe79b9c6688> (accessed on 1 February 2025).
43. Grewal, A.G. Fruiting behavior and fruit quality of leading mango cultivars grown in South Punjab-Pakistan. *Agric. Sci. J.* **2024**, *6*, 59–72. [CrossRef]
44. DJI. DJI Mavic 3M—Specifications. Available online: <https://ag.dji.com/mavic-3-m/specs> (accessed on 23 June 2024).
45. DJI. D-RTK 2—Product Information. Available online: <https://www.dji.com/global/d-rtk-2/info> (accessed on 23 June 2024).
46. PIX4Dmapper. PIX4Dmapper—Documentation. Available online: <https://support.pix4d.com/hc/en-us/articles/115002472186> (accessed on 1 February 2025).
47. Torgbor, B.A.; Sinha, P.; Rahman, M.M.; Robson, A.; Brinkhoff, J.; Suarez, L.A. Exploring the Relationship Between Very-High-Resolution Satellite Imagery Data and Fruit Count for Predicting Mango Yield at Multiple Scales. *Remote Sens.* **2024**, *16*, 4170. [CrossRef]
48. Stephen, S.; Haldar, D.; Patel, N.R. Impact of Various Vegetation Indices on Mango Orchard Mapping Using Object-Based Image Analysis. *J. Geomat.* **2022**, *16*, 159–166. [CrossRef]
49. Verma, H.C.; Ahmed, T. A Comprehensive Analysis of Advanced Change Detection Technique for Planning and Monitoring Sustainable Mango Crop Areas Using Landsat-8 Images. In Proceedings of the 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 15–17 March 2023; pp. 1336–1341.
50. Zang, Y.; Chen, X.; Chen, J.; Tian, Y.; Shi, Y.; Cao, X.; Cui, X. Remote Sensing Index for Mapping Canola Flowers Using MODIS Data. *Remote Sens.* **2020**, *12*, 3912. [CrossRef]
51. Shao, C.; Shuai, Y.; Wu, H.; Deng, X.; Zhang, X.; Xu, A. Development of a Spectral Index for the Detection of Yellow-Flowering Vegetation. *Remote Sens.* **2023**, *15*, 1725. [CrossRef]
52. Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* **2013**, *173*, 74–84. [CrossRef]
53. Rahman, M.; Robson, A.; Brinkhoff, J. Potential of Time-Series Sentinel 2 Data for Monitoring Avocado Crop Phenology. *Remote Sens.* **2022**, *14*, 5942. [CrossRef]
54. Niu, H.; Wang, D.; Ehsani, R.; Chen, Y. Tree-Level Yield Estimation Using UAV-Based Vegetation Indices and Plant Physiology-Informed Machine Learning. In Proceedings of the 2022 International Conference on Unmanned Aircraft Systems (ICUAS), Dubrovnik, Croatia, 21–24 June 2022; pp. 1614–1619.

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.