



**HAL**  
open science

# Hybrid-AI and model ensembling to exploit UAV-Based RGB imagery: an evaluation of sorghum crop's nitrogen content

Hajar Hammouch, Suchitra Patil, Sunita Choudhary, Mounim A El-Yacoubi, Jan Masner, Jana Kholová, Krithika Anbazhagan, Jiří Vaněk, Huafeng Qin, Michal Stočes, et al.

## ► To cite this version:

Hajar Hammouch, Suchitra Patil, Sunita Choudhary, Mounim A El-Yacoubi, Jan Masner, et al.. Hybrid-AI and model ensembling to exploit UAV-Based RGB imagery: an evaluation of sorghum crop's nitrogen content. *Agriculture*, 2024, 14 (10), pp.1682. 10.3390/agriculture14101682. hal-04827587

**HAL Id: hal-04827587**

**<https://hal.science/hal-04827587v1>**

Submitted on 9 Dec 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.







L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## Article

# Hybrid-AI and Model Ensembling to Exploit UAV-Based RGB Imagery: An Evaluation of Sorghum Crop's Nitrogen Content

Hajar Hammouch <sup>1,2,†</sup>, Suchitra Patil <sup>3,4,5,†</sup> , Sunita Choudhary <sup>4,\*</sup>, Mounim A. El-Yacoubi <sup>1</sup> , Jan Masner <sup>6,\*</sup> , Jana Kholová <sup>4,6</sup>, Krithika Anbazhagan <sup>7</sup>, Jiří Vaněk <sup>6</sup>, Huafeng Qin <sup>8</sup> , Michal Stočes <sup>6</sup>, Hassan Berbia <sup>2</sup>, Adinarayana Jagarlapudi <sup>3</sup> , Magesh Chandramouli <sup>9</sup>, Srinivas Mamidi <sup>10</sup>, KVS Prasad <sup>7</sup>  and Rekha Baddam <sup>4</sup>

- <sup>1</sup> SAMOVAR, Telecom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France; hajar.hammouch@telecom-sudparis.eu (H.H.); mounim.el\_yacoubi@telecom-sudparis.eu (M.A.E.-Y.)
  - <sup>2</sup> SSLAB, Ecole Nationale Supérieure d'Informatique et d'Analyse des Systèmes, Mohamed V University, Rabat 10100, Morocco; hassan.berbia@ensias.um5.ac.ma
  - <sup>3</sup> Centre of Studies in Resources Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, Maharashtra, India; suchitrapatil@somaiya.edu (S.P.); adi@iitb.ac.in (A.J.)
  - <sup>4</sup> Crop Physiology and Modeling, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Telangana, India; jana.kholova@icrisat.org (J.K.); rekha.baddam@icrisat.org (R.B.)
  - <sup>5</sup> Department of Information Technology, K. J. Somaiya College of Engineering Vidyavihar, Mumbai 400 077, Maharashtra, India
  - <sup>6</sup> Department of Information Technologies, Czech University of Life Sciences Prague, 165 00 Prague, Czech Republic; vanek@pef.czu.cz (J.V.); stoces@pef.czu.cz (M.S.)
  - <sup>7</sup> International Livestock Research Institute (ILRI), Patancheru, Hyderabad 502 324, Telangana, India; krithika.anbazhagan@icrisat.org (K.A.); k.v.prasad@cgiar.org (K.P.)
  - <sup>8</sup> National Research base of Intelligent Manufacturing Service, Chongqing Technology and Business University, Chongqing 400067, China; qinhuafengfeng@163.com
  - <sup>9</sup> Computer Graphics Technology, Purdue University NW, Hammond, IN 46323, USA; mchandr@pnw.edu
  - <sup>10</sup> Marut Dronetech Private Limited, Gachibowli, Hyderabad 500 032, Telangana, India; srinivas@marutdrones.com
- \* Correspondence: sunita.choudhary@icrisat.org (S.C.); masner@pef.czu.cz (J.M.)  
† These authors contributed equally to this work.



**Citation:** Hammouch, H.; Patil, S.; Choudhary, S.; El-Yacoubi, M.A.; Masner, J.; Kholová, J.; Anbazhagan, K.; Vaněk, J.; Qin, H.; Stočes, M.; et al. Hybrid-AI and Model Ensembling to Exploit UAV-Based RGB Imagery: An Evaluation of Sorghum Crop's Nitrogen Content. *Agriculture* **2024**, *14*, 1682. <https://doi.org/10.3390/agriculture14101682>

Received: 14 June 2024

Revised: 30 August 2024

Accepted: 6 September 2024

Published: 26 September 2024



**Copyright:** © 2024 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/>).

**Abstract:** Non-invasive crop analysis through image-based methods holds great promise for applications in plant research, yet accurate and robust trait inference from images remains a critical challenge. Our study investigates the potential of AI model ensembling and hybridization approaches to infer sorghum crop traits from RGB images generated via unmanned aerial vehicle (UAV). In our study, we cultivated 21 sorghum cultivars in two independent seasons (2021 and 2022) with a gradient of fertilizer and water inputs. We collected 470 ground-truth N measurements and captured corresponding RGB images with a drone-mounted camera. We computed five RGB vegetation indices, employed several ML models such as MLR, MLP, and various CNN architectures (season 2021), and compared their prediction accuracy for N-inference on the independent test set (season 2022). We assessed strategies that leveraged both deep and handcrafted features, namely hybridized and ensembled AI architectures. Our approach considered two different datasets collected during the two seasons (2021 and 2022), with the training set from the first season only. This allowed for testing of the models' robustness, particularly their sensitivity to concept drifts, in the independent season (2022), which is fundamental for practical agriculture applications. Our findings underscore the superiority of hybrid and ensembled AI algorithms in these experiments. The MLP + CNN-VGG16 combination achieved the best accuracy ( $R^2 = 0.733$ , MAE = 0.264 N% on an independent dataset). This study emphasized that carefully crafted AI-based models applied to RGB images can achieve robust trait prediction with accuracies comparable to the similar phenotyping tasks using more complex (multi- and hyper-spectral) sensors presented in the current literature.

**Keywords:** AI; machine learning; UAV; RGB; nitrogen; phenotyping

## 1. Introduction

Research on plant characterization (i.e., phenotyping) is gradually transiting from traditional methods (visual observations, manual measurements) to non-invasive methods facilitated by sensors (reviewed by, e.g., [1–3]). Phenotyping includes plant digitization via sensors (creation of plant digital twin, (Demidchik et al., 2020); [4]) and building mathematical models relating the plant digital twin to the actual plant's functional and structural characteristics. Sensor-based phenotyping has potential to automate, standardize, and optimize the throughput of phenotyping tasks (e.g., [1,5,6]).

RGB imaging, particularly in combination with mobile vectors like UAVs (unmanned aerial vehicles), has been widely used in precision agriculture tasks due to its simplicity and cost-effectivity [5,6]. In this regard, remote sensing based on spectral reflectance indexes specific to the canopy (i.e., vegetation indexes (VIs)) are being intensively studied, e.g., [7].

Although these methods have been around for at least two decades, the repeatability, standardization, and interpretation of these VIs, particularly for precise crop characterization, are still questioned today ([8,9]).

In the early 2010s, there was a tendency to improve the accuracy of trait inference by using more complex sensors (e.g., multi-spectral and hyper-spectral cameras) with considerable success ([10–15]). Nevertheless, using more complex sensing methods complicates image capture protocols and, consequently, image processing methods which could pose considerable obstacles for many end-uses ([8,15–17]). This might be why recently, many researchers have resorted back to RGB imaging whilst also testing more advanced AI-based algorithms to infer the crop traits ([10–14]).

This trend coincides with intense exploration of artificial intelligence (AI) models in agricultural research (e.g., [11–14]). Here, researchers generally assess traditional machine learning (ML) methods (e.g., linear regression (LR), partial least squares (PLSR), principal component analysis (PCA), adaptive boosting (AB), K-nearest neighbor (KNN), random forest (RF), support vector machine (SVM)), or artificial neural networks (ANN) (e.g., multi-layer perceptron (MLP) and convolutional neural networks (CNN)). ANN-based methods are generally reported to outperform the traditional ML methods achieving accuracies suitable for practical applications like crop breeding (generally  $> 0.75 R^2$  e.g., in [13,14]). Nevertheless, most of the studies do not report the model accuracy on external datasets (e.g., application of the model on images obtained in different contexts (field/season)) to ensure the models address the concept drift which is an important drawback for use in agricultural applications [18]. Overfitting of AI-based algorithms to training datasets is a known phenomenon hindering the trustworthiness of AI-based algorithms in applications and is intensively researched by the data-research community e.g., [19].

To address this issue, we have investigated how accurately we could infer the sorghum crop features (N-content of biomass) from simple RGB-imaging technology (UAV-based RGB-imaging) using ML and ANN models when they are integrated with expert knowledge in different ways. Here, we hypothesized that combining expert knowledge (such as VIs) with more complex AI structures can enhance sorghum trait prediction accuracy. Importantly, we generated an additional independent dataset solely to assess the generalization capabilities of these models and their robustness to concept drift. In the end, we aim to find a trustworthy suite of tools for application in sorghum breeding programs that need to monitor sorghum N with high throughput and relevant accuracy.

## 2. Materials and Methods

Two independent field trials were conducted, as described in Section 2.1.1, and multiple UAV flights were carried out as detailed in Section 2.1.2 to capture images of the crop at various growth stages. Laboratory measurements of N content in the imaged crop were performed, creating a ground truth dataset as discussed in Section 2.1.3. The acquired images underwent image-processing techniques and orthomosaic generation (Section 2.2). Specific details about the dataset are provided in Sections 2.3 and 2.4. Evaluating the models' performance involved comparing ground truth observations (biochemical N estimation) with

predictions from trained models, through metrics, namely coefficient of determination ( $R^2$ ), mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ). Concept drift was investigated by applying the models to an independent dataset acquired at a later stage. The training of diverse models is outlined in Section 2.6, along with the ensemble in Section 2.6.5.

## 2.1. Data Acquisition

### 2.1.1. Plant Material and Experiment Details

Two field trials were planted on 26 October 2021 and on 10 November 2022 in post-rainy (rabi) seasons "2021–2022" and "2022–2023" at the International Crop Research Institute for Semi-arid Tropics (ICRISAT, Patancheru, Telangana, India, latitude 17.53° N, longitude 78.27° E). The sorghum crop was raised on alfisol soil organized into 54 plots of 7 m × 5 m for rabi 2021 and 36 plots of 10 m × 5 m for rabi 2022 and was prepared as per standard agronomic practices [20]. Nine (rabi 2021) and twelve (rabi 2022) genotypes of sorghum (*Sorghum bicolor*) contrasting for agronomic characteristics were used ([21]). Each experiment included combinations of two irrigation (fully irrigated or well-watered (WW), limited water supply or water-stressed (WS)) and two fertilization (full N-fertilizer or standard nitrogen (SN), and limited N-fertilizer or low nitrogen (LN)) regimes resulting in three blocks of treatments: (1) well-watered and standard nitrogen (WWSN), (2) water-stressed and standard nitrogen (WSSN), and (3) well-watered and low nitrogen (WWLN). Within each block of treatments, two replications of each genotype-treatment combination were randomized in each block (i.e., split-plot CRBD) in rabi 2021 while there were no replications of genotype-treatment combinations in rabi 2022.

To raise the crop in 2021, a basal dose of diammonium phosphate (DAP) was applied prior to sowing at the rate of 100 kg ha<sup>-1</sup> to WWSN and WSSN treatments in both trials. Top-dressing of urea ~20 days after sowing (DAS) was applied at the rate of 50 kg ha<sup>-1</sup> for standard nitrogen plots and 25 kg ha<sup>-1</sup> for low nitrogen plots and the same dose was repeated ~35 DAS. Similarly, in the rabi 2022, a top dressing of urea was applied at the rate of 50 kg ha<sup>-1</sup> and 15 kg ha<sup>-1</sup> for standard and low nitrogen plots ~20 DAS and 35 kg ha<sup>-1</sup> top dressing of urea ~30 DAS was applied to only standard nitrogen plots. In the rabi 2021, all the plots were irrigated every ~10 days until 50 DAS, and later, the irrigation ceased for WSSN treatments while continued for WWSN and WWLN treatments. Similarly, in the rabi 2022, all plots were well irrigated until 16 DAS and thereafter irrigation was reduced by ~30% and stopped after 46 DAS for WS treatments while continued for WW treatments. A total of 12 circular ground control points (GCP) were placed, covering the field's boundaries for both field trials.

### 2.1.2. UAV Setup and Flight Details for Image Collection

The image collection protocol followed the image acquisition protocol described in [22,23]. The quadcopter DJI Matrice 210 (DJI, Shenzhen, Nanshan District, China, 2023) with an onboard ZENMUSE X5S (DJI, Shenzhen, Nanshan District, China; [24]) RGB camera was used to capture high-resolution RGB imagery. The details of UAV flights taken (UAV settings and camera settings) are given in [22,23] and in Supplementary Table S1, while flight dates (late vegetative and flowering stages of crop) are listed in Table 1. All the flights were taken in a clear and sunny sky with even ambiance and in the daytime window of 10 am to 2 pm. During all the flights, the wind speed was less than 18 km h<sup>-1</sup> which is mandatory to maintain the stability of the UAV and the crop canopy. The front and side overlap of 80 to 90% was maintained with the camera looking down at 90°, the nadir view (Table S1). Each flight taken in this way generated 128 raw RGB images.

**Table 1.** The overview of the plot images generated from raw imagery collected for the present study, including flight dates, ground truth harvesting dates, and field trial season and its use for the model-building exercise (split into training, validation, and test sets).

Field Trial (Section 2.1.1)	Flights Timing (Section 2.1.2)	Corresponding Ground Truth (Section 2.1.3)	Dataset Use (Section 2.3)	Plot Images (Absolute/%)
Season 1 (rabi 2021–2022)	3 December 2021	3 December 2021	Training set	122 (80%)
	21 December 2021	21 December 2021		
	18 January 2022	20 January 2022	Validation set	31 (20%)
Season 2 (rabi 2022–2023)	8 December 2021	3 December 2021	Test set 1	102
	7 January 2022	20 January 2022		
	16 December 2022	14 December 2022	Test set 2	215
	21 December 2022	21 December 2022		
	27 December 2022	28 December 2022		
	5 January 2023	4 January 2023		
11 January 2023	11 January 2023			
18 January 2023	18 January 2023			

### 2.1.3. Ground Truth Collection

In this study, we collected data in a series of sequential harvests of sorghum crops during the vegetative stages of growth (between 7 and 11 weeks after sowing, i.e., before flowering). In the Season 1 (2021–2022) trial, the data collected from the first three harvests was used in this study. Similarly, for the Season 2 trial (2022–2023), the first six sequential harvests were taken in an interval of one week across vegetative crop growth stages before flowering (between 5 and 10 weeks after sowing). The ground truth collection dates are mentioned in Table 1. At each ground truth collection time (Table 1), eight representative plants (subsamples) in an area of 4.8 m<sup>2</sup> were harvested from each plot. These plants were brought to the laboratory, where leaves and stems were separated and dried in the oven for four days at 600 °C. The dried plant matter was ground using a CM 290 Cemotec<sup>TM</sup> laboratory grinder (FOSS, Hilleroed, Denmark; [25]) to a uniform particle size of <1 mm. The analysis for N content was performed at the Livestock Nutritional Service Laboratory, International Livestock Research Institute (ILRI) at ICRISAT campus using near infrared spectroscopy (NIRS) calibrated against conventional laboratory analyses. The NIRS instrument used was a FOSS Analyzer DS2500 (FOSS, Hilleroed, Denmark) with the software package WinISI II (4.8). The NIRS-based procedure for N analysis was used as per global recommended standards [26,27]; the particular analysis for cereal stover quality traits using NIRS has been described in detail in [28,29]).

Statistically, Season 2 had a higher average nitrogen content (2.11) compared to Season 1 (1.71), with a broader range and greater variability. Both datasets showed a roughly normal distribution, but Season 2 included some higher outliers. Overall, Season 2 exhibited a generally higher nitrogen content and more variability. Detailed descriptive statistics are presented in Supplementary Materials Table S7 and Figure S1.

## 2.2. Orthomosaic Generation and Delineation of Crop Plots

### 2.2.1. Quality Check

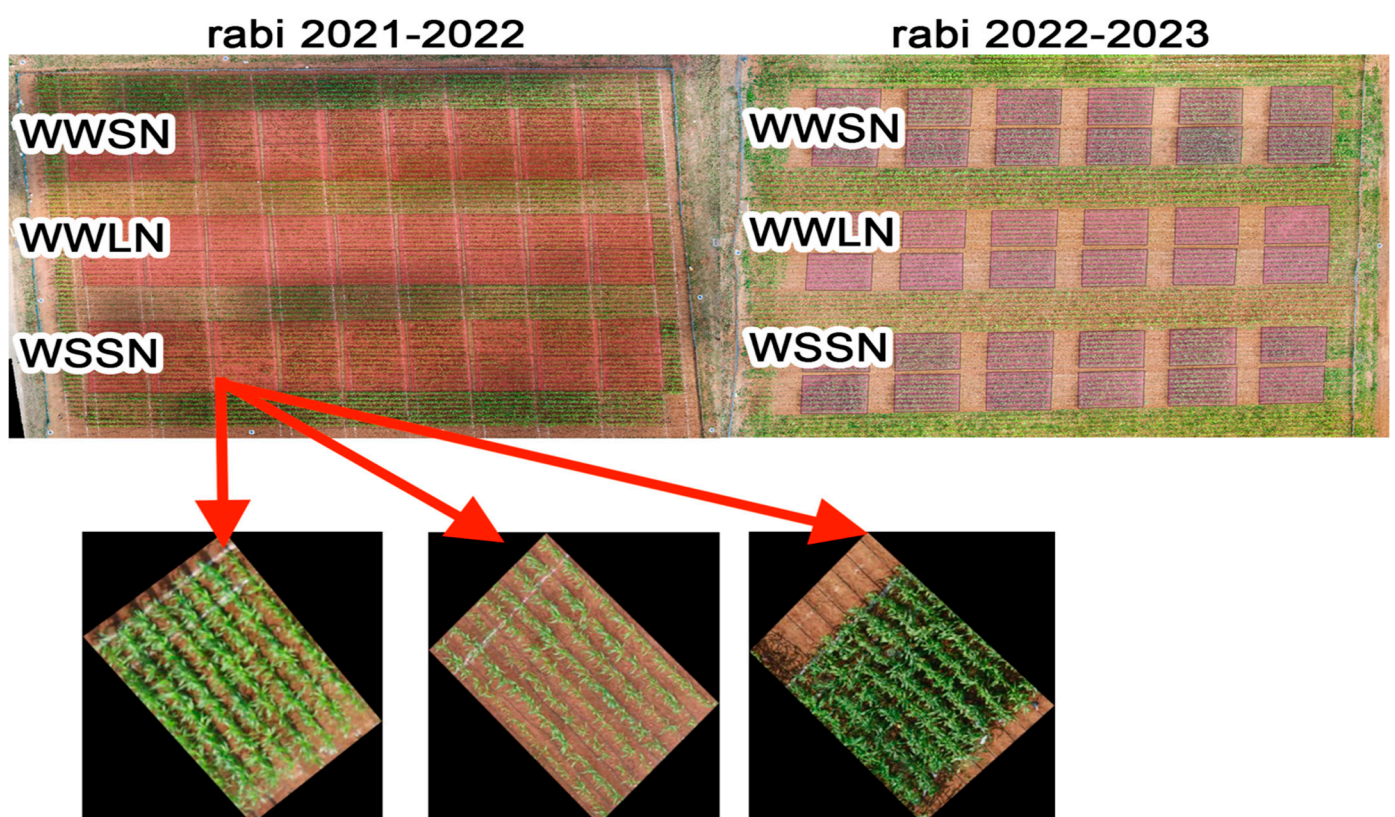
A two-fold quality check was applied to the raw images collected from each flight as defined by [22]. Firstly, the images were passed through the quality check pipeline and the value quality parameters such as the count of underexposed pixels (0–5) and overexposed pixels (200–255), DCT blur ( $<0.2 \times 10^{-3}$ ) [30], NIQE ([31]) (<5), and BRISQE [32] (<35) were computed (for each parameter, the lower value was preferred). The images with values of these parameters in a given range were regarded as quality images. Secondly, these images were also tested for quality using photogrammetry software Agisoft Metashape Professional's (v 2.0) [33] estimate quality option. Agisoft computes image quality based



on the contrast and brightness in images and returns values between 0 and 1. This study used all the images with Agisoft estimated quality > 0.75.

### 2.2.2. Orthomosaic Generation

Once the two-fold quality check was passed, the raw RGB images were stitched using Agisoft to generate the orthomosaic of the entire field for the corresponding growth stage. The first step in this process [33] was to align the raw imagery using Agisoft where common tie points were detected from the overlapped regions of adjacent image pairs using scale-invariant features transformation (SIFT) [34]; these tie points represent a sparse point cloud. Further, using the structure from the motion (SfM) technique [35], a dense 3D point cloud was generated, and rasterized using interpolation to achieve the depth maps. Finally, using these depth maps and 3D point clouds, an orthomosaic of up to 3 to 5mm of horizontal resolution was generated [33]. The orthomosaic represented the entire sorghum field with all plots (Figure 1).



**Figure 1.** The overview of the field experiments, Sample locations of orthomosaic for rabi 2021 and rabi 2022 including treatment on top, and RGB images of sample sorghum plots.

### 2.2.3. Delineation of Plots from Orthomosaic

To delineate the individual plots from this orthomosaic, a plot-wise grid was created and a shapefile was generated using QGIS 2.23 [36]. Using this shapefile, the sorghum plots from the orthomosaic of each flight were delineated. Samples of the generated plot images are shown in Figure 1.

### 2.3. Data Distribution for Comparative Analysis and Model Training

The pre-processed data, discussed in Section 2.2, was divided into distinct sets: the Training set, the Validation set, and Test sets 1 and 2, details of which are provided in Table 1. A total of 15 plots in rabi 2021–2022 and 1 plot in rabi 2022–2023 were excluded from the dataset as the ground-truth value of nitrogen was missing due to poor germination. There was not enough material in the sample to measure its N content. The training/validation set ratio was set to 80:20. This dataset distribution was consistently applied across all models developed within this study, ensuring accurate comparisons. Standardization and normalization procedures were employed on the datasets with both inputs and outputs used to enhance training efficiency while avoiding saturation. This was achieved via standard scaling and computed using  $Z = (x - \mu)/\sigma$ , where  $\mu$  is the mean and  $\sigma$  is the standard deviation of the distribution used to attain the scaled value  $Z$  for each  $x$  value.

### 2.4. Generation of Different Vegetation Indices

Five different vegetation indices (VI) were computed pixel-wise from RGB values of plot images. For each plot, each VI was represented by its median value. The band arithmetic for different VIs such as Excess green index (EXGR), Green chromaticity coordinate (GCC), Green leaf index (GLI; [37]), Green red difference index (GRD), and difference in squared greenness and product of red and blue (RGBVI; [8,38]) were applied to extract the greenness from the crop. The purpose of the selection of these VIs was to compute the green color dominant features of the crop as it is related to its chlorophyll content and is considered an indicator of the crop N status. Although it was not possible to completely eliminate the impact of noise caused by uneven illuminance and reflectance from non-vegetative objects, the VIs were used here to normalize the data by mitigating these effects. This improves the consistency and reliability of the analysis by addressing potential radiometric inconsistencies [8,9]. Additionally, indices EXGR and GRD are known for separating crop pixels from the ground pixels.

### 2.5. Model Evaluation Metrics

To assess our ML models and ANN architectures we computed different standardized metrics, namely coefficient of determination ( $R^2$ ), mean absolute error (MAE), root mean square error (RMSE), and Pearson's correlation coefficient ( $r$ ).

### 2.6. Prediction Models

In this section, we provide a description of the different models and their combinations that we evaluated. A complete overview of all the models in one place can be found in Supplementary Table S3. All the models were trained using early stopping to avoid overfitting.

#### 2.6.1. Multiple Linear Regression (MLR) Model

Linear models can learn the linear relationship between the green pigments in crops and the crop's nutritional status in terms of its N content [39]. Five green color dominant VIs, namely EXGR, GCC, GLI, GRD, and RGBVI, were computed as described in Section 2.4. A single vegetation index may not be able to effectively explore the relationship with the %N content of the crop. Even if it is ideal to have one VI for one crop trait, relying only on one VI may be impossible, as it may suffer from problems due to uneven illumination and other environmental conditions. Hence, we have explored MLR using all five handcrafted features (5 VIs) as predictors and values of %N content as a dependent variable. This model was used as a baseline for the other models (Supplementary Table S2).

### 2.6.2. Hybrid Multi-Layer Perceptron Model

ML models and ANNs can be used effectively for extracting both spatial and spectral features from images. Multi-layer perceptron (MLP) regressors are known for handling small-sized datasets and limited feature sizes. However, their performance depends upon the amount of noise present in the data, the complexity of the relationship between input and output parameters, the choice of hyper-parameters, and the regularization techniques used [40].

An MLP that leverages expert knowledge providing the five VIs (EXGR, GCC, GLI, GRD, and RGBVI) as input and a single output neuron was experimented on with various hidden layers. Initially, the number of neurons for the hidden layer was selected according to the thumb rule stating that the number of hidden neurons can be  $((input\_size * 2/3) + output\_size)$  or less than twice the  $input\_size$  (Xu and Chen, 2008) [41]. This is required to ensure that sufficient numbers of hidden neurons are available to explore the complex relationship between the input and output, avoiding model overfitting and underfitting. Later, we experimented with a number of hidden neurons from 5 to 9, and an SGD optimizer with different batch sizes, dropout rates, and early stopping. MLPs with a single hidden layer were built, with their hyper-parameters tuned. Afterward, different MLP architectures with two hidden layers were created (with 5, 6, 7, 8, and 9 neurons in the first hidden layer and 4, 5, and 6 in the second), and subjected to hyper-parameter optimization.

The final model that converged with minimum validation losses consisted of an input layer ( $5 \times 1$ ), dense layer ( $8 \times 1$ ), dense layer ( $5 \times 1$ ) + dropout (0.3), and output layer ( $1 \times 1$ ). We used an SGD optimizer, tanh activation function in hidden layers, and linear activation function for the output layer. The MSE was used as a loss function to minimize during model training.

### 2.6.3. CNN Model from Scratch

Deep neural networks can automatically extract deep features from images [13]. CNN is a state-of-the-art model architecture used in computer vision that can learn relevant features from images at different convolutional levels; it uses pooling layers similar to the human visual system [42]. A simple CNN model was adopted to accommodate the small size of the training set and the desire to deploy a computationally efficient model in a production environment.

Initially, we built a simple CNN model from scratch. The optimal CNN architecture was obtained via a Greedy optimization method based on various hyper-parameters such as filter size, batch size, optimizer, learning rate, number of epochs, and dropout. The final architecture consists of an input layer ( $500 \times 500 \times 3$ ), 7 convolutional layers ( $32 (3 \times 3) + dropout$ ;  $32 (3 \times 3) + maxPooling$ ;  $64 (3 \times 3) + dropout$ ;  $64 (3 \times 3) + maxPooling$ ;  $128 (3 \times 3) + dropout$ ;  $128 (3 \times 3) + maxPooling$ ), and an output layer ( $1 \times 1$ ) that predicts the N content.

### 2.6.4. Transfer Learning Using State-of-the-Art Pre-Trained Models

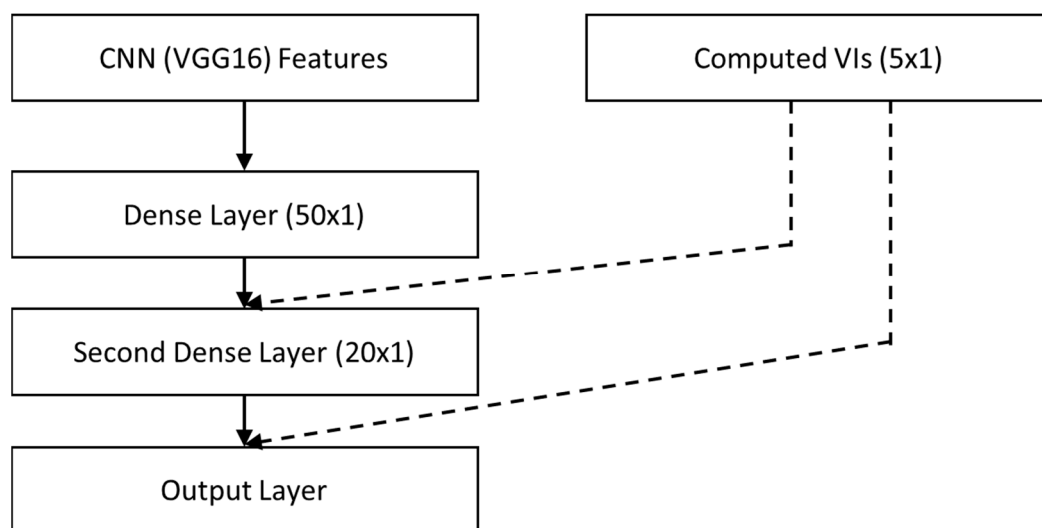
We assessed several architectures using well-established pre-trained classification models of varying sizes available in the Keras Applications library [43]. These models were pre-trained on an extensive ImageNet dataset [44]. We removed the last layer from each model and appended two fully connected layers (50 neurons in the first layer and 20 neurons in the second layer) to make uniform modifications. The training was limited to 200 epochs, using early stopping based on the MSE as a loss function. We utilized the RMSprop optimizer to train the models to maintain comparability.



### 2.6.5. Model Ensembling

We explored two different approaches for combining features/models to attain a robust and precise estimation of N. Our first approach was to explore different possible combinations such as (i) MLP + CNN (VGG16), (ii) CNN (VGG16) + MLR, (iii) MLP + MLR, and (iv) MLP + CNN (VGG16) + MLR. In each combination, the predictions from these models, before applying rescaling, were averaged. The resulting values were rescaled back to the original N values (the data normalization method is described in Section 2.3).

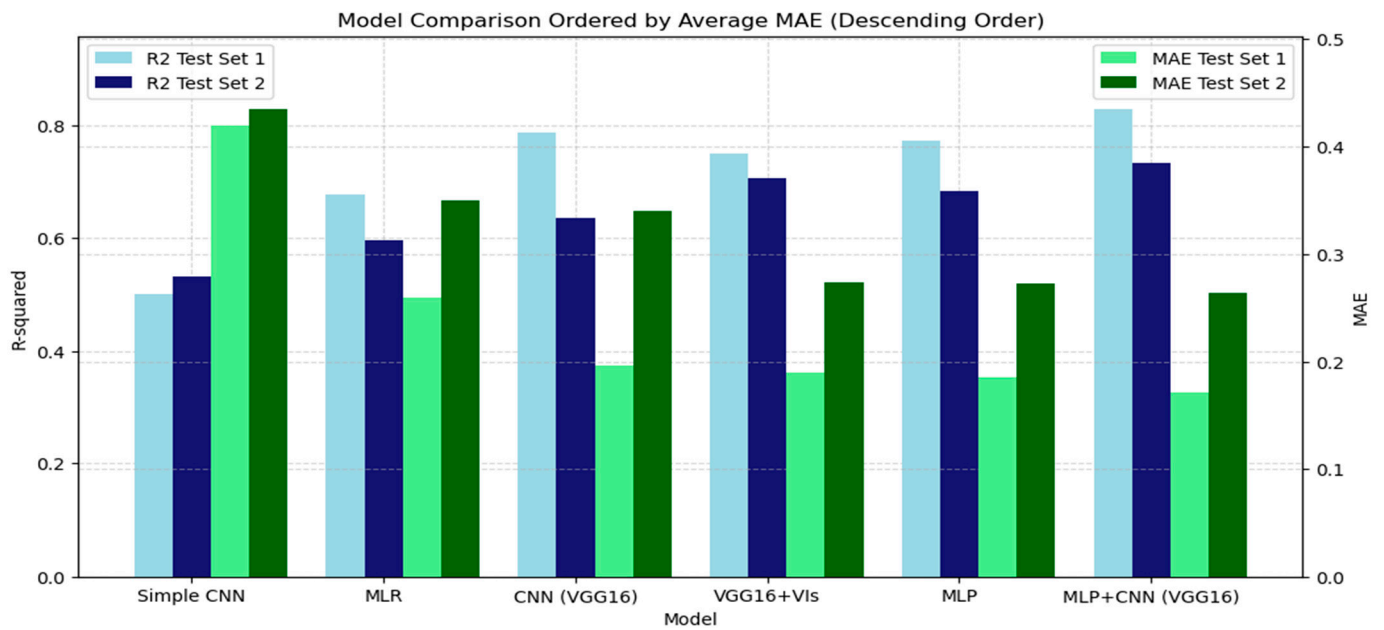
The second approach was based on concatenating, within the same neural net architecture, the VGG16 extracted deep features with the handcrafted computed VIs. To accommodate the large difference of the two feature types in terms of dimensionality, we added two dense layers, with 50 and 20 neurons, respectively, after the last convolution layer. We then added the five VIs either to the output of either the first dense layer or to the output of the second dense layer. These two concatenation schemes are displayed in Figure 2. The optimal final architecture (later referred to as VGG16 + VIs) that gave the best results is obtained when the five DVIs are concatenated to the first dense layer output consisting of 50 neurons.



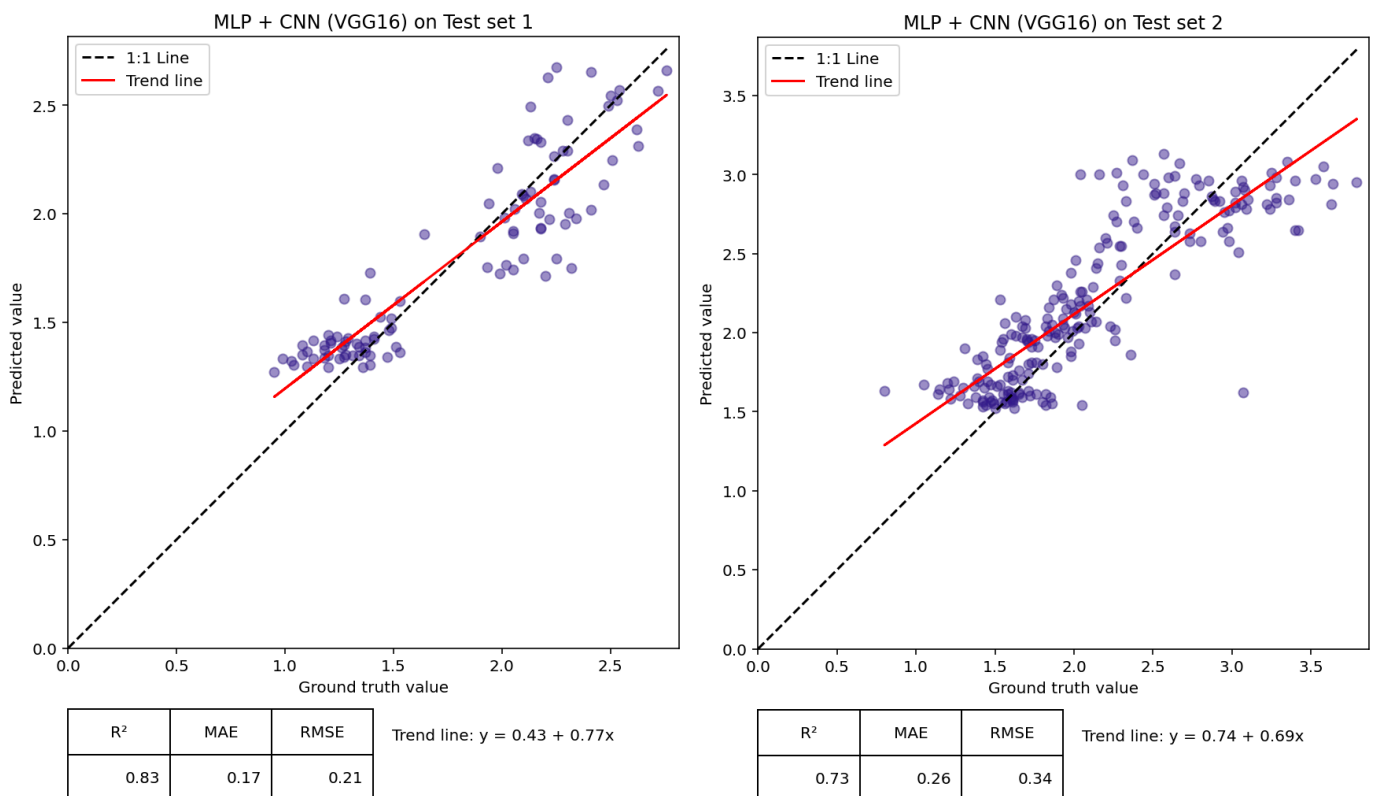
**Figure 2.** Visualization of the model combining the CNN (VGG16) architecture with computed VIs into an additional layer. The combination is later referred to as VGG16 + VIs.

### 3. Results

This work explored and evaluated four models (Sections 2.6.1–2.6.4) and their combinations (Section 2.6.5) on independent Test set 1. Final hyper-parameters of the models can be found in the Supplementary Materials (Table S4 for the MLP, Table S5 and Figure S2 for the Simple CNN, and Table S6 for the pre-trained CNNs). To ensure the repeatability of results, these models were tested and compared on a second independent Test set 2 (as described in Section 2.1.2). The prediction performance on the independent test sets was compared with the other models primarily using standard metrics  $R^2$  and MAE. The results can be seen in Figure 3. Overall, the combination of MLP and CNN (VGG16) was the best model combination when compared with the baseline model (MLR), as it led to a relative improvement of 34.23% for the Test set 1 and 24.75% for the Test set 2 in MAE (Figure 3; scatter plot in Figure 4). For a more comprehensive comparison, detailed metrics are shown in Supplementary Table S2. Detailed descriptive statistics of both ground-truth datasets are in Supplementary Materials Table S7 and Figure S1.



**Figure 3.** Comparative analysis of the prediction potential of four different models on independent Test set 1 (rabi 2021 season) and independent Test set 2 (rabi 2022 season) using performance metrics MAE and R<sup>2</sup>. Models are sorted based on MAE value. Specific values of the metrics are shown in Supplementary Table S2.



**Figure 4.** Prediction performance metrics of the best nitrogen prediction model (MLP + CNN(VGG16)) shown using scatter plots for observed (ground-truth) vs. predicted %N in sorghum. Left side for Test set 1 (rabi 2021–2022), and right side for Test set 2).

## 4. Discussion

In this work, we have assessed the power of advanced analytical model-building techniques for RGB imaging-based phenotyping tasks and discussed its potential value for end-users.

### 4.1. Challenges of Using Image-Based Phenotyping for Precision Applications

Non-invasive image-based plant phenotyping is being used across many research disciplines (e.g., basic plant sciences, genebanks, crop breeding, agronomy, and ecology [45–48]); for each end-use, there are different requirements for technology robustness (e.g., image capture technology, throughput, accuracy, repeatability, generalization, and computation intensity [8,9,49]). Image-based phenotyping of crops e.g., in breeding programs, it is particularly challenging as it requires rapid throughput, high accuracy, repeatability, and time- and cost-efficiency (typically 1000s of relatively similar genotypes in the field conditions, e.g., [45–47,50]).

Despite image-based phenomics being around for decades, the robustness of the methods, particularly for sensors carried by an aerial vehicle (UAV, satellite), is still questioned in terms of repeatability and generalization capacity of trait inference models [8,22,49]. Therefore, apart from the adequate image acquisition protocols (Gattu et al., 2023), it is essential to design a robust trait inference model that deals with residual image variability not tackled during image acquisition [50]. These issues should be on the frontline of interest of the data-research community (e.g., digital responsibility goals, [19]) but are rarely addressed in the literature concerning phenotyping [8,9].

To tackle these issues in plant trait inference models, we used standardized image quality control protocols to acquire images [22]. We tested a range of ML and DL models usually used for similar purposes [10–14]. These ranged from ML methods (MLR and MLP) based on vegetation indices (i.e., EXGR, GCC, GLI, GRD, and RGBVI [37,38]) up to DL methods and ML–DL combinations [51].

Within the DL models, we used a range of existing CNN architectures with light, medium, and large sizes [51] previously deployed for similar tasks [12–14]. These transfer learning schemes were also expected to enhance the model generalization capacity [52] by preventing overfitting on small training datasets [53].

### 4.2. Combining Expert Knowledge and DL Features Improved the Sorghum-N Prediction from RGB Images

In our particular case study, i.e., inferring the sorghum N content from RGB images, we achieved similar levels of prediction accuracies as reported recently in similar remote-sensing studies (typically  $R^2 \sim 0.5\text{--}0.8$  on test set; [4,14,16,17]). We achieved, however, more accurate predictions using VIs in combined model schemes that improved model robustness, in our case, combining the VGG16 with VIs directly in one model or combining VGG16 and MLP models. The highest prediction performance was achieved by combining VGG16 with MLP (VIs)  $R^2 = 0.733$ , and lowest MAE = 0.26 N% on an independent dataset (Test set 2). This points out that combining neural networks (CNNs) with spectral features—VIs (in MLP)—enhanced the model prediction capability on the independent dataset. It is worth noting, nonetheless, that the MLP + CNN (VGG16) ensembling scheme was slightly better compared to the scheme where the VI expert features were injected into the deep neural network architecture directly. This may be explained by the required additional validation set to learn the neural nets weights associated with the combining layer and also by the fact that additional injecting schemes should be investigated in terms of the depth of layers considered for combination and exploring attention strategies.

Overall, we can ascertain that the schemes combining expert knowledge and feature representations from DL architectures are the most effective when developing trait inference models even with a small training dataset.

The size of the utilized dataset was another point of our study (we generated altogether 470 ground truth laboratory N assessments and corresponding crop images). This might

be considered as a medium-sized dataset for any DL modeling task, yet the precise and relevant ground truth observations suitable for modeling are very difficult to generate and the availability of such a dataset is rare (recently, e.g., <https://www.global-wheat.com/> became available (accessed on 10 June 2024)). Therefore, providing the dataset to the community to test their own models (see Data Availability section for annotated dataset) is another important contribution of ours.

In addition, after assessing the performance of the models on the first test set, we showed how these models behaved on the second set, collected much later. This is not regularly considered in the state-of-the-art technology that mostly builds only the test subset with the same data distribution as the training set collected using the same acquisition protocol and operator. This has two drawbacks. First, the repeated optimization of the models on the training sets may implicitly lead to overfitting on the test set. Second, such a test set does not allow for assessing the generalization capabilities of the models on data acquired from different contexts or different time periods, with the potential for causing concept and data drift. This is particularly important for DL-based architectures whose trustworthiness might be compromised by overfitting to the training dataset [54,55].

#### 4.3. Limitations of the Study and Ways Forward

In our case study, two facts might be considered as limitations: (1) we used only RGB images (similar to [14]), while much of the work pursuing the crop N-related traits estimation was recently completed via RGB-NIR, multi-spectral, and hyper-spectral cameras (e.g., [4,8,14,16,17,56]); (2) we used a small part of the dataset to train the N-prediction model (153 data points) which might be considered constraining, especially for DL which is known to substantially improve with increasing training set sizes, e.g., [14]. This limitation was nonetheless mitigated by resorting to effective transfer learning schemes leveraging DL models pre-trained on huge datasets, such as ImageNet. In the future, we will seek to increase the training data size by leveraging our recent adversarial learning-based data augmentation schemes proposed in other tasks [17,57].

We argue that even though remote-evaluation of N-related crop traits is moving towards use of multi-spectral cameras to strengthen the physical foundations of predictions [16,17], we showed similar prediction accuracy is achievable with the simpler RGB camera, particularly when coupled with hybrid and ensembled DL architectures infused with expert knowledge. For example, [16,17] reported a prediction accuracy of N (or N-related cereal traits) with  $R^2 \sim 0.45\text{--}0.65$  with multi-spectral VIs which is well in range of our accuracies using RGB-spectra-based VIs (Supplementary Table S2). Also, the highest achieved prediction accuracies on the test set using RGB ( $R^2 = 0.829$ ) images were similar to the reported accuracies for similar tasks using multi-spectral images and a broad range of data modeling approaches ( $R^2 \sim 0.6\text{--}0.8$ , in [4,16,17]). We also expect these algorithms might serve a wider range of end-users and end-uses as these might be built using simpler imaging methods with less data. Therefore, we will continue exploring these types of models in parallel to others as these might become an effective and practical bridge to close the frequently argued gap in image-interpretation transferability and accuracy [16,17].

The paper's knowledge and results will help us accelerate the implementation of precise, trustworthy UAV phenotyping, particularly in crop breeding. In the future, we intend to further explore the potential of our combining schemes using more complex sensors and in parallel to other data modeling techniques by considering data augmentation techniques to make training of CNNs more effective. To this end, we will leverage our recent work on generative adversarial network (GAN)-based data augmentation techniques that substantially improved our work on predicting the soil moisture dissipation rates from aerial images [58–60].

## 5. Conclusions

Automation of crop characterization tasks (phenotyping) using non-destructive, sensor-based technologies is on its way to becoming routine. Nevertheless, many processes are still



being questioned, e.g., repeatability, robustness, and accuracy of data gathering intensity, modeling techniques or cost-efficiency, just to name a few. Our case study explored the potential of relatively simple imaging technology (UAV-carried RGB camera) and a small dataset for precise crop phenotyping. For our case study, we focused on the sorghum N-content, a critical trait needed, e.g., for breeding dual-purpose sorghum. We collected 470 ground-truth measurements of biomass N content and the corresponding RGB images in two independent seasons (rabi 2021, 2022). We used the first season data (rabi 2021) for building the range of prediction models (ML, DL, and their combinations) and tested their robustness on an independent dataset (rabi 2022).

When fused with expert knowledge (VIs or MLP built upon VIs), DL-based prediction models were generally more accurate compared to the individual models (particularly a combination of MLP + CNN-VGG16;  $R^2 = 0.73$ , MAE = 0.26 %N on independent dataset). Such accuracy already justifies the basic applications, e.g., in crop breeding. Moreover, it was achieved with a small training dataset (153 datapoints). Thus, we demonstrated that the relatively simple proximal sensing methods and a small amount of ground truth observations might already have an added value for particular end-uses when combined with carefully crafted novel model-building methods. This knowledge will help us to accelerate implementation of precise, trustworthy UAV phenotyping, particularly in crop breeding. In the future, we intend to further explore the potential of our combining schemes using more complex sensors and in parallel to other data modeling techniques.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/agriculture14101682/s1>, Table S1. Settings that were used to capture the RGB images in the field trials for UAV and camera with optimal parameters ensuring good quality of raw images for each flight. Table S2. Comparative table of the prediction potential of four different models (best results) on independent Test set 1 (rabi 2021 season) and independent Test set 2 (rabi 2022 season) using performance metrics MAE and  $R^2$ . Models are sorted based on MAE value. Table S3. A comparative table of the prediction potential of all tested models based on the Test set 1. Table S4. Final hyperparameter values in MLP model for N prediction in sorghum deduced by hyperparameter tuning. Table S5. Hyperparameters used for training the architecture that uses the simple CNN (Section 2.6.3). Table S6. Hyperparameters used for training the architecture that uses the pre-trained models built in Keras library. Table S7. Basic descriptive statistics of both data sets. Figure S1. Basic statistical distribution of both data sets. Figure S2. The architecture of CNN from scratch. The first is the input - the image of size 500x500x3. Next are the blocks of feature extraction, which consists of convolutional and pooling layers. Finally, the prediction of nitrogen is the output layer.

**Author Contributions:** Conceptualization, J.K., H.Q., H.B., A.J., M.C. and K.P.; Data curation, S.P., S.C., K.P. and R.B.; Formal analysis, S.C.; Funding acquisition, K.A., J.V., H.Q., M.S. and A.J.; Investigation, M.C.; Methodology, M.A.E.-Y., J.M., J.K., J.V., H.Q. and M.S.; Project administration, J.V. and M.S.; Resources, S.C., K.A. and S.M.; Software, H.H. and S.P.; Supervision, M.A.E.-Y., J.M., J.K. and H.B.; Writing—original draft, H.H., S.P., M.A.E.-Y. and J.M.; Writing—review & editing, H.H., S.P., M.A.E.-Y., J.M. and J.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** Funding for this research work was sourced from ongoing projects funded by the Technology Innovation Hub on Autonomous Navigation and Data Acquisition Systems (TiHAN) of the Indian Institute of Technology Hyderabad (IIT-H) under the National Mission on Interdisciplinary Cyber-Physical Systems (NMI-CPS) of the Department of Science & Technology (DST), Govt of India (Grant number TiHAN-IITH/03/2021-22/32(4) “Leveraging the UAV-based technology for Crop Residue: Important Resource for Crop-Livestock Farming Community” and grant number TiHAN-IITH/03/2022/194 “Technology Transfer: UAV based automated Data processing pipeline for Optimized Aerial Crop Monitoring”); Early Career Research Award from the Department of Science and Technology (DST), Government of India; and the internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. 2023B0005 (Oborově zaměřené datové modely pro podporu iniciativy Open Science a principu FAIR). This research was also supported by the Ministry of Agriculture of the Czech Republic, grant number

QK23020058 (precision agriculture and digitization in the Czech Republic). The results and knowledge included herein have been obtained owing to support from the sabbatical granted to Suchitra Patil by K.J. Somaiya College of Engineering, Mumbai for her PhD research; Eiffel scholarship for Hajar's Ph.D. thesis and the Barrande Fellowship funding for Hajar's internship at Czech University of Life Sciences, Prague.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used in this study (RGB images and ground-truth) are available under doi: 10.5281/zenodo.13733946; <https://zenodo.org/doi/10.5281/zenodo.13733946> (accessed on 9 September 2024).

**Acknowledgments:** This work was supported extensively by the research team of Crop Physiology and Modelling (<https://gems-icrisat.site/> (accessed on 2 September 2024)). A special note of thanks to Amrutha Kumar, Premalatha T., Mallesh R., and Jayalakshmi Ambhati for their assistance in the activities related to field management, ground truth measurements, and crop quality assessment. We would like to thank Prem Kumar from Marut Drones for their support in drone flights. We would also like to thank Padam Kumar from International Livestock Research Institute (ILRI) and ILRI staff for their assistance in crop quality assessment.

**Conflicts of Interest:** Author Srinivas Mamidi was employed by the company Marut Dronetech Private Limited. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Tardieu, F.; Cabrera-Bosquet, L.; Pridmore, T.; Bennett, M. Plant Phenomics, From Sensors to Knowledge. *Curr. Biol.* **2017**, *27*, R770–R783. [[CrossRef](#)] [[PubMed](#)]
2. Li, L.; Zhang, Q.; Huang, D. A Review of Imaging Techniques for Plant Phenotyping. *Sensors* **2014**, *14*, 20078–20111. [[CrossRef](#)]
3. Demidchik, V.V.; Shashko, A.Y.; Bandarenka, U.Y.; Smolikova, G.N.; Przhevalskaya, D.A.; Charnysh, M.A.; Pozhvanov, G.A.; Barkosvkiy, A.V.; Smolich, I.I.; Sokolik, A.I.; et al. Plant Phenomics: Fundamental Bases, Software and Hardware Platforms, and Machine Learning. *Russ. J. Plant Physiol.* **2020**, *67*, 397–412. [[CrossRef](#)]
4. Zhu, R.; Sun, K.; Yan, Z.; Yan, X.; Yu, J.; Shi, J.; Hu, Z.; Jiang, H.; Xin, D.; Zhang, Z.; et al. Analysing the Phenotype Development of Soybean Plants Using Low-Cost 3D Reconstruction. *Sci. Rep.* **2020**, *10*, 7055. [[CrossRef](#)]
5. Das Choudhury, S.; Maturu, S.; Samal, A.; Stoerger, V.; Awada, T. Leveraging Image Analysis to Compute 3D Plant Phenotypes Based on Voxel-Grid Plant Reconstruction. *Front. Plant Sci.* **2020**, *11*, 521431. [[CrossRef](#)]
6. Das Choudhury, S.; Bashyam, S.; Qiu, Y.; Samal, A.; Awada, T. Holistic and Component Plant Phenotyping Using Temporal Image Sequence. *Plant Methods* **2018**, *14*, 35. [[CrossRef](#)] [[PubMed](#)]
7. Huang, S.; Tang, L.; Hupy, J.P.; Wang, Y.; Shao, G. A Commentary Review on the Use of Normalized Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing. *J. For. Res.* **2021**, *32*, 1–6. [[CrossRef](#)]
8. Zeng, Y.; Hao, D.; Huete, A.; Dechant, B.; Berry, J.; Chen, J.M.; Joiner, J.; Frankenberg, C.; Bond-Lamberty, B.; Ryu, Y.; et al. Optical Vegetation Indices for Monitoring Terrestrial Ecosystems Globally. *Nat. Rev. Earth Environ.* **2022**, *3*, 477–493. [[CrossRef](#)]
9. Gracia-Romero, A.; Kefauver, S.C.; Fernandez-Gallego, J.A.; Vergara-Díaz, O.; Nieto-Taladriz, M.T.; Araus, J.L. UAV and Ground Image-Based Phenotyping: A Proof of Concept with Durum Wheat. *Remote Sens.* **2019**, *11*, 1244. [[CrossRef](#)]
10. Niu, Q.; Feng, H.; Li, C.; Yang, G.; Fu, Y.; Li, Z.; Pei, H. Estimation of Leaf Nitrogen Concentration of Winter Wheat Using Uav-Based RGB Imagery. *IFIP Adv. Inf. Commun. Technol.* **2019**, *546*, 139–153. [[CrossRef](#)]
11. Shi, P.; Wang, Y.; Xu, J.; Zhao, Y.; Yang, B.; Yuan, Z.; Sun, Q. Rice Nitrogen Nutrition Estimation with RGB Images and Machine Learning Methods. *Comput. Electron. Agric.* **2021**, *180*, 105860. [[CrossRef](#)]
12. Qiu, Z.; Ma, F.; Li, Z.; Xu, X.; Ge, H.; Du, C. Estimation of Nitrogen Nutrition Index in Rice from UAV RGB Images Coupled with Machine Learning Algorithms. *Comput. Electron. Agric.* **2021**, *189*, 106421. [[CrossRef](#)]
13. Kou, J.; Duan, L.; Yin, C.; Ma, L.; Chen, X.; Gao, P.; Lv, X. Predicting Leaf Nitrogen Content in Cotton with UAV RGB Images. *Sustainability* **2022**, *14*, 9259. [[CrossRef](#)]
14. Alves Oliveira, R.; Marcato Junior, J.; Soares Costa, C.; Näsi, R.; Koivumäki, N.; Niemeläinen, O.; Kaivosoja, J.; Nyholm, L.; Pistori, H.; Honkavaara, E. Silage Grass Sward Nitrogen Concentration and Dry Matter Yield Estimation Using Deep Regression and RGB Images Captured by UAV. *Agronomy* **2022**, *12*, 1352. [[CrossRef](#)]
15. Jabir, B.; El Moutaouakil, K.; Falih, N. Developing an Efficient System with Mask R-CNN for Agricultural Applications. *Agris On-line Pap. Econ. Inform.* **2023**, *15*, 61–72. [[CrossRef](#)]
16. Sun, J.; Wang, L.; Shi, S.; Li, Z.; Yang, J.; Gong, W.; Wang, S.; Tagesson, T. Leaf Pigment Retrieval Using the PROSAIL Model: Influence of Uncertainty in Prior Canopy-Structure Information. *Crop J.* **2022**, *10*, 1251–1263. [[CrossRef](#)]

17. Li, J.; Wijewardane, N.K.; Ge, Y.; Shi, Y. Improved Chlorophyll and Water Content Estimations at Leaf Level with a Hybrid Radiative Transfer and Machine Learning Model. *Comput. Electron. Agric.* **2023**, *206*. [CrossRef]
18. Lu, J.; Liu, A.; Dong, F.; Gu, F.; Gama, J.; Zhang, G. Learning under Concept Drift: A Review. *IEEE Trans. Knowl. Data Eng.* **2019**, *31*, 2346–2363. [CrossRef]
19. Meier, J.J.; Hermsen, K.; Bauer, J.; Eskofier, B.M. Digital Responsibility Goals—A Framework for a Human-Centered Sustainable Digital Economy with a Focus on Trusted Digital Solutions. *Stud. Health Technol. Inform.* **2022**, *293*, 250–259. [CrossRef]
20. Indian Council Of Agricultural Research (ICAR). *Handbook of Agriculture*, 6th Revised Edition; Indian Council of Agricultural Research: New Delhi, India, 2012; ISBN 81-7164-050-8.
21. Upadhyaya, H.D.; Reddy, K.N.; Irshad Ahmed, M.; Dronavalli, N.; Gowda, C.L.L. Latitudinal Variation and Distribution of Photoperiod and Temperature Sensitivity for Flowering in the World Collection of Pearl Millet Germplasm at ICRISAT Genebank. *Plant Genet. Resour.* **2012**, *10*, 59–69. [CrossRef]
22. Priyanka, G.; Choudhary, S.; Anbazhagan, K.; Naresh, D.; Baddam, R.; Jarolimek, J.; Parnandi, Y.; Rajalakshmi, P.; Kholova, J. A Step towards Inter-Operable Unmanned Aerial Vehicles (UAV) Based Phenotyping; A Case Study Demonstrating a Rapid, Quantitative Approach to Standardize Image Acquisition and Check Quality of Acquired Images. *ISPRS Open J. Photogramm. Remote Sens.* **2023**, *9*, 100042. [CrossRef]
23. Luisa Buchailot, M.; Baret, F.; Zaman-Allah, M.A.; Cairns, J.; Klassen, S.; Chapman, S.; Potgieter, A.; Poland, J. Basic Standard Operating Procedures for UAV Phenotyping. Available online: [https://excellenceinbreeding.org/sites/default/files/manual/EiB\\_M4\\_%20SOP-UAV-Phenotyping-12-10-20.pdf](https://excellenceinbreeding.org/sites/default/files/manual/EiB_M4_%20SOP-UAV-Phenotyping-12-10-20.pdf) (accessed on 4 June 2024).
24. Zenmuse X5S—DJI. Available online: <https://www.dji.com/cz/zenmuse-x5s> (accessed on 9 May 2024).
25. Cemotec Laboratory Grinder with No Loss of Moisture. Available online: <https://www.fossanalytics.com/en-in/products/cm-290-cemotec> (accessed on 9 May 2024).
26. Ejaz, I.; He, S.; Li, W.; Hu, N.; Tang, C.; Li, S.; Li, M.; Diallo, B.; Xie, G.; Yu, K. Sorghum Grains Grading for Food, Feed, and Fuel Using NIR Spectroscopy. *Front. Plant Sci.* **2021**, *12*, 720022. [CrossRef] [PubMed]
27. FAO; WHO. *Codex Alimentarius Commission Procedural Manual*; Food and Agriculture Organization (FAO): Rome, Italy, 2023. [CrossRef]
28. Blümmel, M.; Deshpande, S.; Kholova, J.; Vadez, V. Introgression of Staygreen QLT's for Concomitant Improvement of Food and Fodder Traits in Sorghum Bicolor. *Field Crops Res.* **2015**, *180*, 228–237. [CrossRef]
29. Ramana Reddy, Y.; Ravi, D.; Ramakrishna Reddy, C.; Prasad, K.V.S.V.; Zaidi, P.H.; Vinayan, M.T.; Blümmel, M. A Note on the Correlations between Maize Grain and Maize Stover Quantitative and Qualitative Traits and the Implications for Whole Maize Plant Optimization. *Field Crops Res.* **2013**, *153*, 63–69. [CrossRef]
30. De, K.; Masilamani, V. Fast No-Reference Image Sharpness Measure for Blurred Images in Discrete Cosine Transform Domain. In Proceedings of the 2016 IEEE Students' Technology Symposium, TechSym 2016, Kharagpur, India, 30 September–2 October 2016; pp. 256–261. [CrossRef]
31. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [CrossRef]
32. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely Blind” Image Quality Analyzer. *IEEE Signal Process Lett.* **2013**, *20*, 209–212. [CrossRef]
33. Agisoft Agisoft Metashape: User Manuals. Available online: <https://www.agisoft.com/downloads/user-manuals/> (accessed on 9 May 2024).
34. Lindeberg, T. Scale Invariant Feature Transform. *Scholarpedia* **2012**, *7*, 10491. [CrossRef]
35. Hu, Q.; Luo, J.; Hu, G.; Duan, W.; Zhou, H. 3D Point Cloud Generation Using Incremental Structure-from-Motion. *J. Phys. Conf. Ser.* **2018**, *1087*, 062031. [CrossRef]
36. Welcome to the QGIS Project! Available online: <https://www.qgis.org/en/site/> (accessed on 9 May 2024).
37. Louhaichi, M.; Borman, M.M.; Johnson, D.E. Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. *Geocarto Int.* **2001**, *16*, 65–70. [CrossRef]
38. Bendig, J.; Yu, K.; Aasen, H.; Bolten, A.; Bennertz, S.; Broscheit, J.; Gnyp, M.L.; Bareth, G. Combining UAV-Based Plant Height from Crop Surface Models, Visible, and near Infrared Vegetation Indices for Biomass Monitoring in Barley. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *39*, 79–87. [CrossRef]
39. Barbedo, J.G.A. Detection of Nutrition Deficiencies in Plants Using Proximal Images and Machine Learning: A Review. *Comput. Electron. Agric.* **2019**, *162*, 482–492. [CrossRef]
40. Haykin, S.; York, N.; San, B.; London, F.; Sydney, T.; Singapore, T.; Mexico, M.; Munich, C.; Cape, P.; Hong, T.; et al. *Neural Networks and Learning Machines*, 3rd ed.; Pearson Education: London, UK, 2009.
41. Xu, S.; Chen, L. A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining. In Proceedings of the 5th International Conference on Information Technology and Applications, Cairns, Australia, 23–26 July 2008; pp. 683–686.
42. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep Learning in Agriculture: A Survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
43. Keras Keras Applications. Available online: <https://keras.io/api/applications/> (accessed on 17 May 2024).
44. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Li, F.-F. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [CrossRef]

45. Pasala, R.; Pandey, B.B. Plant Phenomics: High-Throughput Technology for Accelerating Genomics. *J. Biosci.* **2020**, *45*, 111. [[CrossRef](#)]
46. Pieruschka, R.; Schurr, U. Plant Phenotyping: Past, Present, and Future. *Plant Phenomics* **2019**, 2019. [[CrossRef](#)]
47. Danzi, D.; Briglia, N.; Petrozza, A.; Summerer, S.; Povero, G.; Stivaletta, A.; Cellini, F.; Pignone, D.; de Paola, D.; Janni, M. Can High Throughput Phenotyping Help Food Security in the Mediterranean Area? *Front. Plant Sci.* **2019**, *10*, 433452. [[CrossRef](#)]
48. Hu, P.; Guo, W.; Chapman, S.C.; Guo, Y.; Zheng, B. Pixel Size of Aerial Imagery Constrains the Applications of Unmanned Aerial Vehicle in Crop Breeding. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 1–9. [[CrossRef](#)]
49. Guo, Y.; Senthilnath, J.; Wu, W.; Zhang, X.; Zeng, Z.; Huang, H. Radiometric Calibration for Multispectral Camera of Different Imaging Conditions Mounted on a UAV Platform. *Sustainability* **2019**, *11*, 978. [[CrossRef](#)]
50. Kholová, J.; Urban, M.O.; Cock, J.; Arcos, J.; Arnaud, E.; Aytakin, D.; Azevedo, V.; Barnes, A.P.; Ceccarelli, S.; Chavarriaga, P.; et al. In Pursuit of a Better World: Crop Improvement and the CGIAR. *J. Exp. Bot.* **2021**, *72*, 5158–5179. [[CrossRef](#)]
51. Xiang, Q.; Zi, L.; Cong, X.; Wang, Y. Concept Drift Adaptation Methods under the Deep Learning Framework: A Literature Review. *Appl. Sci.* **2023**, *13*, 6515. [[CrossRef](#)]
52. Taye, M.M. Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions. *Computers* **2023**, *12*, 91. [[CrossRef](#)]
53. Li, D.; Zhang, H.R. Improved Regularization and Robustness for Fine-Tuning in Neural Networks. *Adv. Neural Inf. Process Syst.* **2021**, *33*, 27249–27262.
54. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)] [[PubMed](#)]
55. Pound, M.P.; Atkinson, J.A.; Townsend, A.J.; Wilson, M.H.; Griffiths, M.; Jackson, A.S.; Bulat, A.; Tzimiropoulos, G.; Wells, D.M.; Murchie, E.H.; et al. Deep Machine Learning Provides State-of-the-Art Performance in Image-Based Plant Phenotyping. *Gigascience* **2017**, *6*, gix083. [[CrossRef](#)] [[PubMed](#)]
56. Patil, S.M.; Choudhary, S.; Kholova, J.; Chandramouli, M.; Jagarlapudi, A. Applications of UAVs: Image-Based Plant Phenotyping. In *Digital Agriculture: A Solution for Sustainable Food and Nutritional Security*; Springer International Publishing: Cham, Switzerland, 2024; pp. 341–367. [[CrossRef](#)]
57. Qin, H.; Xi, H.; Li, Y.; El-Yacoubi, M.A.; Wang, J.; Gao, X. Adversarial Learning-Based Data Augmentation for Palm-Vein Identification. *IEEE Trans. Circuits Syst. Video Technol.* **2023**, *34*, 4325–4341. [[CrossRef](#)]
58. Hammouch, H.; El-Yacoubi, M.; Qin, H.; Berbia, H.; Chikhaoui, M. Controlling the Quality of GAN-Based Generated Images for Predictions Tasks. In Proceedings of the International Conference on Pattern Recognition and Artificial Intelligence, ICPRAI 2022, Paris, France, 1–3 June 2022; Volume 13363, pp. 121–133. [[CrossRef](#)]
59. Hammouch, H.; El-Yacoubi, M.; Qin, H.; Berrahou, A.; Berbia, H.; Chikhaoui, M. A Two-Stage Deep Convolutional Generative Adversarial Network-Based Data Augmentation Scheme for Agriculture Image Regression Tasks. In Proceedings of the 2021 International Conference on Cyber-Physical Social Intelligence, ICCSI 2021, Beijing, China, 18–20 December 2021. [[CrossRef](#)]
60. Hammouch, H.; Mohapatra, S.; El-Yacoubi, M.; Qin, H.; Berbia, H.; Mader, P.; Chikhaoui, M. GANSet—Generating Annotated Datasets Using Generative Adversarial Networks. In Proceedings of the International Conference on Cyber-Physical Social Intelligence, ICCSI 2022, Nanjing, China, 18–21 November 2022; pp. 615–620. [[CrossRef](#)]

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