

Improving soybean yield prediction by integrating UAV nadir and cross-circling oblique imaging



Guangyao Sun^a, Yong Zhang^b, Haochong Chen^c, Lei Wang^b, Mingxue Li^b, Xuhong Sun^b, Shuaipeng Fei^c, Shunfu Xiao^c, Long Yan^d, Yinghui Li^e, Yun Xu^{a,*}, Lijuan Qiu^{e,*}, Yuntao Ma^{c,*}

^a College of Information and Electrical Engineering, China Agricultural University, East Campus, Beijing, China

^b Keshan Branch of Heilongjiang Academy of Agricultural Sciences, Qiqihar, China

^c College of Land Science and Technology, China Agricultural University, Beijing 100193, China

^d Institute of Cereal and Oil Crops, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang, Hebei 050035, China

^e State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

ARTICLE INFO

Keywords:

Data fusion
Unmanned aerial vehicle
Feature dimensionality reduction
Ensemble learning
Shapley value

ABSTRACT

High-throughput estimation of soybean yield using unmanned aerial vehicle (UAV) imagery can help improve the efficiency of soybean breeding. Previous studies have mainly focused on the extraction of vegetation indices and texture features from two-dimensional(2D) orthophotos to construct empirical models of yield, lacking spatial structure information of crops. Therefore, UAV cross-circling oblique (CCO) photography combined with SfM-MVS algorithm was used to reconstruct three-dimensional(3D) soybean canopy structure. Then canopy 3D related phenotypic features are extracted and combined with features from RGB nadir and multispectral images to analyze the capability of different modal data fusion on soybean yield prediction. In addition, Shapley value was used to evaluate the importance of features across different machine learning models. Based on the Shapley value, a bagging-stacking ensemble learning framework was developed using **Lasso**, **Random Forest (RF)**, **Ridge Regression (RR)**, and **XGBoost** as base learners for yield prediction. The performance of the traditional stacking method was evaluated and compared with weighted average methods such as Bayesian Model Averaging (BMA) and Entropy Weighted Average (EWA) as meta-learners in the ensemble framework. The results show that 3D canopy structure of soybean can be obtained from UAV CCO photography. The inclusion of 3D structural features can improve the accuracy of yield estimation. Among different modal data combinations, the highest estimation accuracy was achieved when combining RGB nadir features with CCO 3D features. The performance of the above base learners was improved by 8.8%, 3.5%, 7.1%, and 8.0% respectively when using the Shapley value method. The accuracy of yield prediction applied on independent dataset of year 2023 was further calculated by using the bagging-stacking ensemble learning method. When using BMA as the meta-learner, the best performance is obtained with an R^2 of 0.7. Therefore, UAV CCO photography with SfM-MVS algorithm provides a new approach to obtain high-quality point clouds of the crop canopy at low cost. UAV multimodal data combined with ensemble learning models allow accurate estimation of yield prediction in breeding materials of soybeans at the plot scale.

1. Introduction

Soybean has extensive applications and values in food, feed, industry, and medicine (Amol, 2022). Under limited agricultural resources, the breeding of high-yielding, dense and stress-tolerant soybean varieties can significantly increase food production to meet the growing global food demand (Guo et al., 2022). The growth condition of farmland relies on labor-intensive experiments and observations

traditionally to assess their yield potentials. Utilizing remote sensing methods to predict crop yield before harvest allows for a rapid and accurate assessment of crop growth performance. This is of significant importance for enhancing breeding efficiency, standardizing agricultural management, and optimizing resource allocation. Therefore, rapidly obtaining crop phenotypic information is of great significance for global food security and sustainable agricultural development (Tardieu et al., 2017).

* Corresponding authors.

E-mail addresses: sdxuyun@163.com (Y. Xu), qulijuan@caas.cn (L. Qiu), yuntao.ma@cau.edu.cn (Y. Ma).

In recent years, unmanned aerial vehicle (UAV) imaging technology has been widely used in crop phenotyping at the plot scale, due to its convenience, flexibility, low cost, and high resolution (Maes and Steppe, 2019; Feng et al., 2021). UAV can be equipped with various sensors such as RGB, multispectral, thermal infrared, and LiDAR (Wang et al., 2022) to estimate crop phenotypic traits. Structural, textural (Mardanisamani et al., 2019; Yue et al., 2019), spectral, and thermal features extracted from multi-source sensors have been widely used for estimation of crop aboveground biomass (Ogawa et al., 2021; Che et al., 2022), yield (Fei et al., 2023b), nitrogen content (Ding et al., 2022), as well as monitoring of lodging (Shu et al., 2023) and disease (Cao et al., 2018) at the plot scale.

Three-dimensional(3D) point clouds can provide more comprehensive information about the 3D structure of crops, such as canopy volume, plant height, canopy morphology, compared to two-dimensional (2D) images that suffering from scale inaccuracies and occlusion issues (Blackman and Yuan, 2020). LiDAR sensors are commonly used for point cloud acquisition (Cai et al., 2022), which is expensive and difficult to be widespread usage in agriculture. The combination of Structure from Motion (SfM) and Multi-View Stereo (MVS) methods (SfM-MVS) enables cost-effective acquisition of canopy point clouds. However, it is challenging to obtain an adequate number of viewpoints in field environments on the ground due to terrain limitations and crop distribution.

UAV mounted on different sensors can capture images of crops from various angles and heights with flexibility and maneuverability, allowing them to fly over obstacles in fields to adapt to different imaging requirements. UAV cross-circling oblique (CCO) uses tilted on-board lens at a specified angle from the vertical to capture omni-directional photography from multiple angles and has been applied in the fields of geographic mapping, architectural design, urban planning (Zhang and Zhu, 2023). UAV-mounted lens usually shoots the ground at a vertical angle in traditional aerial photography. Therefore, the use of UAVs with low-cost RGB cameras by CCO photography of crops and 3D reconstruction based on SfM algorithms can be used to obtain canopy point clouds at plot-scale in a cost-effective manner (Xiao et al., 2023a).

The fusion of multimodal data can improve the reliability and accuracy of data (Maimaitijiang et al., 2020; Shu et al., 2022a) and enhance the expression of data. Collection of UAV image at different crop growth stages can monitor crop growth more comprehensively and improve the estimation performance of crop phenotypic information (Sun et al., 2022b; Wang et al., 2023). Under multi-temporal and multi-modal conditions, the dimensionality of phenotypic features increases dramatically. A large amount of redundant and irrelevant information exist in the feature dataset (Zhang et al., 2021), which seriously affects the computational efficiency and model performance of phenotypic traits estimation. Therefore, effective feature selection tools are needed to minimize feature dimensionality while ensuring the accuracy of phenotypic trait estimation (Liu and An, 2020; Fei et al., 2022). Furthermore, the contribution of the same feature can significantly vary among different machine learning models. In ensemble learning, multiple heterogeneous weak learners exist. If the same dimensionally reduced features are used, it becomes challenging to meet the requirements of all weak learners (Tasci et al., 2022). Therefore, individual feature selection should be conducted for each model to maximize the performance of each weak learner (Chen et al., 2022).

Machine learning methods have shown excellent performance in phenotypic traits estimation through the fusion of multi-source data. Common machine learning algorithms include Random Forest (RF), Multiple Linear Regression (MLR), and Deep Neural Networks (DNN) (Fu et al., 2019; Jin et al., 2020). However, individual models are significantly affected by noise points and outliers, and are prone to overfitting when the training samples are limited (Pal, 2007). Ensemble learning, which combines the predictions of multiple weak learners using different methods, typically exhibits better performance and generalization capabilities (Wang et al., 2020a). Common ensemble learning methods include Bagging (bootstrap aggregating) and Stacking

(Feng et al., 2020). The core idea of Bagging is to repeatedly and randomly sample from the original training dataset to create multiple training subsets with replacement. Each sampled subset is used to train a model, and the final prediction results of multiple base learners are averaged to reduce variance. (Boukir et al., 2021). Stacking integrates the predictions of heterogeneous learners from the first layer using a meta-learner, and reduces the overfitting risk of the individual models (Ju et al., 2018; Wang et al., 2020b). Machine learning regressors such as MLR and RF can be used as meta-learners in the stacking process, as well as weighted averaging methods such as Bayesian Model Averaging (BMA) and Entropy Weighted Averaging (EWA) (Malik et al., 2021). To our knowledge, few research are reported on using multi-modal data combined with bagging and stacking to predict the soybeans yields.

Therefore, the main objectives are as follows: (1) To apply the UAV CCO photography technique combined with the SfM-MVS algorithm to obtain point clouds of soybean canopies. (2) To evaluate the effectiveness of multi-sensor data fusion using UAV and the Shapley value method for the feature dimension reduction in machine learning models. (3) Combining Shapley values, bagging and stacking to develop an ensemble learning framework after feature selection optimization to improve yield prediction accuracy with independent dataset.

2. Materials and methods

2.1. Experimental location and Setup

The experimental was performed in two consecutive years of 2022 and 2023 at Keshan Branch of Heilongjiang Academy of Agricultural Sciences, Qiqihar City, China (48°01'21.7"N 125°51'06.5"E) (Fig. 1). Keshan County has a continental monsoon climate with significant monsoon influences. It has a rainy and hot season with abundant sunshine throughout the year. The average temperature during the soybean growing season is 16°C. The effective accumulated temperature is approximately 2400 degree-days, and the frost-free period lasts for about 122 days. Precipitation is concentrated in the months of June, July, and August, with an average precipitation of around 380 mm during the soybean growing season. Keshan County is in the typical black soil region of Heilongjiang Province, known for its fertile land and good physical and chemical properties, making it an important grain-producing area.

The experiment in 2022 and 2023 involved 263 and 251 soybean varieties with two replications for each variety and each year. Most of these varieties originated in northeastern China, with 26 percent of the varieties being round-leaved and the rest lanceolate. Ninety-two percent of the varieties have a fertility of about 110 days, and the rest are between 115 and 120 days. There were three rows of plants within each plot. The length of each row was 2 m, with a spacing of 0.2 m between rows. All plots were harvested on October 10th, 2022 and October 10th, 2023, more than 130 days after emergence, to make sure all the varieties reaching their maturity. Field irrigation, fertilization and pest management were carried out according to the optimal local levels. After harvest, all soybean grains from each plot were cleaned to remove impurities and weighed in kg/ha.

2.2. Data acquisition

The UAV images are collected on July 1st, July 18th, August 20th, and September 7th in 2022. These dates correspond to 7, 30, 47, 81, and 99 days after emergence of the soybean plants. In 2023, flights were conducted at four identical times after emergence.

2.2.1. RGB nadir image acquisition and preprocessing

The RGB nadir images were captured by a DJI Phantom 4RTK UAV equipped with an FC6310R camera, which has an effective pixel count of 20 million and a resolution of 5472 × 3648. The UAV was flown at an altitude of 30 m, with a forward and side overlap rate of 80%. The

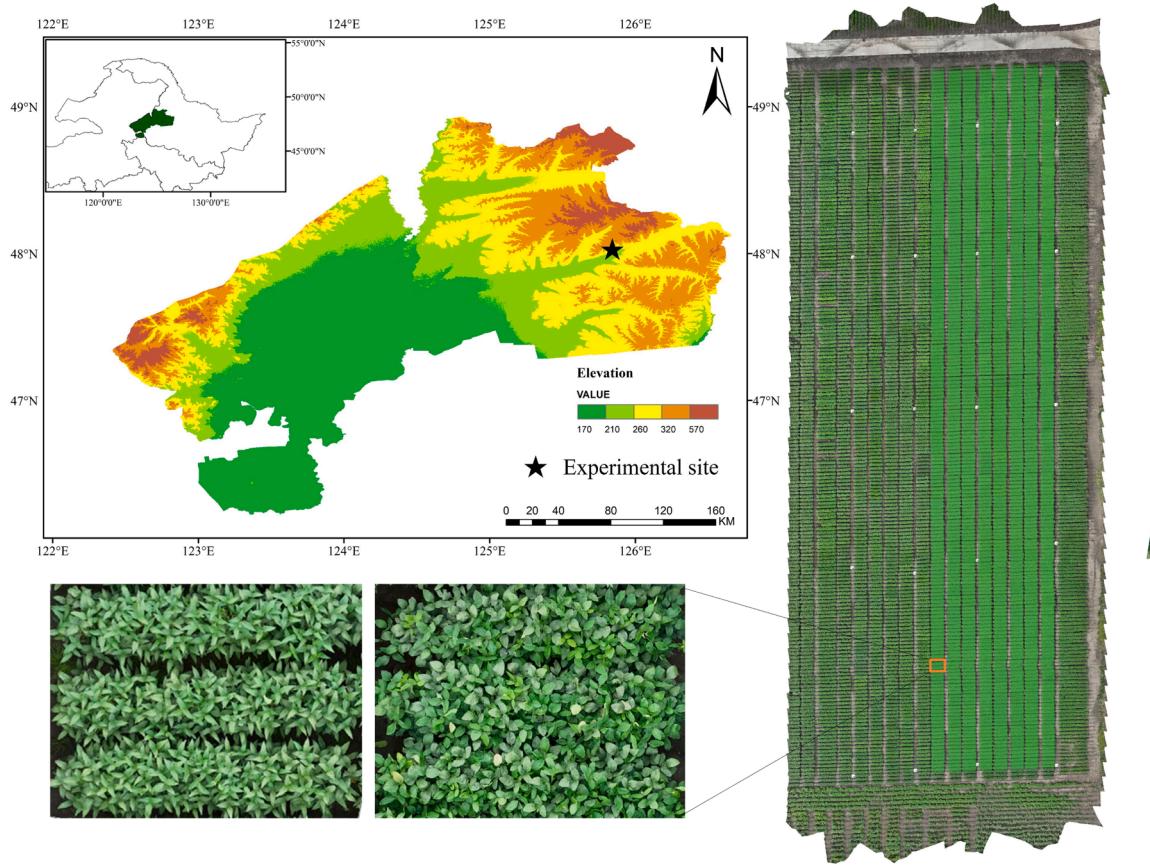


Fig. 1. Location of the experimental site and soybean plot organization.

subsequent image stitching was performed using Agisoft PhotoScan Professional (Agisoft LLC, St. Petersburg, Russia). After aligning the photos and generating a dense point cloud, a digital elevation model and a digital orthophoto image were generated.

2.2.2. Multispectral images acquisition and preprocessing

The multispectral images were captured using a Parrot Sequoia multispectral camera installed on a Parrot Bluegrass UAV from Parrot company in Paris, France. The images have a resolution of 1290×960 pixels. Four spectral bands were simultaneously collected, including green (centered at 550 nm with a bandwidth of 40 nm), red (centered at 660 nm with a bandwidth of 40 nm), red-edge (centered at 735 nm with a bandwidth of 10 nm), and near-infrared (centered at 790 nm with a bandwidth of 40 nm). The UAV flight was set at an altitude of 30 m, with a forward and side overlap rate of 80%. The spectral data collection and radiation correction follow the standard workflows (Lobos et al., 2014; Hernandez et al., 2015). To avoid differences in light intensity at different flight times, a field reference panel is used for translating the DN value into reflectance before and after each flight. The subsequent image stitching was performed using Pix4D mapper (Pix4D, Lausanne, Switzerland) to generate a digital orthophoto image with reflectance.

2.2.3. CCO data acquisition and preprocessing

Raw images of the canopy point cloud were acquired by DJI Phantom 4RTK via CCO photography. The UAV's flight altitude was set to 5 m above the canopy. The CCO flight path was set with circles by a certain number of waypoints along the path for image capture. The radius of the circle was calculated using the Pythagorean theorem based on the right-angled triangle formed by the center point of the circle, the projected ground point of the circle and the UAV. When the camera's tilt angle was set to -45° from horizontal, the radius of the circle was equal to the

flight height. To ensure the reconstruction accuracy, a certain amount of overlap was maintained between the circular paths. After several attempts, it was found that setting 20 waypoints per circle and a 50% overlap between paths yielded good reconstruction accuracy. Image acquisition was conducted on days with low wind conditions. Before each flight, 20 rulers with a known height (100 mm) and 20 whiteboards with a known length (300 mm) were placed evenly in the field to verify the vertical and horizontal reconstruction accuracy of the point cloud. The stitching of the multi-view images captured along the CCO path was performed using Agisoft PhotoScan Professional, and point cloud of the soybean canopy was reconstructed using the Multi-View Stereo and Structure from Motion (MVS-SFM) algorithm. Preprocessing of the point cloud was conducted, including soil removal, noise reduction, and point cloud segmentation, to obtain point clouds for each plot. The workflow of the CCO path photography and point cloud preprocessing is shown in Fig. 2.

2.3. Feature extraction

2.3.1. Extraction of vegetation indices

Studies showed that spectral indices calculated based on DN values of RGB images can be used to estimate crop phenotypic traits (Bendig et al., 2015). Nadir images are commonly used for vegetation index extraction which only provide planar (2D) information and cannot capture the reflectance information of vegetation in the vertical distribution. In contrast, 3D point clouds contain complete color information of the canopy. This is particularly important for detecting small-scale vegetation changes or analyzing uneven crop growth conditions. Fifteen vegetation indices were extracted based on RGB nadir images and point cloud (Table 1), respectively, and 18 vegetation indices based on multispectral orthophoto (Table 2). These vegetation indices were

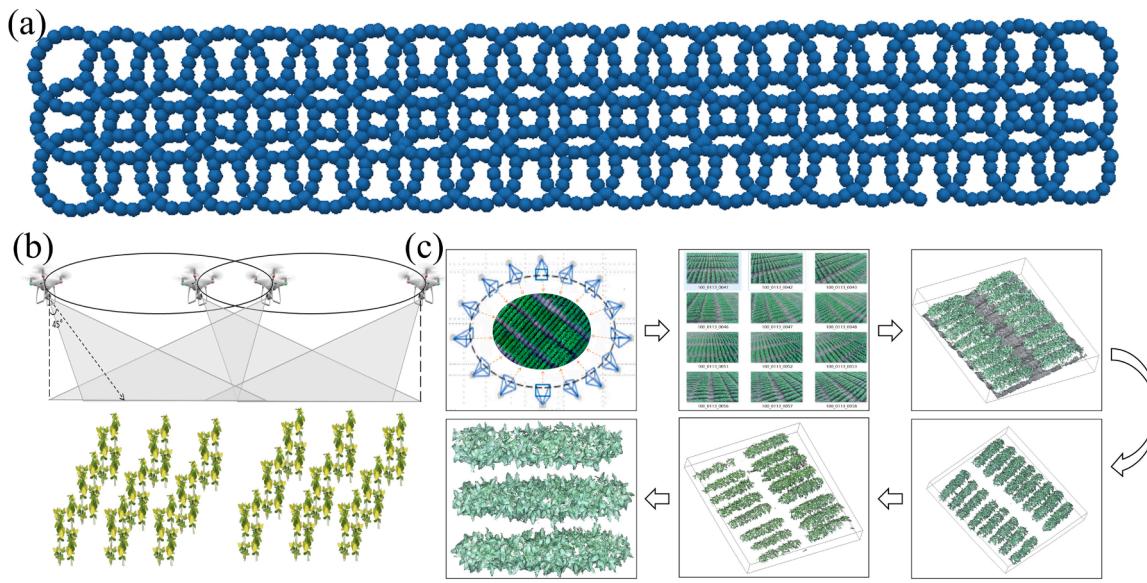


Fig. 2. (a) The aerial view of the UAV's CCO path in the experimental area, with 50% overlap between the circles. (b) Side view of a CCO route consisting of two single circles with a 50% inter-circle overlap. (c) 3D reconstruction and preprocessing workflow of the point cloud.

Table 1
Calculation formula of vegetation index with RGB images.

Vegetation indices	Definition	References
g, r, b	The DN value of each band	/
EXR	$1.4 * r - g$	(Meyer and Neto, 2008)
EXG	$2 * g - r - b$	(Woebbecke et al. 1995)
EXGR	$3 * g - 2.4 * r - b$	(Steward and Tian, 1998)
MGRVI	$(g^2 - r^2) / (g^2 + r^2)$	(Bendig et al. 2015)
NGRDI	$(g - r) / (g + r)$	(Rasmussen et al. 2016)
RGRI	r/g	(Verrelst et al. 2008)
PPRb	$(g - b) / (g + b)$	(Metternicht, 2003)
CIVE	$0.441 * r - 0.881 * g + 0.385 * b + 18.78$	(Kataoka et al. 2003)
VARI	$(g - r) / (g + r - b)$	(Gitelson et al. 2002)
WI	$(g - b) / (r - g)$	(Woebbecke et al. 1995)
GLA	$(2 * g - r - b) / (2 * g + r + b)$	(Guijarro et al. 2011)
RGBVI	$(g^2 - b * r) / (g^2 + b * r)$	(Gamone and Surfus, 1999)
VEG	$g^{(r^k b^{1-k})}, k = 0.667$	(Hague et al. 2006)
COM	$0.25 * EXG + 0.3 * EXGR + 0.33 * CIVE + 0.12 * VEG$	(Guijarro et al. 2011)
COM2	$0.36 * EXG + 0.47 * CIVE + 0.17 * VEG$	(Guijarro et al. 2011)

Note: g: green; r: red; b: blue.

calculated from the DN values of each band of RGB images or the reflectance of each band in the multispectral images. After removing the ground pixels, the average values for all pixels of each plot in each band is taken as the reflectance of the plot.

2.3.2. Extraction of texture features

Texture features refer to recognizable visual effects such as patterns, textures, and surface qualities formed by a group of pixels in a digital image. Texture feature extraction from crop images can be done using the gray-level co-occurrence matrix (GLCM) (Liu et al., 2021; Chen et al.,

Table 2
Calculation formula of vegetation index with multispectral images.

Vegetation indices	Definition	References
g, r, re, nir	The DN value of each band	/
CI	$(nir/re) - 1$	(Gitelson et al. 2003)
DVI	$nir - r$	(Tucker, 1979)
GNDVI	$(nir-g)/(nir+g)$	(Gitelson et al. 1996)
GRVI	$(g - r)/(g + r)$	(Daughtry et al. 2000; Hague et al. 2006)
MCARI	$((re - r) - 0.2 * (re - g)) * (re/r)$	(Daughtry et al. 2000)
MNVI	$(1.5 * (nir^2 - r)) / (nir^2 + r + 0.5)$	(Gong et al. 2003)
MSR	$(nir/r - 1) / (\sqrt{nir/r} + 1)$	(Chen and Gillieson, 2009)
MTCI	$(nir-re) / (re-r)$	(Gitelson et al. 2006)
NDRE	$(nir-re) / (nir+re)$	(Gitelson and Merlyak, 1997)
NDVI	$(nir - r) / (nir + r)$	(Rouse et al. 1974)
NLI	$(nir^2 - r) / (nir^2 + r)$	(Chen et al. 2018)
OSAVI	$(1.16 * (nir - r)) / (nir + r + 0.16)$	(Rondeaux et al. 1996)
RDVI	$(nir-r) / (\sqrt{nir+r})$	(Roujean and Breon, 1995)
RVI1	nir/r	(Tucker, 1979)
RVI2	nir/g	(Tian et al. 2007)
TO	$3 * ((reg - r) - 0.2 * (reg - g) * (reg/r)) / OSAVI$	(Haboudane et al. 2002)
SAVI	$1.5 * (nir - r) / (nir + r + 0.5)$	(Qi et al. 1994)
TVI	$60 * (nir - g) - 100 * (r - g)$	(Brogue and Leblanc, 2001)

Note: g: green; r: red; re: red-edge; nir: near-infrared.

2022). These features can be used to characterize the differences between different crop varieties or growth stages, to reflect the growth and health status of crops under different environmental or management conditions. The RGB nadir image with soil removal was converted into a grayscale image. The correlation coefficient between image texture and aboveground biomass was significantly affected by the image ground resolution and less affected by the computational window size (Yue et al., 2019). (Xiao et al., 2021) showed that there were differences in the correlation coefficients between the texture features and the target variables under different window sizes. However, such differences were not significant. Therefore, we adopted the commonly used 7×7 window size with the sliding step value of 2 based on previous studies (Wang

et al., 2021; Shu et al., 2022a).

2.3.3. Extraction of 3D structure features

Structural features include plant height, canopy convex hull volume, roughness, canopy cover, canopy width, reconstruction points of canopy point cloud and vegetation index of point cloud (Fig. 3 4). The average value of the top 10% highest points in the point cloud is used as the crop upper boundary, which is then subtracted from the ground elevation to obtain the average plant height per plot (Shu et al., 2022c). Point cloud roughness refers to the irregularity of the point cloud surface, indicating its roughness or smoothness. Roughness is an important attribute that can be used to describe the morphological characteristics of the point cloud, such as surface smoothness, texture, shape, etc. The detailed process for calculating roughness can be found in Herrero-Huerta (Herrero-Huerta et al., 2020). Canopy coverage refers to the extent to which the point cloud covers the surface of an object, specifically the ratio of points in the point cloud to the points on the object's surface. It is calculated by dividing the area covered by the canopy point cloud by the area of the plot. Canopy convex hull volume refers to the volume of the geometric body formed by the convex hull generated from the canopy point cloud of the plot. The convex hull is the smallest convex polyhedron formed by the point cloud and can be used to characterize the size or shape of the crop canopy (Zhou et al., 2022). Canopy width is the extent or thickness of the crop canopy in the horizontal direction. Assuming that the monopoly direction is the X-axis, the canopy width at each X-coordinate can be determined by calculating the difference between the maximum and minimum values of the Y-coordinate at that X-coordinate. After removing the background, three rows of soybeans are split into individual rows to calculate the width of each row. The

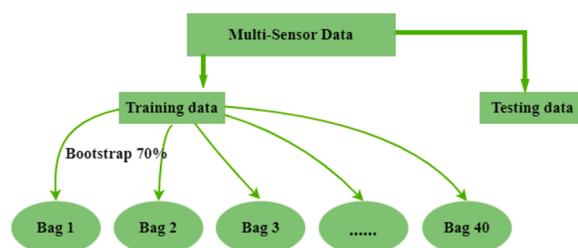


Fig. 4. Construction of 40 bootstrap samples for bagging.

average width of these three rows is then taken as the canopy width of the plot. The effective number of reconstructed canopy point cloud is the number of successfully reconstructed point clouds in a plot under the same treatment, which can reflect the difference in leaf area and shading condition of the crop canopy. RGB vegetation index of 3D canopy point cloud reflects the reflectance information of three bands of RGB for the whole canopy, which can evaluate the plant vigor more comprehensively.

2.4. Modeling methods

Ensemble learning framework of one feature selection optimization based on multimodal data is proposed here through a combination of Shapley value, bagging and stacking. The ensemble framework consists of forty base learners with different training samples. Each base learner evaluates the importance of all candidate features in its own training sample by combining Shapley values. Candidate features with large

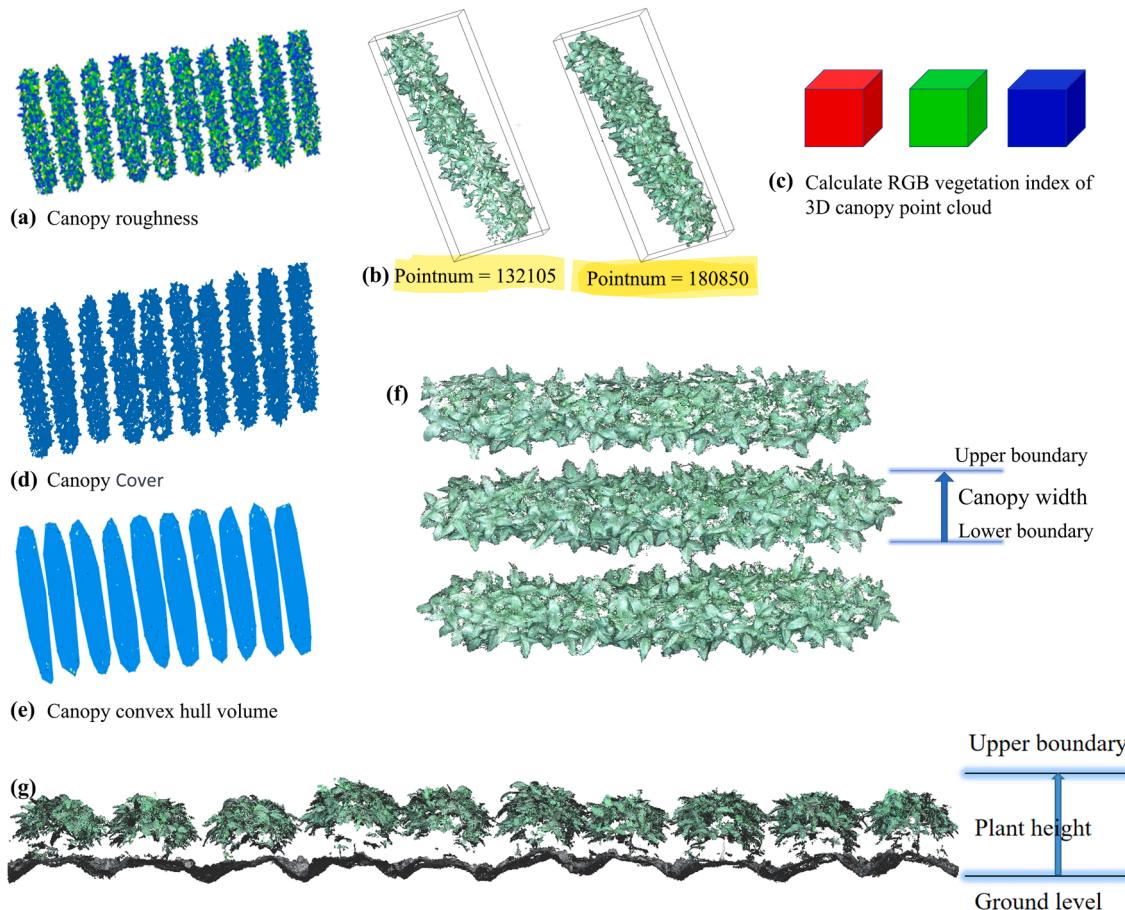


Fig. 3. (a) Canopy roughness (b) reconstruction points of canopy point cloud (c) RGB vegetation index of 3D canopy point cloud (d) canopy coverage (e) canopy convex hull volume (f) plant height.

Shapley values is considered as strong features to facilitate the model's prediction process. The predicted values of these forty models are subsequently used as variables for training and validation in the second layer of the ensemble framework to obtain the final estimated soybean yield.

2.4.1. Shapley value-based feature selection

Shapley value is a concept used in cooperative games to measure the extent to which each player contributes to the total gain obtained in a cooperative game. In a base machine learning model, different subsets of features are divided. The Shapley value of each feature is calculated based on the predictions of each feature subset, measuring the contribution of each feature to these predictions (Sun et al., 2022a; Zhou et al., 2022). The Shapley value for each feature is calculated by marginal contribution, weighted, and summed over all possible combinations of feature values:

$$\phi_j(\text{val}) = \sum_{S \subseteq \{1, \dots, p\} \setminus \{j\}} \frac{|S|!(p - |S| - 1)!}{p!} (\text{val}(S \cup \{j\}) - \text{val}(S))$$

where S is a subset of the features used in the model, x is a vector of feature values of the instances to be explained, and p is the number of features,

$\text{val}_x(S)$ is a prediction of the feature values in the set S that are

However, as more features are added, the number of possible feature sets increases exponentially. Therefore an approximation with Monte-Carlo sampling is used :

The instances of interest x, features j and the number of iterations M are selected. For each iteration, random instance z from the data is selected and a random order of features is generated. Two new instances x_{+j} and x_{-j} are created by combining the values of the instances of interest x and the sample z. x_{+j} and x_{-j} , x_{+j} is the instance of interest, but all values in the order after feature j are replaced by feature values from the sample z. The instance x_{-j} is the same as x_{+j} , but has feature j replaced by the value for feature j from the sample z. The difference in the prediction from the black box is calculated :

$$\hat{\phi}_j = \frac{1}{M} \sum_{m=1}^M \left(\hat{f}(x_{+j}^m) - \hat{f}(x_{-j}^m) \right)$$

All these differences were averaged to derive the final Shapley Value for each feature:

$$\phi_j(x) = \frac{1}{M} \sum_{m=1}^M \hat{\phi}_j^m$$

This calculation process is described in [Algorithm 1](#) :

Algorithm 1. Evaluating the Importance of Candidate Features Using the Shapley Value.

Input: Given a training data matrix X, model f, integer M, feature count p

Output: $\phi_j(x)$: the Shapley value of feature f_j .

For m in range (1, M+1):

 Draw random instance z from the data matrix X

 Choose a random permutation o of the feature values

 Order instance x: $x_o = (x_{(1)}, \dots, x_{(j)}, \dots, x_{(p)})$

 Order instance x: $x_o = (x_{(1)}, \dots, x_{(j)}, \dots, x_{(p)})$

For j in range(p) **do**:

 With j: $x_{+j} = (x_{(1)}, \dots, x_{(j-1)}, x_{(j)}, z_{(j+1)}, \dots, z_{(p)})$

 Without j: $x_{-j} = (x_{(1)}, \dots, x_{(j-1)}, z_{(j)}, z_{(j+1)}, \dots, z_{(p)})$

 Compute marginal contribution: $\phi_j^m = \hat{f}(x_{+j}) - \hat{f}(x_{-j})$

End

End

Calculate Shapley value as the average: $\phi_j(x) = \frac{1}{M} \sum_{m=1}^M \phi_j^m$

marginalized rather than features not included in the set S, which can be expressed as:

$$\text{val}_x(S) = \int \hat{f}(x_1, \dots, x_p) d\mathbb{P}_{x \notin S} - E_X(\hat{f}(X))$$

All possible feature sets must be evaluated with and without the jth feature to calculate the exact Shapley value.

Obviously, the larger the Shapley values of a feature, the greater its impact on the model. Considering the global importance of features, the average of the absolute Shapley values for each feature across all samples is calculated (Nohara et al., 2019). Each base machine learning model in the ensemble learning here is independently feature-reduced based on the magnitude of Shapley values. A comparison is made with two common feature reduction methods: Pearson correlation coefficient and recursive feature elimination (RFE). The maximum Shapley value

feature is selected from each weak learner, and the impact of the number of selected important features by each weak learner on the model accuracy is explored.

2.4.2. Shapley bagging-stacking Ensemble workflow

The complexity and parameter selection of individual models may affect the generalization ability of the models. In addition, certain machine learning algorithms are based on stochastic processes, such as random forest. Ensemble learning can combine the performance of multiple models and tends to have stronger generalization ability with estimation performance. Bagging and stacking are two popular integration methods. Bagging method performs multiple resampling in the original samples to improve model generalization and to reduce the impact of noise and outliers on the model. However, bagging algorithms tend to use homogeneous models (Zhang et al., 2022), which makes the model prone to fall into local optimal solutions and limits the expressiveness of the model. The Stacking method uses the predictions of multiple base learners as new training data to train a meta-learner for the final prediction. This approach leverages the strengths of multiple base learners to create a more powerful ensemble model (Xu et al., 2022). The drawback of stacking methods is that each training iteration uses the entire training set, which makes it more susceptible to noise and outliers. This lack of diversity in the training data also affects generalization ability of the model to some extent.

The ensemble framework here consists of multiple base learners based on the bagging method. The base learners include four types of machine learning models: Lasso, RR, RF, and XGBoost. This selection increases the diversity of models, allowing model to capture features and patterns in the data from different perspectives. Subsequently, each model is subjected to 10 random sampling from the training set, with a sampling ratio of 70%. This process generates forty training sample subsets (Fig. 5). Each subset is independently used to train a model, which enhances the diversity of training samples in the ensemble framework.

Then, the Shapley value method is used to perform strong feature selection on these 40 models with different training samples (Fig. 5), aiming to increase feature diversity and to maximize the performance of each base learner while reducing computational costs. Finally, the predictions from these 40 base learners are treated as new features and used

for training and validation in the second layer of the ensemble framework, resulting in the final estimation of soybean yield.

2.4.3. Selection of meta-learner

The model performances are evaluated for Multiple linear regression (MLR), nonlinear model RF, two weighted averaging methods BMA and EWA when used as a meta-learner in the second layer of the ensemble framework.

BMA characterizes the uncertainty of a model by calculating the probability distribution for a sample through the marginal likelihood, and calculates the posterior probability of its corresponding class in a new sample through a Bayesian formula (Kejzlar et al., 2023). The posterior probability is used to assign weights to each weak learner for weighted averaging (Hamill, 2007; Fei et al., 2023a). The probability of category y given M base models and samples is expressed as:

$$P(y|x) = \sum_{i=1}^M P(M_i|x)P(y|M_i,x)$$

$P(M_i|x)$ denotes the probability of the category y given the model M_i and the sample. In general, $P(M_i|x)$ can be calculated by Bayes' theorem :

$$P(M_i|x) = \frac{P(x|M_i)P(M_i)}{\sum_{j=1}^M P(x|M_j)P(M_j)}$$

where $P(M_i|x)$ denotes the probability distribution of model M_i for the sample.

Finally, the prediction of corresponding category for a new sample can be obtained by a weighted average of the posterior probabilities of all possible models:

$$f(x) = \sum_{i=1}^M P(M_i|x)f_i(x)$$

The EWA method calculates the information entropy of each indicator based on the degree of confusion within the sample, calculates the information utility value, and assigns weights according to the magnitude of the utility value. Better performance can be obtained on different datasets (Chowdhury et al., 2022). Models with very similar/different structure are avoided, thus improving the model robustness and generalization. The information entropy of the prediction results of each weak learner is calculated as follows:

$$e_j = k \sum_{i=1}^n y_{ij} \ln y_{ij}$$

where k is a constant, generally taken as $1/\ln n$, such that the information entropy lies between [0,1]; y_{ij} is the weight of the predicted value of the ith sample in the prediction result of the jth weak learner.

The smaller the entropy value, the greater the amount of information, and the information utility value is expressed as:

$$d_j = 1 - e_j$$

The greater the information utility, the greater the weight assigned to get the weight of each weak learner:

$$w_j = d_j / \sum_{i=1}^m d_i$$

In summary, the workflow of ensemble learning is shown in Fig. 6.

2.5. Model performance evaluation

The evaluation metrics include coefficient of determination (R^2) and root mean square error (RMSE). The 2022 dataset was used as training set. Subsequently, tenfold cross-validation is adopted in feature selection and model parameter optimization. During cross-validation process, different varieties are assigned as the test set in turn, and ultimately the average performance of all test sets is taken for the final performance evaluation (Cernezel et al., 2014). This process helps to ensure the model transferability among varieties in its applications. To validate the

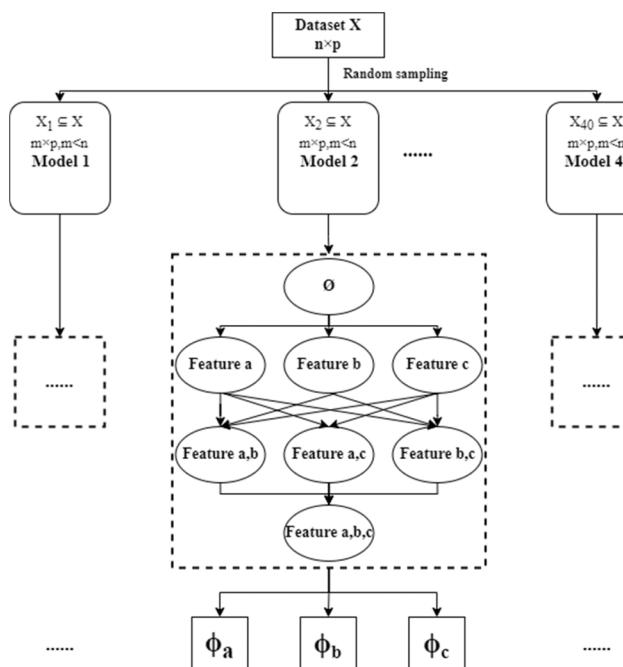


Fig. 5. Shapley value calculation process for each feature.

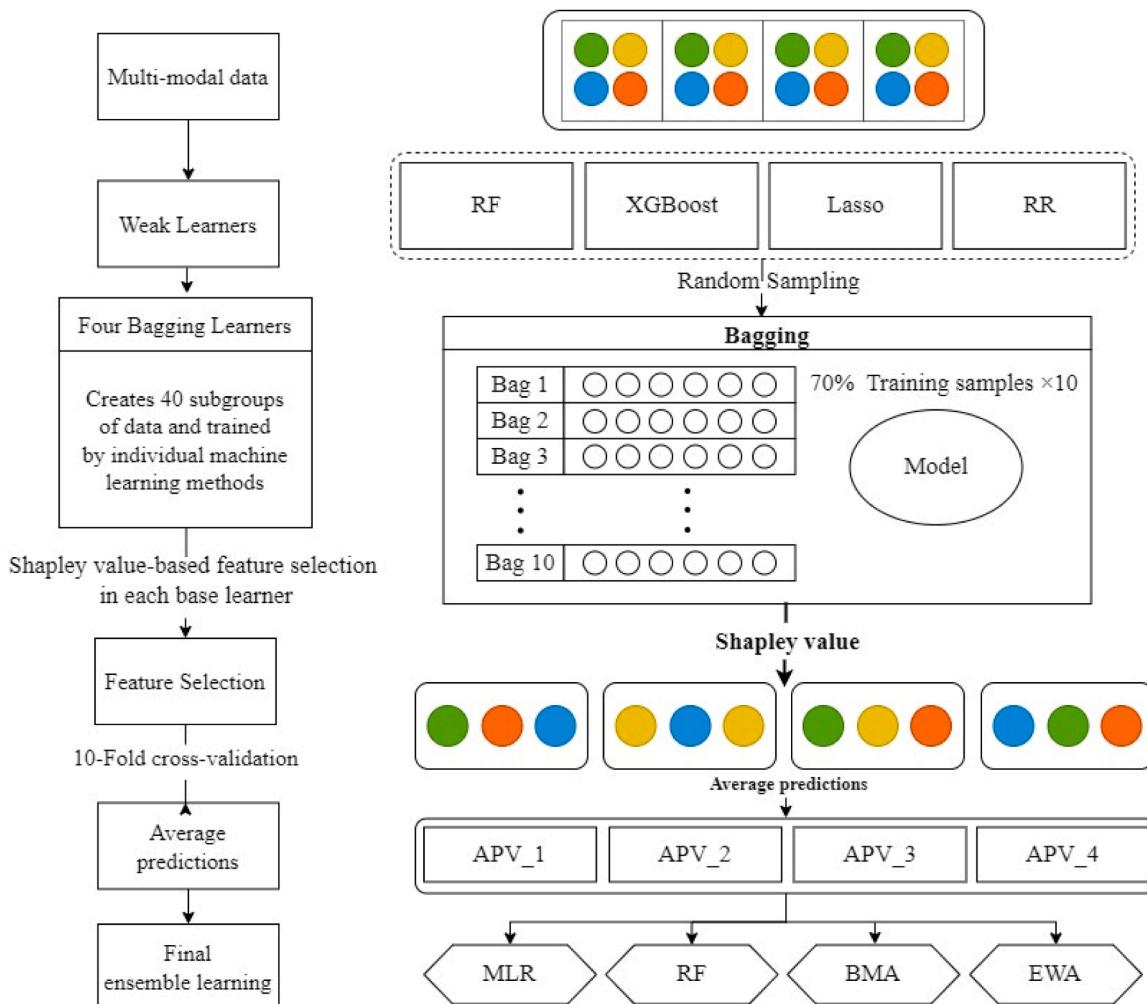


Fig. 6. Workflow of a multi-sensor data fusion and ensemble learning. Lasso: Least Absolute Shrinkage and Selection Operator. RF: Random Forest. RR: Ridge Regression. XGB: eXtreme Gradient Boosting. APV: Average Predicted Value. MLR: Multiple Linear Regression. BMA: Bayesian Model Averaging. EWA: Entropy Weighted Average.

transferability of the yield prediction model among years, the 2023 dataset was used as an independent validation dataset.

3. Results

3.1. Reconstruction accuracy of 3D canopy from the CCO route

To explore the reconstruction accuracy of the CCO method in the horizontal and vertical directions, point clouds obtained by the CCO/nadir method are used to estimate the heights of the rulers and the lengths of the whiteboards placed on the test site, and compared with the manual measurements. The results show that in the horizontal direction, both CCO and nadir accurately estimate the lengths of the whiteboards, with CCO slightly better than nadir (Fig. 7). The RMSE for the four growth stages of CCO method were 10.4 mm, 12.6 mm, 11.9 mm, and 13.7 mm, respectively. In the vertical direction, CCO method still achieves better accuracy for height estimations, with RMSE values of 24.5 mm vs. 64.3 mm, 28.8 mm vs. 55.1 mm, 25.3 mm vs. 67.1 mm, and 19.6 mm vs. 53.5 mm for CCO method vs. nadir method at the four growth stages, respectively.

3.2. Model accuracy on multimodal data fusion and multiple growth stages

The mean values of R^2 and RMSE for the models constructed with

different combinations are given in Table 3. After the integration of all feature variables, the R^2 of the base learner for estimating soybean yield was among 0.50 ~ 0.57. The order of estimation accuracy is RGB nadir > 3D > multispectral, when only considering modeled individual three variables. To compare the model performance before and after feature fusion, the performance of the yield estimation models constructed by each base learner was analyzed. The accuracy was slightly decreased after adding multispectral features to the features calculated with RGB nadir images. R^2 of XGB decreased the most from 0.50 to 0.47 among different algorithm. The estimation accuracy was improved by combining 3D features with either RGB nadir or multispectral images. All four base models achieved the best estimation accuracy when combining RGB nadir features with 3D features, the R^2 was among 0.53 ~ 0.59. The results show that the integration of UAV nadir and CCO imaging can effectively reflect crop yields and improve the estimation accuracy of the model.

The performance of soybean yield prediction models based on different collection dates is compared in Table 4. The best prediction models for each time except 30 DAE achieved with $R^2 > 0.30$. The highest accuracy ($R^2 = 0.51$) of yield prediction is achieved at 91 DAE when soybeans are at the grain filling stage. Yield prediction accuracy is higher when multiple growth stages are combined together than individual growth stages.

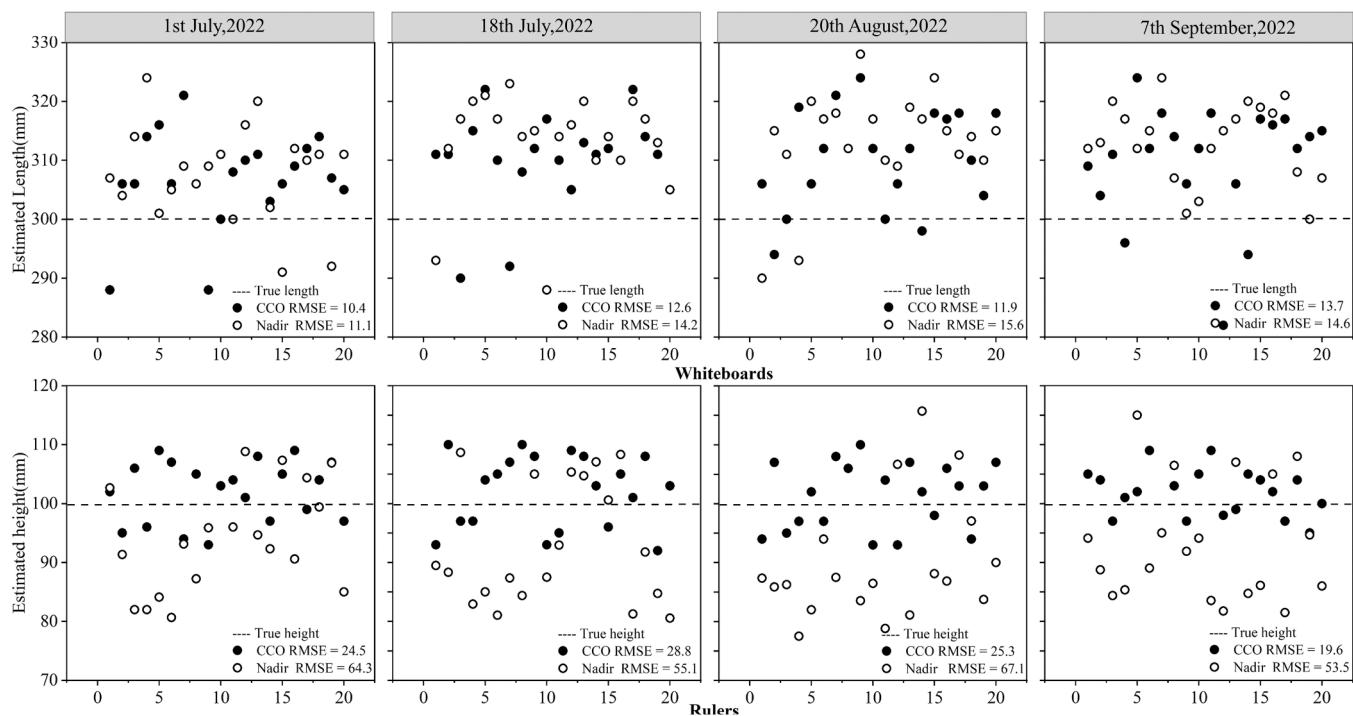


Fig. 7. Comparison of CCO and nadir method for estimating the ruler heights and whiteboard lengths.

Table 3
Validation of different data combinations for yield estimation.

Feature type	Metrics	RGB nadir	MS	3D	RGB nadir+MS	RGB nadir+3D	MS+3D	All
Lasso	R ²	0.57	0.50	0.53	0.56	0.59	0.56	0.57
	RMSE	653.73	713.52	685.81	690.83	644.25	667.81	659.93
RF	R ²	0.56	0.48	0.56	0.55	0.58	0.57	0.57
	RMSE	664.76	744.12	659.88	673.11	650.34	654.30	666.14
RR	R ²	0.58	0.51	0.54	0.58	0.59	0.56	0.56
	RMSE	640.98	700	680.14	646.58	644.98	659.37	657.09
XGB	R ²	0.50	0.43	0.50	0.47	0.53	0.47	0.5
	RMSE	719.65	767.55	713.73	733	685.46	727.12	731.59

Note: RMSE (kg/ha); RGB nadir: features extracted from RGB nadir images; MS: features extracted from multispectral images; 3D: features extracted from three-dimensional point cloud. Lasso: Least Absolute Shrinkage and Selection Operator. RF: Random Forest. RR: Ridge Regression. XGB: eXtreme Gradient Boosting.

Table 4
Validation of different collection dates for yield estimation.

Date	Metrics	Lasso	RF	RR	XGBoost
30 DAE	R ²	0.19	0.20	0.15	0.17
	RMSE	792.01	798.27	813.78	790.16
47 DAE	R ²	0.29	0.32	0.34	0.27
	RMSE	721.28	665.97	662.72	717.85
81 DAE	R ²	0.39	0.36	0.33	0.39
	RMSE	641.19	674.32	656.93	691.16
91 DAE	R ²	0.51	0.51	0.50	0.44
	RMSE	664.89	611.08	700.54	787.34
All times	R ²	0.59	0.59	0.59	0.52
	RMSE	638.34	644.36	625.39	704.58

Note: RMSE (kg/ha); DAE: Days after emergence; RF: Random Forest. RR: Ridge Regression.

3.3. Analysis of feature selection

3.3.1. Performance evaluation of Shapley value-based feature selection

To investigate the impact of different feature reduction on the model performance, the 280-dimensional features extracted from the UAV images are subjected to Pearson correlation analysis with the plot yield. The features with the highest correlation coefficients are selected. RFE

recursive elimination and calculation of Shapley values are also performed on each weak learner to output the most important features and to evaluate the impact of the number of important features on the model performance.

Fig. 8 shows the comparison of the effect of different feature reduction methods with the base learner on the training set. Shapley values and RFE significantly improve the model performance after important feature filtering compared by using all feature variables. Among them, the Shapley values-based method performs the best, the performance of all models except XGBoost has been improved during feature dimension reduction. All four base learners achieved maximum accuracy at a feature dimension of 60, improving R² by 8.8%, 3.5%, 7.1%, and 8.0%, respectively, compared to that with all 280-dimensional features. Lasso has the highest prediction accuracy with R² of 0.62. The Pearson method performs poorly, significantly reduces the model performance at low feature dimensions. Overall, the selection of important features for different modal variables using Shapley values can effectively improve the estimation accuracy and reduce the computational cost.

3.3.2. Shapley values for different features in the base model

The Shapley Value matrix of 60 significant features is calculated on the training set of each base model. The **Fig. 9** shows the number and

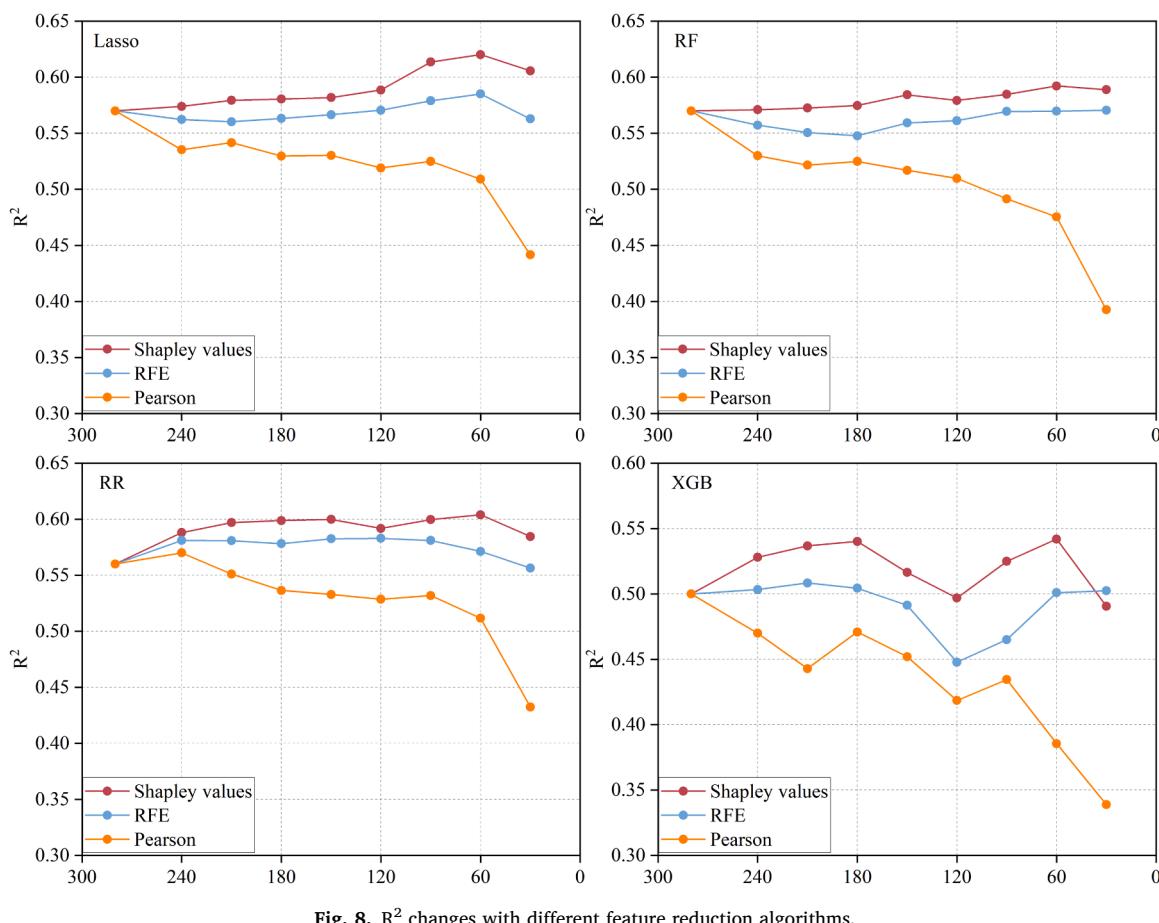


Fig. 8. R^2 changes with different feature reduction algorithms.

influence of significant features for different modalities in each base model. The vegetation index and texture features extracted from RGB nadir images have the greatest influence on all four base models. In the Lasso and RR linear models, the Shapley values indicate that the multispectral features are more important. However, the Shapley values of multispectral features are much lower than the other modal features in the RF and XGB models. Three-dimensional features achieved high Shapley value on all four base learners and were much more important than multispectral features on both RF and XGB models. Among the important features of the different base learners, 44 out of 60 important features are the same for Lasso and RR, including MCARI_1st July, mean_7th September, Pointnum_7th September, etc., and 36 out of 60 important features are the same for RF and XGBoost, including EXR_7th September, r_7th September, asm_18th July, etc. Overall, CCO features can effectively provide supplementary information for soybean phenotypic traits estimation. The Shapley value brings better transparency to the machine learning model, and provides a framework to identify more important values on yield prediction and a better understanding about the underlying predictors.

3.4. Performance of ensemble learning

Fig. 10 provides a visual representation of the accuracy analysis in soybean yield prediction using the Shapley bagging-stacking Ensemble algorithm applied on independent dataset of year 2023. The best prediction accuracy ($R^2 = 0.70$, RMSE = 586.86 Kg/ha) was achieved by utilizing BMA as the meta learner, with an average accuracy improvement of 12.9% compared to the base learner Lasso ($R^2 = 0.62$). The improvement used with RF, MLR, and EWA as the meta-learner was 8.1%, 9.7%, and 8.1%. According to the distribution of errors in each region, the varieties with yields of 2000–5000 Kg/ha showed a lower

bias in estimation compared to the varieties with yields of 0–2000 Kg/ha. These results suggested that the constructed Shapley bagging-stacking Ensemble workflow, especially when BMA is utilized as the meta learner, can be used for improved soybean yield prediction.

4. Discussion

4.1. UAV CCO photography

Previous studies have shown that UAVs using different flight paths and carrying different sensors can fulfill different application requirements (Xie and Yang, 2020). Compared with UAV nadir imaging, UAV CCO imaging can avoid field obstacles (Zhou and Zhang, 2020), adjust flight height and shooting angles based on specific crop conditions, and cost-effectively obtain more comprehensive surface information. In terms of the reconstruction accuracy of the reference whiteboards/rulers (Fig. 7), the reconstruction CCO was slightly better than nadir in the horizontal direction. Since our whiteboards were placed on bare ground, nadir imaging similarly accomplished a high reconstruction accuracy. When photographing crops in the field environment where there is a lot of shading, nadir photography will be more challenging (Blackman and Yuan, 2020). In the vertical direction, CCO achieves much higher reconstruction accuracy than nadir. Nadir imaging can obtain the elevation information of the top pixels of the crop through depth estimation to extract the height of the crop (Teshome et al., 2023), but it is difficult for nadir imaging to obtain the pixels on the side of the target due to the limitation of the shooting angle (Che et al., 2020). CCO imaging acquires multiple views by flying around and adjusting the tilt angle of the lens, which helps to reconstruct the crop structure more completely.

UAV CCO imaging can acquire multi-view images of the entire plot

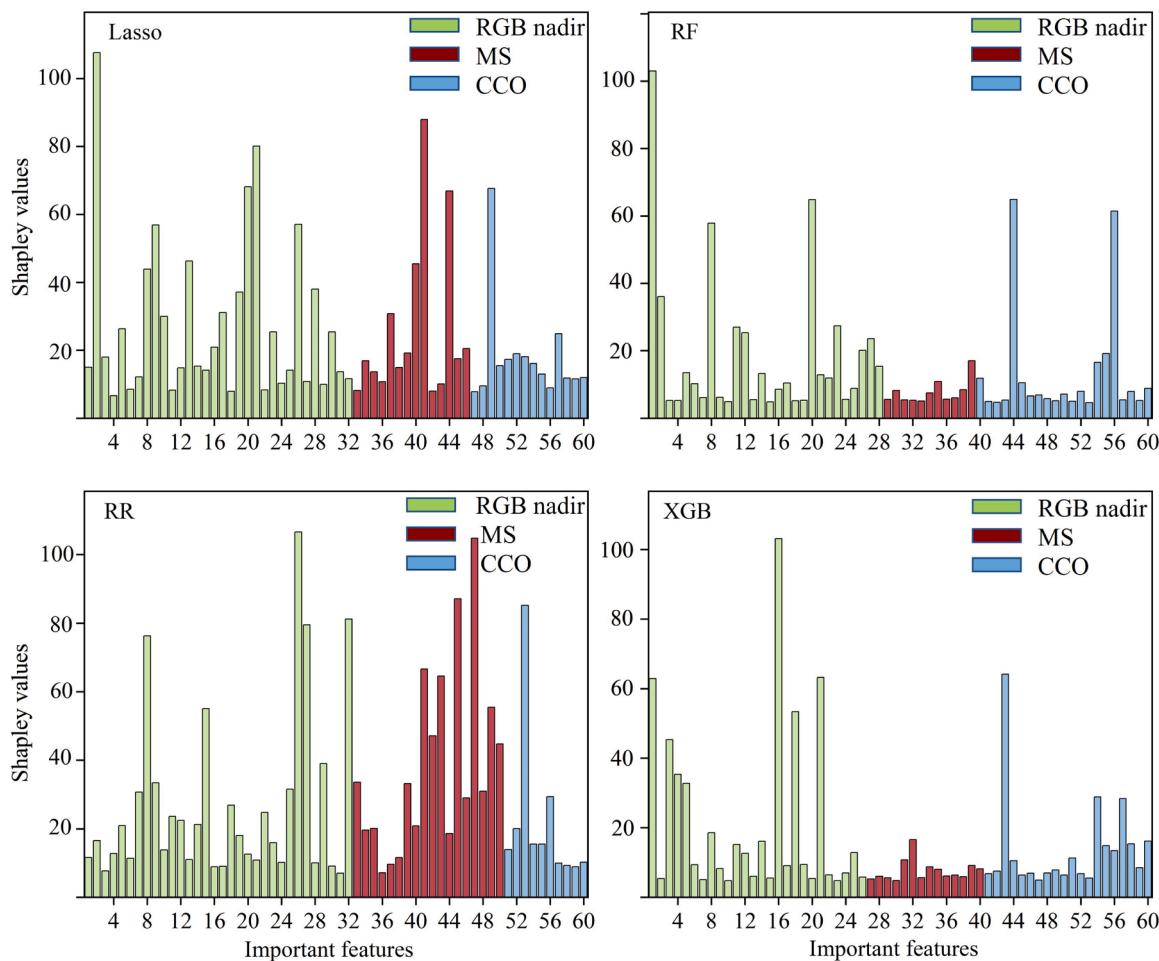


Fig. 9. Shapley values distribution of different modal. RGB nadir: features extracted from RGB nadir images; MS: features extracted from multispectral images; CCO: features extracted from three-dimensional point cloud.

within 3 h for soybean covering 2892 m², when flying at an altitude of 5 m above the canopy, with 85% intra-circle overlap and 50% inter-circle overlap. Crop structures may vary, and settings of flight parameter should be adapted accordingly to fit specific application scenarios (Hu et al., 2019). Due to the relatively low altitude of the crops, it is often necessary to reduce the flight altitude to ensure image clarity to obtain an accurate 3D point cloud, but this will also increase the image acquisition time. Therefore, using high-quality cameras or coordinating multiple UAVs for simultaneous operation is an effective approach to further ensure the efficiency of CCO data acquisition (Xiao et al., 2023b).

4.2. The potential of Shapley values

There is a large amount of redundant and irrelevant information in the feature dataset for multi-temporal and multi-modal data (Yang et al., 2020). Effective tools are needed to filter out features that have a close relationship with soybean yield. Different base learners have different sensitivities to feature sets in ensemble learning. Pearson correlation coefficient and RFE are commonly used for important feature screening (Liu and An, 2020; Shu et al., 2022b). The performance of these two methods have been evaluated and compared Shapley value. Overall, the Pearson method is the least effective among the four base learners, as the method only considers the linear relationship between the independent and dependent variables and ignores the existence correlation between multiple independent variables (La et al., 2019). RFE performs better in Lasso and RF models, and effectively reduces the feature dimension.

However, the selected features did not significantly improve the model accuracy, as RFE focuses on the importance of individual features and ignores the combination of features that may have a positive effect on model performance (Benkessirat et al., 2019).

In contrast to the traditional uniform feature reduction for all base learners, the Shapley values method can evaluate the importance of each feature by calculating its Shapley values (Marcilio and Eler, 2021). Shapley values method considers not only the individual impact of each feature on the model, but also the interactions between features, and therefore performed well in crop model interpretation and important feature selection (Jones et al., 2022). Among the 64 important features selected by Shapley value, Lasso and RR have more identical features, and XGBoost and RF have more identical features, which is related to the structure of the model. Lasso and RR belong to the same category of linear regression model. Both RF and XGBoost are ensemble models based on decision trees. There are fewer similar features between regression models and ensemble models, which shows the necessity that different important features should be used for different models.

4.3. The potential of bagging-stacking ensemble learning

Inspired by the advantages of both Bagging and Stacking ensemble learning methods, heterogeneous learners are used to samples bootstrap from the training set. An ensemble meta-learner is then employed in the second layer to integrate the prediction results of different samples trained by heterogeneous models. The feature selection capability of Shapley values is also incorporated into the ensemble framework. In the

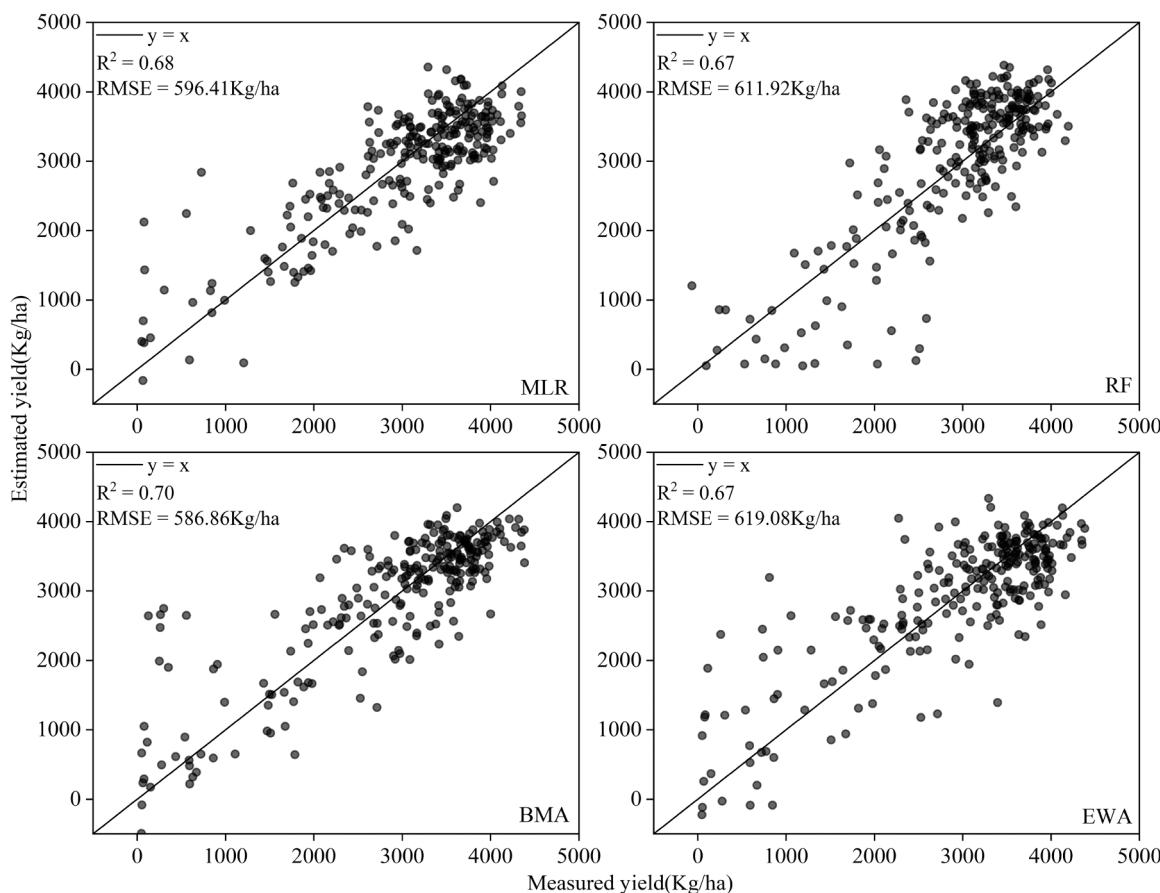


Fig. 10. Performance comparison of different meta-learners in Shapley bagging-stacking Ensemble workflow applied on independent dataset of year 2023.

prediction results on independent dataset of year 2023, the Shapley bagging-stacking model maintained a reliable performance, proving the robustness of the model. According to the distribution of yield predictions, the model's performance was lower for varieties with yields ranging from 0 to 2000 Kg/ha compared to the varieties with yields ranging from 2000 to 5000 Kg/ha. This may be attributed to the relatively small number of training samples in this range. In further research, it may be necessary to consider enhancing the diversity and scale of the dataset to improve the Shapley bagging-stacking model's predictive capabilities across various yield ranges.

In the selection of the meta-learner, the BMA method clearly outperforms other models (Fig. 10). BMA can calculate the posterior probabilities and provide estimates of the credibility for each model (Berkhout et al., 2023), thus minimize the impact of model uncertainty on the prediction results and provide more accurate prediction intervals. The MLR model performs better when used as the meta-learner in the two stacking methods, as the predictions of the weak learners are more stable under multiple sampling conditions. MLR is less affected by randomness and noise compared with RF. The EWA method measures the uncertainty of the model by calculating the information entropy. The weights are dynamically adjusted based on the credibility and information content of the weak learners. This method also demonstrates great potential as a meta-learner in ensemble learning (Hino and Murata, 2013). In theory, there is no clear superiority or inferiority among the BMA, EWA, and stacking method, when considering the correct data generation model and low noise. The performance of these ensemble methods may vary on different datasets (Zian et al., 2021).

In future research, the focus will be on the improvement of the interpretability of Shapley bagging-stacking Ensemble workflow. Model interpretability contributes to a more accurate understanding of crop growth dynamics and environmental impacts (Han et al., 2022), thereby

providing support for agricultural management and resource allocation. Besides, the yield prediction presented here is expandable to more traits and can be continuously improved as new data become available. Instead of retraining the model from scratch, the model weights can be easily updated by retraining the model on new multi-modal data, including remote sensing images (Schwalbert et al., 2020), hyperspectral data (Shu et al., 2022b) and soil information (Wang, 2019). Through the Shapley bagging-stacking model, high-dimensional agricultural information can be efficiently filtered to facilitate breeding efforts and sustainable food production.'

4.4. Model accuracy on multiple growth stages

Data from four different growth stages were used to estimate soybean yield, with the highest estimation accuracy achieved during the grain filling stage (Table 4). Grain filling is the stage when soybean transfers organic materials such as starch and protein produced by photosynthesis from the nutrient organs to the kernels. Previous studies show that information from UAV in this stage is closely related to the final thousand grain weight, so the accuracy of yield estimation in this stage is the highest (Zhang et al., 2019; Joshi et al., 2023). Ren et al. (2023) investigated the effect of growth stages on yield estimation, and the results showed that integration of multiple stages were the best for yield estimation. This is because the rate of crop canopy development is also an important factor affecting yield. The multiple-stages conjunction complements the information gaps related to growth rate, enabling the model to capture crop growth changes in agricultural scenarios and thus achieve higher yield prediction accuracy.

The research exclusively concentrated on predicting soybean yield, collecting image data from multiple sensors at four different times. However, there are still some differences in growth stages between

varieties which can introduce errors into the model predictions. Further research will be done to incorporate growth stages information from the field observation to evaluate the growth duration effect on yield prediction.

5. Conclusions

This study proposes a method for obtaining point clouds of soybean canopy by using UAV combined with CCO route design, assesses the contribution of fusing point cloud data and feature variables from different sensors to soybean yield prediction. Compared to using all features, the use of the Shapley value method to reduce feature dimensions to 64 significantly improves the accuracy of yield predictions, the R^2 increases from 0.57 to 0.62. A heterogeneous bagging and stacking method are employed to predict soybean yield on independent dataset of year 2023. The results demonstrate that Shapley value based ensemble framework enhance the prediction accuracy of soybean yield, when using BMA as the meta-learner, the R^2 reaches 0.70, which is a 12.9% improvement in accuracy compared to the best-performing base learner, Lasso. By considering both model and sample diversity and ensuring optimized feature selection for each base learner, the current proposed method provides possibility of establishing a high-precision phenotypic trait extraction model.

CRediT authorship contribution statement

Xu Yun: Validation, Supervision. **Li Yinghui:** Resources. **Yan Long:** Resources. **Xiao Shunfu:** Software. **Fei Shuaipeng:** Conceptualization. **Sun Xuhong:** Investigation. **Li Mingxue:** Data curation. **Wang Lei:** Resources, Data curation. **Chen Haochong:** Formal analysis, Data curation, Conceptualization. **Ma Yuntao:** Supervision, Resources, Funding acquisition, Formal analysis. **Zhang Yong:** Supervision, Resources, Investigation, Data curation. **Qiu Lijuan:** Validation, Resources. **Sun Guangyao:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The authors declare the following financial interests/personal relationships which may be considered as potential competing interests.

Data Availability

Data will be made available on request.

Acknowledgements

This research was funded by Pinduoduo-China Agricultural University Research Fund (No.PC2023B01004), the Inner Mongolia Department of Science and Technology (202204711910816, 202304711910152), National Key R & D Project (2021YFD1201602).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Amol, V., 2022. Crop management and therapeutic efficacy of soybean (*Glycine max*) on health. *J. Food, Agric. Environ.* 20, 58–62.
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G., 2015. 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.* 39, 79–87.
- Benkessirat, A., Benblidia, N., & Ieee (2019). Fundamentals of feature selection: an overview and comparison. In, *16th IEEE/ACS International Conference on Computer Systems and Applications (AICCSA)*. Abu Dhabi, U ARAB EMIRATES.
- Berkhout, S.W., Haaf, J.M., Gronau, Q.F., Heck, D.W., & Wagenaarmakers, E.-J. (2023). A tutorial on Bayesian model-averaged meta-analysis in JASP. *Behavior Research Methods*.
- Blackman, R., & Yuan, F. (2020). Detecting Long-Term Urban Forest Cover Change and Impacts of Natural Disasters Using High-Resolution Aerial Images and LiDAR Data. *Remote Sensing*, 12.
- Boukir, S., Feng, W., & Ieee Comp, S.O.C. (2021). Boundary bagging to address training data issues in ensemble classification. In, *25th International Conference on Pattern Recognition (ICPR)* (pp. 9975–9981). Electr Network.
- Broge, N.H., Leblanc, E., 2001. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sens. Environ.* 76, 156–172.
- Cai, Z., Liang, J., Hou, K., & Liu, S. (2022). LiDAR Point Cloud Image Interpolation via Sep arable Convolution. In, *41st Chinese Control Conference (CCC)* (pp. 6709–6713). Hefei, PEOPLES R CHINA.
- Cao, F., Liu, F., Guo, H., Kong, W.W., Zhang, C., He, Y., 2018. Fast Detection of *Sclerotinia sclerotiorum* on Oilseed Rape Leaves Using Low-Altitude Remote Sensing Technology. *Sensors* 18.
- Cernezel, A., Rozman, I., & Brumen, B. (2014). Comparisons between Three Cross-Validation Methods for Measuring Learners' Performances. In, *24th International Conference on Information Modelling and Knowledge Bases* (pp. 77–87). Kiel, GERMANY.
- Che, Y., Wang, Q., Zhou, L., Wang, X., Li, B., Ma, Y., 2022. The effect of growth stage and plant counting accuracy of maize inbred lines on LAI and biomass prediction. *Precis. Agric.* 23, 2159–2185.
- Che, Y.P., Wang, Q., Xie, Z.W., Zhou, L., Li, S.W., Hui, F., Wang, X.Q., Li, B.G., Ma, Y.T., 2020. Estimation of maize plant height and leaf area index dynamics using an unmanned aerial vehicle with oblique and nadir photography. *Ann. Bot.* 126, 765–773.
- Chen, C., Qin, Y., Chen, H., Cheng, J., He, B., Wan, Y., Zhu, D., Gao, F., Zhou, X., 2022. Machine learning to differentiate small round cell malignant tumors and non-small round cell malignant tumors of the nasal and paranasal sinuses using apparent diffusion coefficient values. *Eur. Radiol.* 32, 3819–3829.
- Chen, H., Huang, W., Li, W., Niu, Z., Zhang, L., & Xing, S. (2018). Estimation of LAI in Winter Wheat from Multi-Angular Hyperspectral VNIR Data: Effects of View Angles and Plant Architecture. *Remote Sensing*, 10.
- Chen, Y., Gillieson, D., 2009. Evaluation of Landsat TM vegetation indices for estimating vegetation cover on semi-arid rangelands: a case study from Australia. *Can. J. Remote Sens.* 35, 435–446.
- Chowdhury, N.K., Kabir, M.A., Rahman, M.M., Islam, S.M.S., 2022. Machine learning for detecting COVID-19 from cough sounds: An ensemble-based MCDM method. *Comput. Biol. Med.* 145.
- Daughtry, C.S.T., Walther, C.L., Kim, M.S., de Colstoun, E.B., McMurtrey, J.E., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239.
- Ding, F., Li, C.C., Zhai, W.G., Fei, S.P., Cheng, Q., Chen, Z., 2022. Estimation of Nitrogen Content in Winter Wheat Based on Multi-Source Data Fusion and Machine Learning. *Agric. -Basel* 12.
- Fei, S., Chen, Z., Li, L., Ma, Y., Xiao, Y., 2023a. Bayesian model averaging to improve the yield prediction in wheat breeding trials. *Agric. For. Meteorol.* 328.
- Fei, S., Hassan, M.A., Xiao, Y., Su, X., Chen, Z., Cheng, Q., Duan, F., Chen, R., Ma, Y., 2023b. UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat. *Precis Agric.* 24, 187–212.
- Fei, S., Li, L., Han, Z., Chen, Z., Xiao, Y., 2022. Combining novel feature selection strategy and hyperspectral vegetation indices to predict crop yield. *Plant Methods* 18.
- Feng, L., Chen, S., Zhang, C., Zhang, Y., He, Y., 2021. A comprehensive review on recent applications of unmanned aerial vehicle remote sensing with various sensors for high-throughput plant phenotyping. *Comput. Electron. Agric.* 182.
- Feng, L., Li, Y., Wang, Y., Du, Q., 2020. Estimating hourly and continuous ground-level PM2.5 concentrations using an ensemble learning algorithm: The ST-stacking model. *Atmos. Environ.* 223.
- Fu, P., Meacham-Hensold, K., Guan, K., Bernacchi, C.J., 2019. Hyperspectral Leaf Reflectance as Proxy for Photosynthetic Capacities: An Ensemble Approach Based on Multiple Machine Learning Algorithms. *Front. Plant Sci.* 10.
- Gamon, J.A., Surfus, J.S., 1999. Assessing leaf pigment content and activity with a reflectometer. *N. Phytol.* 143, 105–117.
- Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160, 271–282.
- Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298.
- Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87.
- Gitelson, A.A., Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18, 2691–2697.
- Gitelson, A.A., Vina, A., Verma, S.B., Rundquist, D.C., Arkebauer, T.J., Keydan, G., Leavitt, B., Ciganda, V., Burba, G.G., Suyker, A.E., 2006. Relationship between gross

- primary production and chlorophyll content in crops: Implications for the synoptic monitoring of vegetation productivity. *J. Geophys. Res. -Atmospheres* **111**.
- Gong, P., Pu, R.L., Biging, G.S., Larrieu, M.R., 2003. Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data. *Ieee Trans. Geosci. Remote Sens.* **41**, 1355–1362.
- Guizarro, M., Pajares, G., Riomerros, I., Herrera, P.J., Burgos-Artizzu, X.P., Ribeiro, A., 2011. Automatic segmentation of relevant textures in agricultural images. *Comput. Electron. Agric.* **75**, 75–83.
- Guo, B., Sun, L., Ren, H., Sun, R., Wei, Z., Hong, H., Luan, X., Wang, J., Wang, X., Xu, D., Li, W., Qiu, L.-J., 2022. Soybean genetic resources contributing to sustainable protein production. *Theor. Appl. Genet.* **135**, 4095–4121.
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **81**, 416–426.
- Hague, T., Tillett, N.D., Wheeler, H., 2006. Automated crop and weed monitoring in widely spaced cereals. *Precis. Agric.* **7**, 21–32.
- Hamill, T.M., 2007. Calibrated surface temperature forecasts from the Canadian ensemble prediction system using Bayesian model averaging. *Mon. Weather Rev.* **135**, 4226–4230.
- Han, L., Yang, G.J., Yang, X.D., Song, X.Y., Xu, B., Li, Z.H., Wu, J.T., Yang, H., Wu, J.W., 2022. An explainable XGBoost model improved by SMOTE-ENN technique for maize lodging detection based on multi-source unmanned aerial vehicle images. *Comput. Electron. Agric.* **194**.
- Hernandez, J., Lobos, G.A., Matus, I., del Pozo, A., Silva, P., Galleguillos, M., 2015. Using Ridge Regres. Models Estim. Grain Yield Field Spectr. Data Bread Wheat (*Triticum aestivum L.*) Grown Three Water Regimes *Remote Sens.* **7**, 2109–2126.
- Herrero-Huerta, M., Bucksch, A., Puttonen, E., & Rainey, K.M. (2020). Canopy Roughness: A New Phenotypic Trait to Estimate Aboveground Biomass from Unmanned Aerial System. *Plant Phenomics*, 2020.
- Hino, H., Murata, N., 2013. Information estimators for weighted observations. *Neural Netw.* **46**, 260–275.
- Hu, P.C., Guo, W., Chapman, S.C., Guo, Y., Zheng, B.Y., 2019. Pixel size of aerial imagery constrains the applications of unmanned aerial vehicle in crop breeding. *Isprs J. Photogramm. Remote Sens.* **154**, 1–9.
- Jin, X., Li, Z., Feng, H., Ren, Z., Li, S., 2020. Deep neural network algorithm for estimating maize biomass based on simulated Sentinel 2A vegetation indices and leaf area index. *Crop. Journal* **8**, 87–97.
- Jones, E.J., Bishop, T.F.A., Malone, B.P., Hulme, P.J., Whelan, B.M., Filippi, P., 2022. Identifying causes of crop yield variability with interpretive machine learning. *Comput. Electron. Agric.* **192**.
- Joshi, D.R., Clay, S.A., Sharma, P., Rekabdarkolaee, H.M., Kharel, T., Rizzo, D.M., Thapa, R., & Clay, D.E. (2023). Artificial intelligence and satellite-based remote sensing can be used to predict soybean (*Glycine max*) yield. *Agronomy Journal*.
- Ju, C., Bibaut, A., van der Laan, M., 2018. The relative performance of ensemble methods with deep convolutional neural networks for image classification. *J. Appl. Stat.* **45**, 2800–2818.
- Kataoka, T., Kaneko, T., Okamoto, H., Hata, S., Ieee, I., 2003. Crop growth estimation system using machine vision. In: *IEEE/ASME Int. Conf. Adv. Intell. Mechatron.* 1079–1083.
- Kejzlar, V., Bhattacharya, S., Son, M., Maiti, T., 2023. Black Box Variational Bayesian Model Averaging. *Am. Stat.* **77**, 85–96.
- La, C.S., Tao, Y.S., Xu, F.Y., Ng, W.W.Y., Jia, Y.W., Yuan, H.L., Huang, C., Lai, L.L., Xu, Z., Locatelli, G., 2019. A robust correlation analysis framework for imbalanced and dichotomous data with uncertainty. *Inf. Sci.* **470**, 58–77.
- Liu, H.P., An, H.J., 2020. Preliminary tests on the performance of MLC-RFE and SVM-RFE in Lansat-8 image classification. *Arab. J. Geosci.* **13**.
- Liu, S., Jin, X., Nie, C., Wang, S., Yu, X., Cheng, M., Shao, M., Wang, Z., Tuohuti, N., Bai, Y., Liu, Y., 2021. Estimating leaf area index using unmanned aerial vehicle data: shallow vs. deep machine learning algorithms. *Plant Physiol.* **187**, 1551–1576.
- Lobos, G.A., Matus, I., Rodriguez, A., Romero-Bravo, S., Araus, J.L., del Pozo, A., 2014. Wheat genotypic variability in grain yield and carbon isotope discrimination under Mediterranean conditions assessed by spectral reflectance. *J. Integr. Plant Biol.* **56**, 470–479.
- Maes, W.H., Steppé, K., 2019. Perspectives for Remote Sensing with Unmanned Aerial Vehicles in Precision Agriculture. *Trends Plant Sci.* **24**, 152–164.
- Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., Fritschi, F.B., 2020. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sens. Environ.* **237**.
- Malik, O.A., Faisal, M., Hussein, B.R., & Ieee (2021). Ensemble Deep Learning Models for Fine-grained Plant Species Identification. In, 8th IEEE Asia-Pacific Conference on Computer Science and Data Engineering (IEEE CSDE). Electr Network.
- Marcello, W.E., Eler, D.M., 2021. Explaining dimensionality reduction results using Shapley values. *Expert Syst. Appl.* **178**.
- Mardanisamani, S., Maleki, F., Kassani, S.H., Rajapaksa, S., Duddu, H., Wang, M., Shiraffe, S., Ryu, S., Josuttes, A., Zhang, T., Vail, S., Pozniak, C., Parkin, I., Stavness, I., Eramian, M., & Ieee (2019). Crop Lodging Prediction from UAV-Acquired Images of Wheat and Canola using a DCNN Augmented with Handcrafted Texture Features. In, 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2657–2664). Long Beach, CA.
- Metternicht, G., 2003. Vegetation indices derived from high-resolution airborne videography for precision crop management. *Int. J. Remote Sens.* **24**, 2855–2877.
- Meyer, G.E., Neto, J.C., 2008. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **63**, 282–293.
- Nohara, Y., Matsumoto, K., Soejima, H., Nakashima, N., & Assoc Comp, M.. (2019). Explanation of Machine Learning Models Using Improved Shapley Additive Explanation. In, 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (ACM-BCB) (pp. 546–546). Niagara Falls, NY.
- Ogawa, D., Sakamoto, T., Tsunematsu, H., Kanno, N., Nonoue, Y., Yonemaru, J.-i., 2021. Haplotype analysis from unmanned aerial vehicle imagery of rice MAGIC population for the trait dissection of biomass and plant architecture. *J. Exp. Bot.* **72**, 2371–2382.
- Pal, M. (2007). Ensemble Learning with Decision Tree for Remote Sensing Classification. In, Conference of the World-Academy-of-Science-Engineering-and-Technology (pp. 735–737). Bangkok, THAILAND.
- Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A MODIFIED SOIL ADJUSTED VEGETATION INDEX. *Remote Sens. Environ.* **48**, 119–126.
- Rasmussen, J., Ntakos, G., Nielsen, J., Svensgaard, J., Poulsen, R.N., Christensen, S., 2016. Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots? *Eur. J. Agron.* **74**, 75–92.
- Ren, P., Li, H., Han, S., Chen, R., Yang, G., Yang, H., Feng, H., & Zhao, C. (2023). Estimation of Soybean Yield by Combining Maturity Group Information and Unmanned Aerial Vehicle Multi-Sensor Data Using Machine Learning. *Remote Sensing*, 15.
- Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **55**, 95–107.
- Roujean, J.L., Breon, F.M., 1995. ESTIMATING PAR ABSORBED BY VEGETATION FROM BIDIRECTIONAL REFLECTANCE MEASUREMENTS. *Remote Sens. Environ.* **51**, 375–384.
- Schwalbert, R.A., Amado, T., Corassa, G., Pott, L.P., Prasad, P.V.V., Ciampitti, I.A., 2020. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. *Agric. For. Meteorol.* **284**.
- Shu, M., Fei, S., Zhang, B., Yang, X., Guo, Y., Li, B., Ma, Y., 2022a. Application of UAV Multisensor Data and Ensemble Approach for High-Throughput Estimation of Maize Phenotyping Traits. *Plant Phenomics* **2022**, 9802585.
- Shu, M.Y., Bai, K., Meng, L., Yang, X.H., Li, B.G., Ma, Y.T., 2023. Assessing maize lodging severity using multitemporal UAV-based digital images. *Eur. J. Agron.* **144**.
- Shu, M.Y., Dong, Q.Z., Fei, S.P., Yang, X.H., Zhu, J.Y., Meng, L., Li, B.G., Ma, Y.T., 2022b. Improved estimation of canopy water status in maize using UAV-based digital and hyperspectral images. *Comput. Electron. Agric.* **197**.
- Shu, M.Y., Shen, M.Y., Dong, Q.Z., Yang, X.H., Li, B.G., Ma, Y.T., 2022c. Estimating the maize above-ground biomass by constructing the tridimensional concept model based on UAV-based digital and multi-spectral images. *Field Crops Res.* **282**.
- Steward, B.L., & Tian, L.F. (1998). Real-time machine vision weed-sensing. *ASAE Annual International Meeting*, Orlando, Florida, USA, 12–16 July, 1998., 11 pp.–11 pp.
- Sun, J., Yu, H., Zhong, G., Dong, J., Zhang, S., Yu, H., 2022a. Random shapley forests: cooperative game-based random forests with consistency. *IEEE Trans. Cybern.* **52**, 205–214.
- Sun, Z., Li, Q., Jin, S., Song, Y., Xu, S., Wang, X., Cai, J., Zhou, Q., Ge, Y., Zhang, R., Zang, J., & Jiang, D. (2022b). Simultaneous Prediction of Wheat Yield and Grain Protein Content Using Multitask Deep Learning from Time-Series Proximal Sensing. *Plant Phenomics*, 2022.
- Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., Bennett, M., 2017. Plant phenomics, from sensors to knowledge. *Curr. Biol.* **27**, R770–R783.
- Tasci, E., Zhuge, Y., Kaur, H., Camphausen, K., Krause, A.V., 2022. Hierarchical voting-based feature selection and ensemble learning model scheme for glioma grading with clinical and molecular characteristics. *Int. J. Mol. Sci.* **23**.
- Teshome, F., Bayabil, H.K., Hoogenboom, G., Schaffer, B., Singh, A., Ampatzidis, Y., 2023. Unmanned aerial vehicle (UAV) imaging and machine learning applications for plant phenotyping. *Comput. Electron. Agric.* **212**.
- Tian, Y., Zhou, D., Yao, X., Cao, W., Zhu, Y., 2007. Monitoring leaf nitrogen in rice using canopy reflectance spectra. *4th Int. Symp. . Intell. Inf. Technol. Agric.* **639**–649.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **8**, 127–150.
- Verrelst, J., Schaeppman, M.E., Koetz, B., Kneubuehler, M., 2008. Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sens. Environ.* **112**, 2341–2353.
- Wang, J., Wang, P., Tian, H., Tansey, K., Liu, J., Quan, W., 2023. A deep learning framework combining CNN and GRU for improving wheat yield estimates using time series remotely sensed multi-variables. *Comput. Electron. Agric.* **206**.
- Wang, K., Liu, X., Zhao, J., Gao, H., Zhang, Z., & Ieee, 2020a. Application research of ensemble learning frameworks. , Chin. Autom. Congr. (CAC) 5767–5772.
- Wang, K.K., Liu, X.D., Zhao, J.M., Gao, H.W., Zhang, Z., & Ieee, 2020b. Application research of ensemble learning frameworks. *Chin. Autom. Congr. (CAC)* 5767–5772.
- Wang, P. (2019). Interaction effect of nitrogen form and planting density on plant growth and nutrient uptake in maize seedlings (vol 18, pg 1120, 2019). *Journal of Integrative Agriculture*, 20, IV-IV.
- Wang, Q., Che, Y., Shao, K., Zhu, J., Wang, R., Sui, Y., Guo, Y., Li, B., Meng, L., Ma, Y., 2022. Estimation of sugar content in sugar beet root based on UAV multi-sensor data. *Comput. Electron. Agric.* **203**.
- Wang, Z.X., Nie, C.W., Wang, H.W., Ao, Y., Jin, X.L., Yu, X., Bai, Y., Liu, Y.D., Shao, M.C., Cheng, M.H., Liu, S.B., Wang, S.Y., Tuohuti, N., 2021. Detection and analysis of degree of maize lodging using UAV-RGB image multi-feature factors and various classification methods. *ISPRS Int. J. Geo-Inf.* **10**.
- Woebbecke, D.M., Meyer, G.E., Vonbargen, K., Mortensen, D.A., 1995. Color indexes for weed identification under various soil, residue, and lighting conditions. *Trans. Asae* **38**, 259–269.
- Xiao, S., Ye, Y., Fei, S., Chen, H., Zhang, B., Cai, J., Che, Z., Wang, Y., Ghafoor, Q., Bi, A., Shao, K., Wang, G., Guo, R., Li, Y., Zhang, B., Chen, Z., R., Ma, Y., 2023a. High-throughput calculation of organ-scale traits with reconstructed accurate 3D canopy structures using a UAV RGB camera with an advanced cross-circling oblique route. *Isprs J. Photogramm. Remote Sens.* **201**, 104–122.

- Xiao, S.F., Ye, Y.L., Fei, S.P., Chen, H.C., Zhang, B.Y., Li, Q., Cai, Z.B., Che, Y.P., Wang, Q., Ghafoor, A., Bi, K.Y., Shao, K., Wang, R.L., Guo, Y., Li, B.G., Zhang, R., Chen, Z., Ma, Y.T., 2023b. High-throughput calculation of organ-scale traits with reconstructed accurate 3D canopy structures using a UAV RGB camera with an advanced cross-circling oblique route. *Isprs J. Photogramm. Remote Sens.* **201**, 104–122.
- Xiao, Y., Dong, Y., Huang, W., Liu, L., & Ma, H. (2021). Wheat Fusarium Head Blight Detection Using UAV-Based Spectral and Texture Features in Optimal Window Size. *Remote Sensing*, **13**.
- Xie, C.Q., Yang, C., 2020. A review on plant high-throughput phenotyping traits using UAV-based sensors. *Comput. Electron. Agric.* **178**.
- Xu, C., Ding, J., Qiao, Y., Zhang, L., 2022. Tomato disease and pest diagnosis method based on the Stacking of prescription data. *Comput. Electron. Agric.* **197**.
- Yang, S., Gu, L., Li, X., Jiang, T., Ren, R., 2020. Crop Classification Method Based on Optimal Feature Selection and Hybrid CNN-RF Networks for Multi-Temporal Remote Sensing Imagery. *Remote Sens.* **12**.
- Yue, J., Yang, G., Tian, Q., Feng, H., Xu, K., Zhou, C., 2019. Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices. *ISPRS J. Photogramm. Remote Sens.* **150**, 226–244.
- Zhang, Q.M., Liu, P.S., Wang, X., Zhang, Y.Q., Han, Y., Yu, B., 2021. StackPDB: Predicting DNA-binding proteins based on XGB-RFE feature optimization and stacked ensemble classifier. *Appl. Soft Comput.* **99**.
- Zhang, X.Y., Zhao, J.M., Yang, G.J., Liu, J.G., Cao, J.Q., Li, C.Y., Zhao, X.Q., & Gai, J.Y. (2019). Establishment of Plot-Yield Prediction Models in Soybean Breeding Programs Using UAV-Based Hyperspectral Remote Sensing. *Remote Sensing*, **11**.
- Zhang, Y., Lin, M., Yang, Y., Ding, C., 2022. A hybrid ensemble and evolutionary algorithm for imbalanced classification and its application on bioinformatics. *Comput. Biol. Chem.* **98**, 107646–107648.
- Zhang, Z.X., Zhu, L.X., 2023. A Review on unmanned aerial vehicle remote sensing: platforms, sensors, data processing methods, and applications. *Drones* **7**.
- Zhou, L., Li, X.J., Zhang, B., Xuan, J., Gong, Y.L., Tan, C., Huang, H.G., & Du, H.Q. (2022). Estimating 3D Green Volume and Aboveground Biomass of Urban Forest Trees by UAV-Lidar. *Remote Sensing*, **14**.
- Zhou, X.M., Zhang, X.L., 2020. Individual tree parameters estimation for plantation forests based on UAV oblique photography. *IEEE Access* **8**, 96184–96198.
- Zian, S., Kareem, S.A., Varathan, K.D., 2021. An empirical evaluation of stacked ensembles with different meta-learners in imbalanced classification. *IEEE Access* **9**, 87434–87452.