Statistics and Applications {ISSN 2454-7395 (online)}
Volume 24, No. 1, 2026 (New Series), pp 275–296
http://www.ssca.org.in/journal

A Hyperspectral and Deep Learning Approach for Wheat Yield Prediction

Mohit Kumar[1,2], Alka Arora[2], Sudeep Marwaha[2], Viswanathan Chinnusamy[3], Sudhir Kumar[3], Soumen Pal[2], Mrinmoy Ray[2] and Rajkumar Dhakar[3]

[1] *The Graduate School,*
ICAR-Indian Agricultural Research Institute, New Delhi, 110012, India
[2] *ICAR-Indian Agricultural Statistics Research Institute, New Delhi, 110012, India*
[3] *ICAR-Indian Agricultural Research Institute, New Delhi, 110012, India*

Received: 04 June 2025; Revised: 10 October 2025; Accepted: 03 November 2025

Abstract

Accurate detection of wheat spikes and reliable yield prediction are critical for optimizing crop production and resource management. This study presents an integrated framework for spike detection and yield estimation using pseudo-RGB images derived from hyperspectral data. A YOLOv8 model was trained on 1,050 images, achieving high precision, recall, and mean average precision values. The bounding boxes and masks generated by YOLOv8 were used to quantify spike count and spike area, while six vegetation indices were extracted from hyperspectral images acquired at the booting stage. Three multiple linear regression models were developed for yield prediction: one based on spike features, another on vegetation indices, and a third combining both. The combined model achieved the highest accuracy, with a five-fold cross-validation R^2 of 0.902 $\pm$ 0.007, RMSE of 1.739 $\pm$ 0.133 g, and MAE of 1.289 $\pm$ 0.066 g. Compared with previous approaches, the proposed framework demonstrated improved performance, highlighting the value of integrating spike morphology and spectral data for yield prediction. Overall, the study shows that hyperspectral imaging can simultaneously provide morphological and physiological traits, reducing reliance on high-resolution RGB data in wheat phenotyping.

Key words: Deep learning; YOLOv8; Hyperspectral imaging; Yield prediction; Vegetation indices.

AMS Subject Classifications: 62K05, 05B05

1. Introduction

Wheat (*Triticum aestivum L.*) is one of the most widely cultivated cereal crops globally, serving as a staple food source and playing a central role in global food security (Curtis and Halford, 2014). Accurate and timely prediction of wheat yield is essential for agricultural planning, market forecasting, and food policy development. Traditional approaches for

Corresponding Author: Alka Arora
Email: : alka.office27@gmail.com

predicting yield are based on manual sampling and statistical modelling, which are labour-intensive, time-consuming, and often limited in spatial and temporal coverage (Kumar *et al.*, 2016; Sishodia *et al.*, 2020). Recent advances in remote sensing and image-based phenotyping have enabled non-destructive, scalable, and data-driven approaches to monitor crop traits and forecast yield (Khaki and Wang, 2019; Kumar *et al.*, 2024; Muruganantham *et al.*, 2022).

One of the most direct morphological indicators of wheat yield is the spike, which contains the grain-producing florets. Traits such as spike count and spike area are closely related to grain number and biomass, making them valuable phenotypic markers for yield prediction (Misra *et al.*, 2022; Qiongyan *et al.*, 2017). High-resolution RGB imagery captured by ground-based platforms or unmanned aerial vehicles (UAVs) has been widely used to detect and quantify wheat spikes using deep learning and computer vision techniques (Hasan *et al.*, 2018; Misra *et al.*, 2020; Zang *et al.*, 2022). A study by Misra *et al.* (2022) highlighted the potential of the spike count and spike area extracted from RGB images as predictors of wheat yield. However, these approaches are highly dependent on high-resolution RGB images, which can be computationally expensive and time-consuming due to large data volumes (Arora *et al.*, 2023; Dagar *et al.*, 2024). Furthermore, RGB-based methods often overlook important physiological parameters such as plant stress, canopy structure, and biochemical composition that are critical to yield evaluation (Zhou *et al.*, 2023).

The spectral indices derived from remote sensing data offer a valuable alternative for yield prediction, as they serve as representations of the physiological and structural traits of the plant. Indices such as the Normalized Difference Vegetation Index (NDVI), DVI, IPVI, and SAVI capture information on chlorophyll concentration, canopy density, and photosynthetic efficiency, all of which are closely related to biomass accumulation and grain yield (Aboelghar *et al.*, 2014; Xue and Su, 2017). Some studies demonstrated that NDVI is a reliable predictor of wheat yield (Liu *et al.*, 2023) and emphasize that combining multiple vegetation indices can improve prediction accuracy by capturing complementary aspects of crop physiology and canopy structure (Su *et al.*, 2023; Zhou *et al.*, 2023). However, a key limitation of spectral indices is that they do not directly capture spike-related morphological traits, such as spike count and spike area, which have a direct and strong correlation with grain yield.

To bridge this gap, integrating spectral and morphological information could lead to more comprehensive and accurate yield prediction models (Patrignani and Ochsner, 2015). While one option is to combine high-resolution RGB imagery with hyperspectral or multi-spectral data, this approach significantly increases data volume and processing complexity. Moreover, acquiring high-resolution RGB images requires additional equipment and setup. To address these challenges, this study proposes using hyperspectral imagery (HSI) alone to extract both morphological and physiological traits. Specifically, HSI data spanning the 400–1000 nm spectral range is used to generate low-resolution pseudo-RGB images for spike detection and counting. This approach eliminates the need for high-resolution RGB cameras and simplifies the workflow by leveraging a single data source for both morphological and spectral information. In addition to spike detection, the study proposes to compute multiple vegetation indices from the same HSI data and combine these with spike count and spike area to enhance yield prediction accuracy. This integrated approach contributes to the development of scalable, cost-effective phenotyping tools and explores the potential of HSI-derived imagery to complement or replace high-resolution RGB-based methods in precision

agriculture.

Recent advancements in deep learning have further enhanced plant trait extraction from imagery. YOLOv8, the state-of-the-art version of the You Only Look Once (YOLO) architecture, offers improved speed and accuracy in real-time object detection and instance segmentation (Jocher *et al.*, 2023). Its enhanced performance makes it suitable for agricultural applications where fine-scale object localization is needed. In wheat phenotyping, YOLO-based models have shown strong results for spike detection using aerial and ground-level RGB images (Fang and Yang, 2024; Wen *et al.*, 2024). Therefore, this study proposes the use of YOLOv8 to segment wheat spikes from pseudo-RGB images generated from hyperspectral data, enabling accurate spike count and area estimation that are the key features for yield modelling.

To provide an appropriate environment for evaluating the proposed approach, this study was conducted in a phenomics facility. Such facilities are designed for experiments where large numbers of genotypes are grown with replications and screened within a single season or multiple growing seasons, creating a demand for rapid and accurate assessment of yield-related traits. The integration of hyperspectral imaging with deep learning offers an automated, scalable, and precise framework that reduces the need for labour-intensive manual phenotyping and facilitates efficient yield prediction. By addressing the bottleneck of large-scale genotype screening, the methodology directly supports crop improvement and breeding programs. The specific objectives of this study are to: (1) Evaluate the feasibility of spike detection and counting using low-resolution pseudo-RGB images generated from HSI data. (2) Extract key morphological spike traits, particularly spike count and spike area. (3) Compute spectral reflectance indices from the hyperspectral data. (4) Assess the performance of combined morphological and spectral indices for predicting the wheat yield.

2. Materials and methods

The materials and methods section details the acquisition of a hyperspectral imaging (HSI) dataset of wheat and the extraction of key features for yield prediction. These features include spectral vegetation indices and spike-based morphological traits, all derived from the HSI data cube. Pseudo-RGB images were generated from the HSI cube and used for spike detection and segmentation by using the YOLOv8 model to extract spike morphological traits, spike count, and area. In parallel, spectral indices were computed directly from the hyperspectral bands to capture the physiological and spectral characteristics of the plants. A detailed workflow outlining the overall methodology is presented in Figure 1.

2.1. Dataset

The experiment was conducted at the Nanaji Deshmukh Plant Phenomics Centre (NDPPC), ICAR-IARI, Pusa, New Delhi, India, during three winter wheat-growing seasons (2021–22, 2022–23, and 2023–24). In phenomics facilities, wheat varieties and genotypes are grown under controlled environments that minimize variability and ensure uniform plant development.

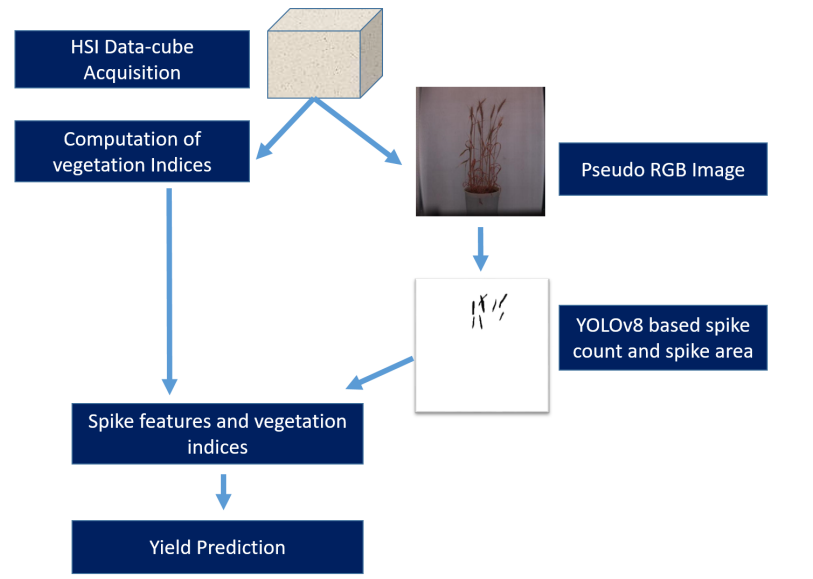

Figure 1: Workflow of the proposed methodology for wheat yield prediction using hyperspectral imagery

A total of 66 wheat varieties and genotypes were included, representing both released cultivars (*e.g.*, HD-2687, PBW343, LOK1, DBW51, Raj4037) and breeding lines (*e.g.*, Core-set accessions, LSP3043, KACHU#1, C518RF), and the detailed names are provided in Annexure Table 5. Each entry was grown in six replications, while the widely cultivated checks HD-2967, HD-3086, HD-3237, and HD-3271 were represented with more than six replications to strengthen comparisons. The six replications were grown under contrasting conditions, comprising two non-stressed controls, two subjected to nitrogen deficiency, and two exposed to drought stress, enabling robust assessment of genotypic performance across different environmental conditions.

Across the three years, the dataset comprised 440 plants in 2023–24, 90 plants in 2022–23, and 90 plants in 2021–22, giving a total of 620 plants. Each plant was imaged at two developmental stages, booting and physiological maturity, producing 1240 hyperspectral images (620 plants $\times$ 2 stages).

Hyperspectral imaging was performed using a Headwall E-Series VNIR camera, operating in band-sequential format (BSL). The system covered the 400–1000 nm spectral range and recorded reflectance across 940 spectral bands, providing detailed information on canopy structure, pigment composition, and physiological traits. The dataset served two main purposes: (i) computation of spectral vegetation indices from booting-stage images, and (ii) conversion of maturity-stage hyperspectral data into pseudo-RGB images for spike detection.

For spike detection model development, 350 maturity-stage images were used for training. The trained model was applied to the remaining 270 maturity-stage test images to detect and segment spikes, ensuring that the images used for yield prediction were completely independent of the training set. Spike features extracted from these test images were then combined with vegetation indices derived from the booting stage and used in the regression analysis for yield prediction.

2.2. Data pre-processing

One of the essential steps after acquiring raw hyperspectral data is to perform geometric and radiometric correction to ensure data consistency and accuracy. Radiometric calibration was performed using Equation 1 (Xue and Su, 2017). Where r_0 represents the corrected reflectance data, r_{raw} denotes the raw hyperspectral data, r_{dark} is the dark reference, and r_{white} refers to the reference white panel data. The Lemnatech Imaging System captures dark reference frames (shutter closed, no light) and white reference frames (Spectralon panel, 99% reflectance) for each imaging cycle. These references are directly linked to the corresponding plant images, ensuring that environmental noise, illumination variability, and sensor drift are minimized.

$$r_0 = \frac{r_{raw} - r_{dark}}{r_{white} - r_{dark}} \tag{1}$$

This calibration step normalized reflectance values across all images, correcting for variations caused by environmental factors and instrumental inconsistencies. Dark frame subtraction reduced electronic noise and sensor offsets, while normalization with the white reference ensured spectral consistency across sessions. To assess image quality, hyperspectral cubes were visually inspected for striping artifacts, saturation, or band misalignment. As the imaging was performed in a standardized LemnaTec facility, no defective images were identified. Additionally, signal-to-noise ratios (SNR) were monitored across representative bands (*e.g.*, 550, 680, 800 nm), confirming spectral stability and reliability throughout the dataset. In the next step, pseudo-RGB images generated from spectral bands (Figure 2) were used for the spike detection and morphological trait extraction (spike count and projected spike area). The image had a width of 400 pixels and a height of 348 pixels.

Figure 2: Pseudo RGB of the wheat plant generated from HSI data

2.3. Spike detection and segmentation

YOLOv8 (Varghese and Sambath, 2024), a state-of-the-art deep learning framework, was used for wheat spike detection and segmentation. This latest version of YOLO intro-

duces architectural improvements such as anchor-free detection, dynamic label assignment, and transformer-based feature aggregation (Jocher *et al.*, 2023; Fang and Yang, 2024), making it well-suited for complex plant phenotyping tasks. In this study, pseudo-RGB images generated from hyperspectral data were provided as input to YOLOv8, and the model was trained on a manually annotated dataset to ensure accurate localization and mask generation.

The YOLOv8-seg architecture optimizes a composite loss function consisting of (i) Complete IoU (CIoU) loss for bounding box regression, (ii) Binary Cross-Entropy (BCE) loss for objectness confidence, and (iii) BCE loss for class prediction. For segmentation, an additional mask loss is computed as the pixel-wise BCE between predicted and ground-truth masks. These loss components collectively balance localization accuracy, object classification, and mask generation. Weighting of the losses followed the YOLOv8 default configuration (box: 7.5, cls: 0.5, dfl: 1.5), empirically optimized by Ultralytics for robust detection and segmentation. The YOLOv8 model was trained for 150 epochs with a batch size of 16, an initial learning rate of 0.001, and the stochastic gradient descent (SGD) optimizer with a momentum of 0.937. Anchor-free detection and dynamic label assignment were enabled by default, and early stopping with a patience of 30 epochs was applied to prevent over-fitting.

2.4. Dataset preparation

The VGG Image Annotator (VIA) tool (Dutta and Zisserman, 2019), developed by the Visual Geometry Group at the University of Oxford, was used to manually label pseudo-RGB images to make the training dataset for wheat spike detection and segmentation. VIA is a web-based, open-source annotation tool that supports polygon, rectangle, and point-based annotations, making it well-suited for detailed labelling of images. Each wheat spike was manually outlined using a polygonal to capture its shape precisely (Figure 3a). All annotated spikes were labelled under the class "spike" to support both detection and segmentation tasks in YOLOv8. Before annotation, the dataset was split into a training set of 350 images and a test set of 270 images. The training images (Figure 3b) were augmented through horizontal and vertical rotations (Figure 3c), horizontal flipping (Figure 3d), and brightness adjustments (Figure 3e), thereby enhancing data diversity and expanding the training dataset to 1,050 images. The final annotations were exported in COCO JSON format, which is compatible with the YOLOv8 training pipeline and includes image IDs, polygon coordinates, and class labels.

2.5. Evaluation metrics for spike detection and segmentation

The performance of the YOLOv8 model in spike detection and segmentation was assessed using standard object detection and segmentation evaluation metrics (Everingham *et al.*, 2010; Lin *et al.*, 2014).

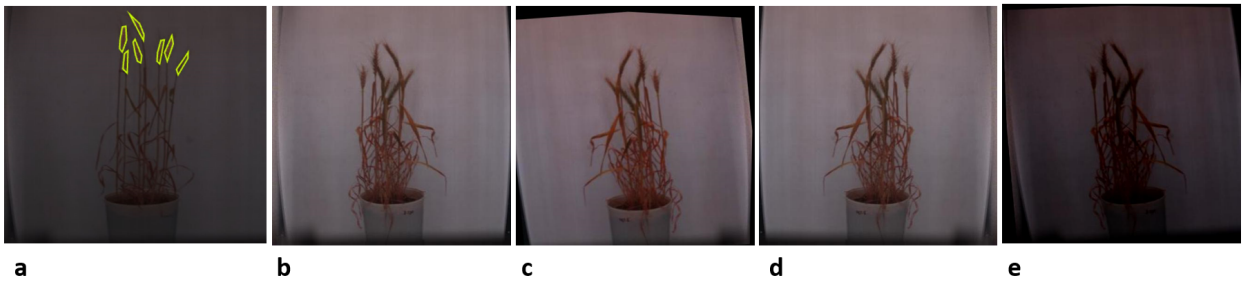

Figure 3: Illustration of the annotation and augmentation process for wheat spike detection. (a) Manual polygonal annotation of wheat spikes using the VIA tool; (b) sample pseudo-RGB image; (c) rotation augmentation; (d) horizontal flip; (e) brightness adjustment. These steps expanded the training dataset to 1,050 images and ensured greater variability for robust YOLOv8 model training

2.5.1. Spike detection metrics

Precision (P) measures the proportion of correctly identified spikes (true positives) among all spikes predicted by the model.

$$P = \frac{TP}{TP + FP} \tag{2}$$

Recall (R) measures the proportion of actual spikes correctly identified by the model.

$$R = \frac{TP}{TP + FN} \tag{3}$$

F1 Score is the harmonic mean of precision and recall, providing a balanced measure of accuracy.

$$F1 = \frac{2 \times (P \times R)}{P + R} \tag{4}$$

Mean Average Precision (mAP) evaluates detection performance across multiple Intersection over Union (IoU) thresholds. It is computed by averaging the precision over recall levels and then averaging across classes (in this case, spikes).

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \tag{5}$$

2.5.2. Spike segmentation metrics

The *Jaccard Index*, commonly referred to as Intersection over Union (IoU), is a widely used metric for evaluating the accuracy of segmentation models. It quantifies the overlap between the predicted segmentation mask and the ground truth mask. The *Dice coefficient* is a metric for overlap between predicted and ground truth segmentation masks.

$$JaccardIndex = \frac{A \cap B}{A \cup B} \tag{6}$$

$$Dice = \frac{2 \cdot |A \cap B|}{|A| + |B|} \tag{7}$$

Where A is the set of pixels (or area) in the predicted spike mask. B is the set of pixels (or area) in the ground truth spike mask.

2.6. Features calculation for yield prediction

2.6.1. Spike count and spike area

The trained YOLOv8 model was applied to the test set of pseudo-RGB images to detect and segment wheat spikes. The model generated binary masks and bounding boxes for each spike. The number of spikes in each image was counted, and the area of each spike was calculated by counting the pixels in its mask. For each image, the spike count and total spike area were recorded, creating a dataset for yield prediction.

2.6.2. Spectral vegetation indices

For the HSI images collected at the booting stage, six spectral reflectance indices were calculated (Table 1). The selection of these indices was based on their proven association with wheat growth dynamics and yield prediction in earlier studies (Aboelghar *et al.*, 2014). Vegetation indices such as NDVI, RVI, SAVI, DVI, and IPVI are widely recognized for their ability to capture canopy vigour, chlorophyll content, and biomass accumulation, which are physiological traits that strongly correlate with final grain yield. The booting stage was specifically chosen because it represents a critical developmental phase where the crop's photosynthetic capacity and canopy structure are highly indicative of yield potential. Previous studies have demonstrated that indices derived at this stage can reliably predict wheat yield at both field and regional scales (Ren *et al.*, 2008; Sultana *et al.*, 2014; Xie *et al.*, 2020). These works collectively establish that early-stage spectral information complements later-stage morphological traits and thereby enhances predictive accuracy.

2.7. Yield prediction based on computed features

Grain yield was determined individually for 270 plants included in the test set. The dataset comprised 90 plants per year across three consecutive growing seasons (2021–22, 2022–23, and 2023–24). Each plant was harvested separately, and the yield was recorded using a precision balance to ensure measurement accuracy.

An ablation study was designed to evaluate the relative contribution of morphological and spectral features to yield prediction. Three feature sets were derived from the hyperspectral data: (i) morphological traits obtained from maturity-stage images, specifically spike count and spike area; (ii) spectral features represented by six vegetation indices computed from booting-stage hyperspectral imagery (Table 1); and (iii) a combined feature set in-

Table 1: Spectral reflectance indices and their formulas

Spectral reflectance index	Formula	References
NDVI	$\frac{NIR - Red}{NIR + Red}$	Baret and Guyot (1991)
SAVI	$\frac{(800\,\text{nm} - 670\,\text{nm})}{(800\,\text{nm} + 670\,\text{nm} + 0.5)}(1 + 0.5)$	Huete (1988)
GVI	$\frac{NIR - Green}{NIR + Green}$	Aboelghar *et al.* (2014)
DVI	$NIR - Red$	Richardson and Everitt (1992)
RVI	$\frac{NIR}{Red}$	Jordan (1969)
IPVI	$\frac{NIR}{NIR + Red}$	Crippen (1990)

tegrating both morphological and spectral traits to capture complementary structural and physiological information.

For each feature set, a separate regression model was constructed: Model 1 (morphological features only), Model 2 (spectral features only), and Model 3 (combined features). All models were developed using multiple linear regression, and model training was performed under a five-fold cross-validation scheme to obtain robust estimates of predictive accuracy and reduce the risk of over-fitting.

Uncertainty in predictions was quantified using 95% prediction intervals derived from the fitted regression models. These intervals provide estimates of the range within which future yield observations are expected to fall, incorporating both model error and residual variability. To address potential multicollinearity among vegetation indices, LASSO regression was used as a regularized comparator. This approach penalizes redundant predictors, thereby facilitating feature selection and improving model stability.

Model performance was evaluated using three standard metrics: the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Comparative analysis of the ablation models, the LASSO regression, and the prediction interval outputs enabled a rigorous assessment of the relative importance of individual feature sets and their combined predictive capacity.

3. Results

3.1. Spike detection and segmentation

The YOLOv8 model was trained using pseudo-RGB images derived from hyperspectral data for the task of wheat spike detection and segmentation. Model training was conducted on an NVIDIA V100 GPU with 32 GB RAM, ensuring adequate computational resources for efficient learning and convergence. The training dataset consisted of 1,050 augmented images, while 270 images were reserved for testing. During model training, 20% of the training images (210 images) were used as the validation set.

Model performance was monitored through distinct sets of loss and evaluation metrics,

which are presented in five categories: (i) training loss, (ii) validation loss, (iii) box and mask precision/recall, and (iv) mAP@50 and mAP@50:95 (v) learning rate (Figures 4–8).

(i) Training loss

The training loss curves, including box loss, segmentation loss, classification loss, and distribution focal loss (DFL), showed a consistent downward trend across epochs, demonstrating that the model converged smoothly during training (Figure 4).

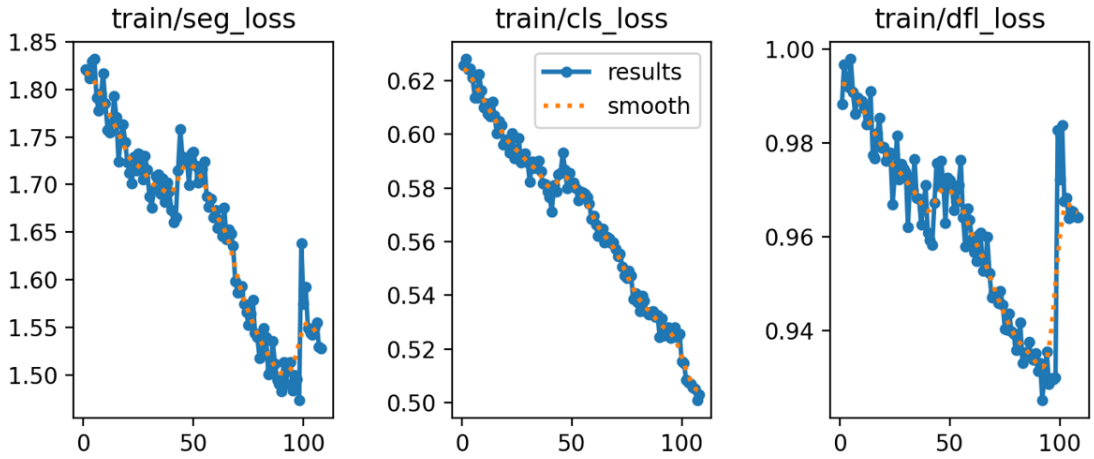

Figure 4: Training loss curves of the YOLOv8 model, including segmentation loss, classification loss, and distribution focal loss (DFL)

(ii) Validation loss

Validation losses for box regression, segmentation, and classification exhibited some fluctuations but remained overall stable throughout the training process, reflecting good generalization of the model to unseen data (Figure 5).

(iii) Box and mask precision and recall

Evaluation metrics for bounding box (B) and mask (M) predictions showed high reliability in spike detection. Precision and recall values ranged between 0.85 to 0.88 for bounding boxes and 0.82 to 0.87 for masks, indicating consistent and accurate predictions across both detection and segmentation tasks (Figure 6).

(iv) mAP@50 and mAP@50:95

The mean average precision at IoU=0.50 (mAP50) reached approximately 0.86 for bounding boxes and 0.83 for masks, while the stricter mAP@50:95 values were slightly lower but remained consistent, confirming robust detection performance across multiple IoU thresholds (Figure 7).

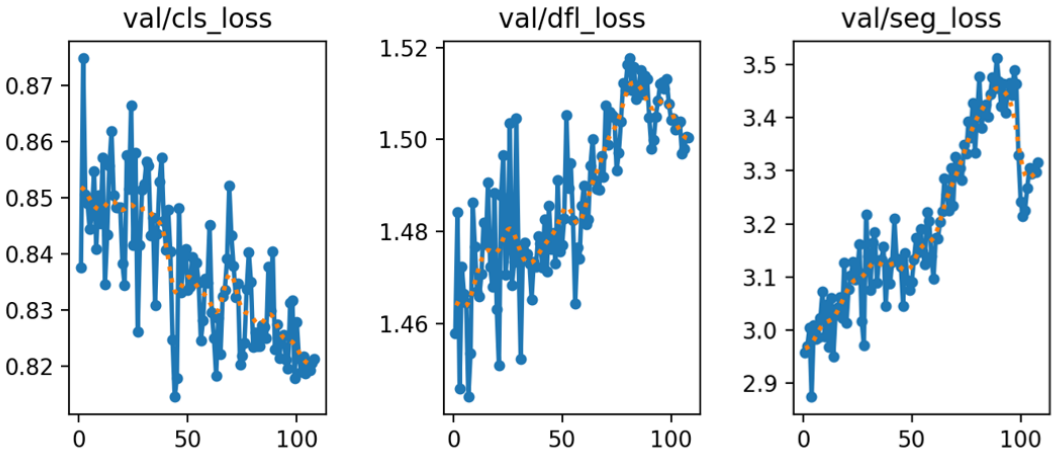

Figure 5: Validation loss curves of the YOLOv8 model, including classification loss, distribution focal loss (DFL), and segmentation loss

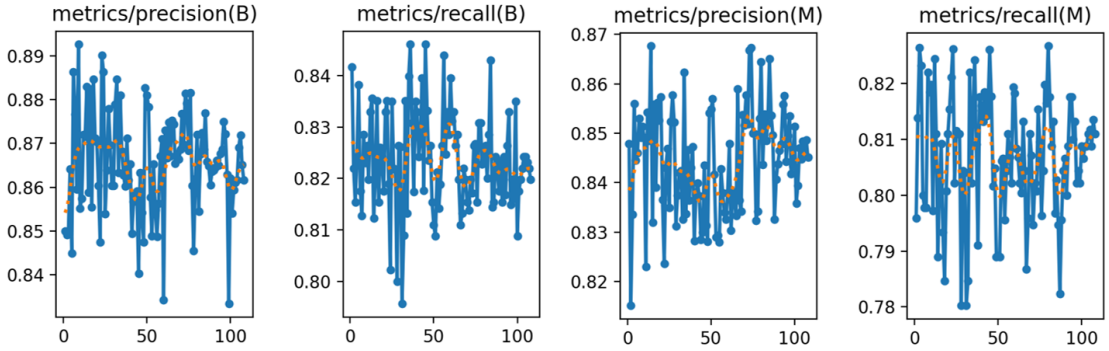

Figure 6: Precision and recall metrics for bounding box (B) and mask (M) predictions during YOLOv8 training, demonstrating high reliability in wheat spike detection and segmentation

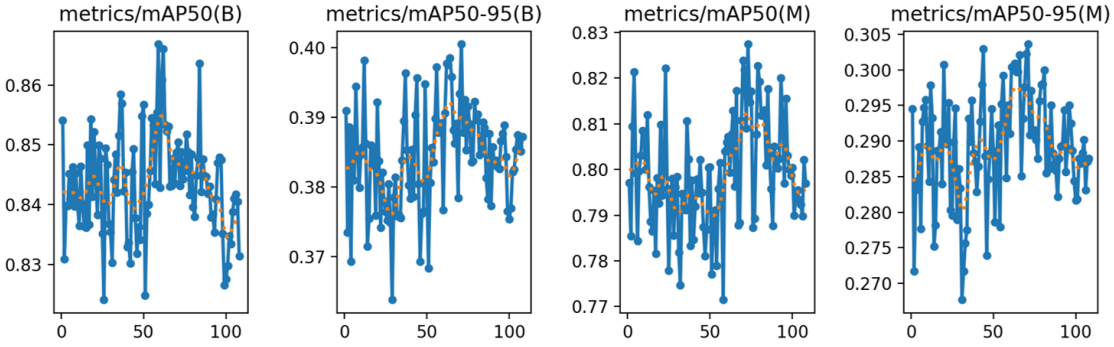

Figure 7: Mean average precision at IoU thresholds (mAP@50 and mAP@50:95) for bounding box (B) and mask (M) predictions

(v) Learning rate

The learning rate schedules for the parameter groups pg0, pg1, and pg2, which correspond to different layers of the YOLOv8 architecture (backbone, neck, and head), exhibited a smooth decay pattern throughout training (Figure 8). A stable and gradually decreasing learning rate is critical to avoid oscillations in the loss landscape and ensures convergence toward an optimal solution. The smooth decay observed across all parameter groups indicated that the optimization process was well-regularized, reducing the likelihood of over-fitting while improving generalization on unseen data..

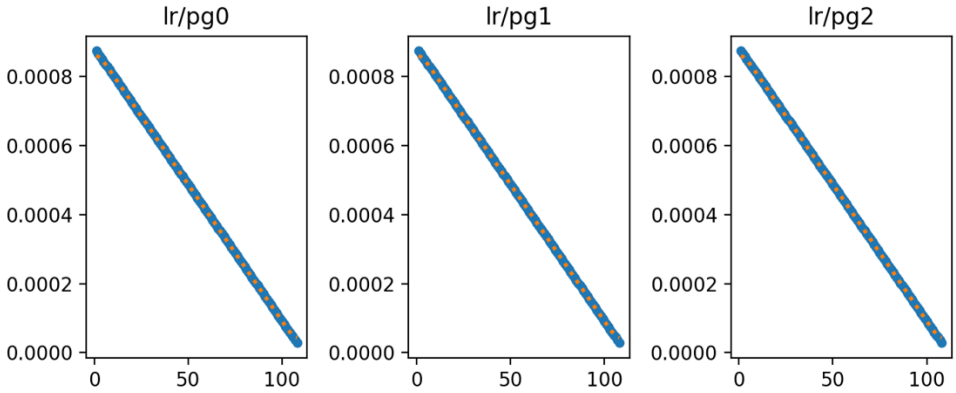

Figure 8: Learning rate schedules for the parameter groups pg0, pg1, and pg2 in the YOLOv8 model

Following training, the YOLOv8 model was evaluated on the test set, with results summarized in Table 2. The model achieved a precision of 0.83, a recall of 0.89, and an F1 score of 0.86, reflecting reliable detection performance. The mAP@50 was 0.89, while the Dice coefficient (0.94) and Jaccard Index (0.88) further confirmed the high accuracy of segmentation. These findings demonstrate that the trained YOLOv8 model effectively learned discriminative features for wheat spike detection and segmentation, achieving robust performance on unseen test images.

Table 2: YOLOv8 evaluation metrics on the test set

Metric	Precision	Recall	F1 Score	mAP@50	Dice coefficient	Jaccard Index
Value	0.83	0.89	0.86	0.89	0.94	0.88

Figure 9 presents qualitative examples of the model's predictions on four test images, where bounding boxes (B) and masks (M) are superimposed on the spikes. The results show that the model consistently detected and segmented wheat spikes, with bounding boxes and masks closely aligned with actual spike regions. The qualitative outputs complement the quantitative performance metrics in Table 2, together confirming that the YOLOv8 model can reliably detect and segment wheat spikes, which is essential for accurate spike counting and area estimation.

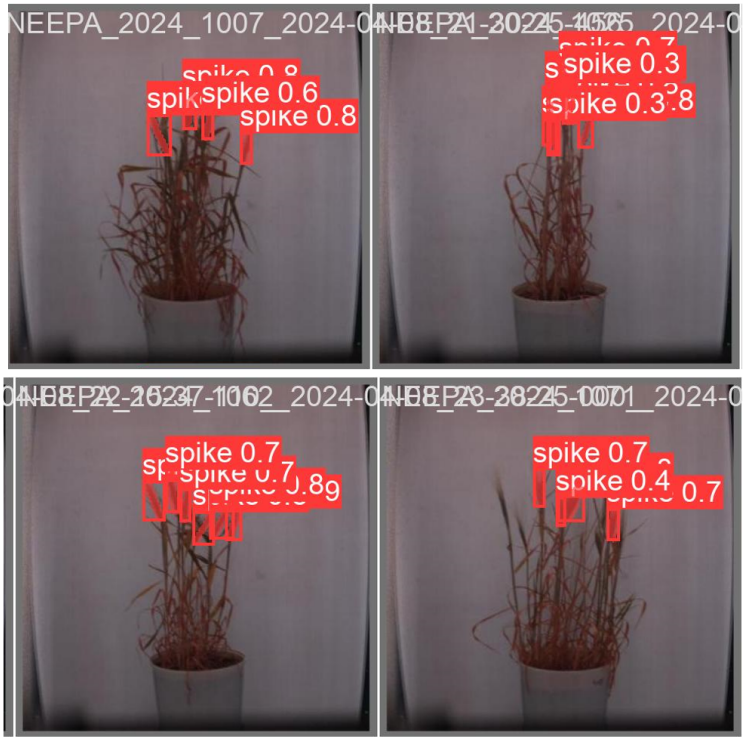

Figure 9: YOLOv8 predictions on four test images, showing bounding boxes and segmentation masks superimposed on wheat spikes. The model accurately localized and segmented spikes, aligning closely with ground-truth regions

The results from training, validation, and test metrics indicate that the model successfully captured the structural characteristics of wheat spikes, enabling its reliable application for downstream spike trait analysis. The YOLO model demonstrated computational efficiency, requiring only approximately 6 milliseconds per image on an NVIDIA V100 GPU with 32 GB RAM, while the same task took 0.5 seconds per image on an Intel Core i7 processor with 16 GB RAM, evaluated over a set of 200 images for inferencing. Although GPUs are typically used for model training, inference is often performed on CPUs in practical scenarios; thus, the CPU performance underscores the model's feasibility for downstream applications under standard computational environments, ensuring faster predictions suitable for high-throughput analysis.

3.2. Yield prediction

The trained YOLOv8 model was applied to a test set of 270 pseudo-RGB images to extract spike-based features for yield prediction. These features included the number of spikes per plant and the total projected spike area. In parallel, vegetation indices were computed from the booting stage hyperspectral images, providing spectral indicators of plant health and canopy structure. After harvesting, the grain yield (in grams) was recorded for the 270 test set plants, allowing for a comparison of model performance. The descriptive statistics of the features and plant yield are provided in Table 3.

Table 3: Descriptive statistics of the features for the yield prediction

Statistic	NDVI	SAVI	GVI	DVI	RVI	IPVI	Spike count	Spike area	Yield (g)
Min	0.61	0.46	0.44	0.46	3.10	0.78	0	0	0
Max	0.73	0.59	0.59	0.61	4.36	0.85	12	12916	36
Mean	0.64	0.49	0.49	0.50	3.33	0.80	7.35	7549.62	17.92
SD	0.02	0.02	0.02	0.02	0.19	0.01	2.32	2343.84	5.61

Figure 10 presents the Pearson correlation heat map illustrating the relationships among vegetation indices, spike features, and grain yield. Spike area and spike count exhibit strong positive correlations with grain yield, with spike area showing the highest correlation coefficient (r = 0.66). This highlights their critical role in yield prediction. Among the vegetation indices, NDVI, SAVI, DVI, RVI, and IPVI demonstrate moderate positive correlations with yield (r = 0.52–0.62), confirming their utility as spectral predictors. Notably, the vegetation indices themselves are highly intercorrelated, particularly SAVI, DVI, RVI, and IPVI ($r > 0.70$), suggesting redundancy among some spectral features. In contrast, the weak or negative correlations observed between vegetation indices and spike features indicate that these two groups of predictors provide largely complementary information, which may improve model robustness when integrated for yield estimation.

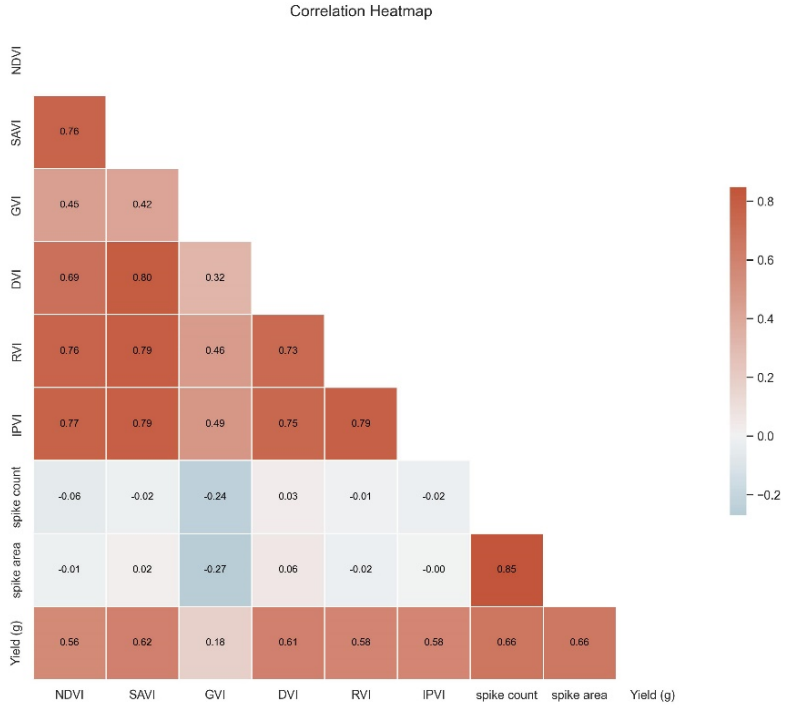

Figure 10: Pearson correlation heat-map showing relationships among vegetation indices, spike features, and grain yield (g)

Three regression models were constructed to evaluate the predictive ability of spike-based morphological features and hyperspectral vegetation indices for wheat yield estimation. The first model used spike traits (spike count and spike area), the second relied solely on vegetation indices, and the third integrated both feature types.

The spike feature-based model (Figure 11a) showed a moderate relationship between predicted and actual yield, with an R^2 of 0.453 ± 0.068, an RMSE of 4.1 ± 0.30 g, and an MAE of 3.291 ± 0.194 g (Table 4). The 95% prediction intervals were relatively wide, indicating variability in prediction accuracy when using morphological traits alone.

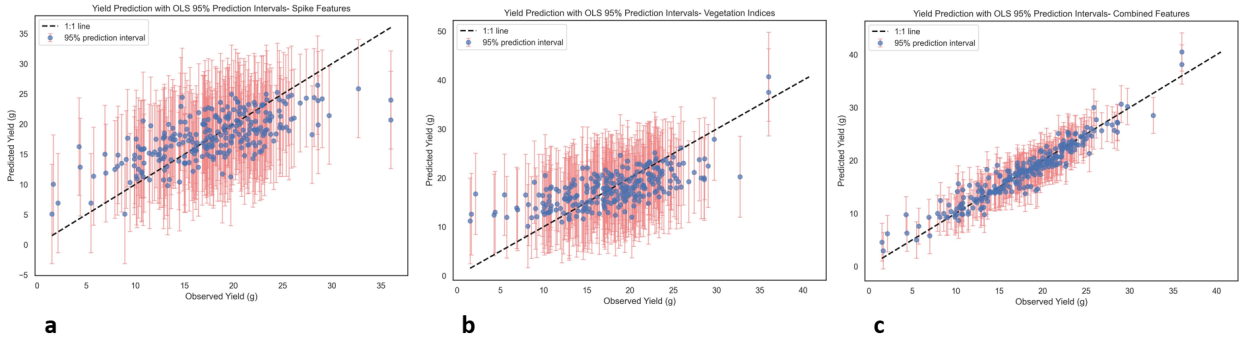

Figure 11: Actual versus predicted yield plot for (a) based on spike features, (b) based on vegetation indices, (c) based on combined spike and vegetation indices features

The vegetation index-based model (Figure 11b) demonstrated slightly weaker predictive performance, with an R^2 of 0.416 ± 0.036, an RMSE of 4.247 ± 0.284 g, and an MAE of 3.377 ± 0.321 g. The prediction intervals in this model were also broad, reflecting greater uncertainty and scatter around the 1:1 line.

Table 4: Cross-validation results (mean ± standard deviation) for yield prediction using vegetation indices, spike features, and their combination across three years of data

Model	R^2 (mean ± std)	RMSE (mean ± std)	MAE (mean ± std)
Indices-only	0.416 ± 0.036	4.247 ± 0.284	3.377 ± 0.321
Spikes-only	0.453 ± 0.068	4.100 ± 0.300	3.291 ± 0.194
Combined	0.902 ± 0.007	1.739 ± 0.133	1.289 ± 0.066

By contrast, the combined model integrating both vegetation indices and spike features (Figure 11c) achieved the best performance, with an R^2 of 0.902 ± 0.007, an RMSE of 1.739 ± 0.133 g, and an MAE of 1.289 ± 0.066 g. Importantly, the prediction intervals in this model were much narrower and closely aligned with the 1:1 line, demonstrating both improved accuracy and reduced uncertainty in yield estimation. To further examine prediction reliability, an error analysis was conducted across the test dataset. Overall, most predictions closely matched the observed yields, with residuals distributed symmetrically around zero. However, larger deviations were evident at the extremes of the yield range, particularly for plants with very low or very high grain yield.

This pattern can be clearly observed in the scatter plots of predicted versus observed yield (Figures 11a–c), where the data points at the lower and upper ends deviate more widely from the 1:1 line. These findings indicate that while the combined model achieved

high overall accuracy, prediction uncertainty increased for outlier cases, likely due to the limited representation of extreme-yielding plants in the training dataset.

As observed in the correlation heat-map (Figure 10), vegetation indices were highly correlated with each other, indicating potential multicollinearity that could affect model stability. To address this, we further applied Lasso regression for feature selection and coefficient shrinkage (Figure 12). The analysis confirmed that spike traits were the dominant predictors, with spike count (2.41) and spike area (1.66) showing the largest coefficients. Among vegetation indices, SAVI (1.04), DVI (0.85), and NDVI (0.80) retained moderate importance, while RVI (0.66), IPVI (0.48), and GVI (0.40) contributed smaller but still non-zero effects. Notably, none of the coefficients were reduced to zero, suggesting that despite their high interdependence, all vegetation indices added complementary spectral information for yield prediction.

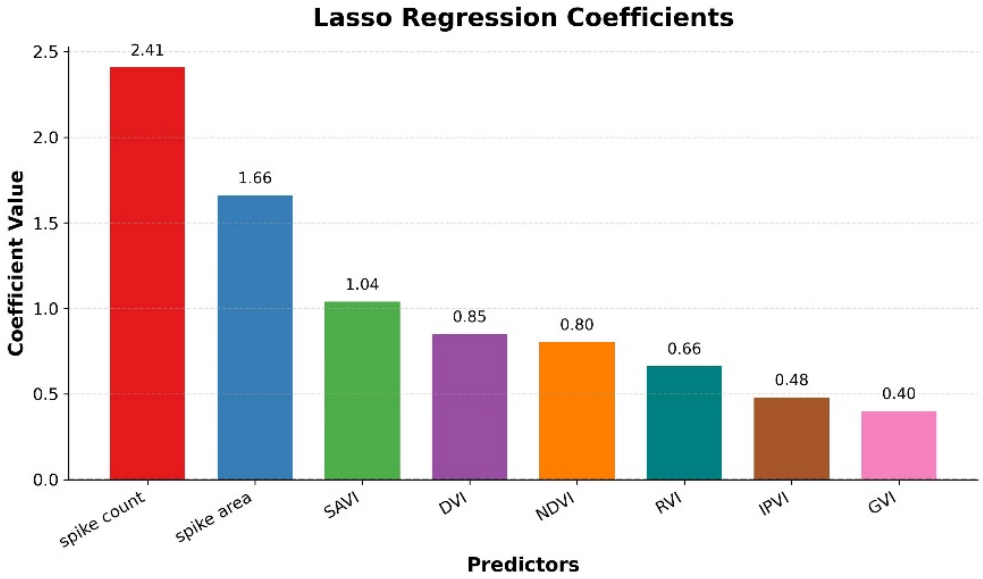

Figure 12: LASSO regression coefficients showing the relative importance of spike traits and vegetation indices

In the post-lasso OLS analysis, each predictor was evaluated under the null hypothesis that its coefficient equals zero ($H_0 : \beta = 0$), against the alternative hypothesis that the coefficient is non-zero ($H_1 : \beta \neq 0$). The null hypothesis was rejected for all predictors, confirming that NDVI, SAVI, GVI, DVI, RVI, IPVI, spike count, and spike area each exerted a statistically significant influence on yield. These findings indicate that both spike-related traits and spectral vegetation indices contribute meaningfully to yield prediction, with the integration of both trait- and index-based information enhancing model robustness.

Taken together, the linear regression and Lasso analyses demonstrate that spike-derived morphological traits explain the majority of yield variability, while vegetation indices enhance predictive robustness by providing additional spectral cues. Overall, the results (Table 4, Figures 11–12) highlight the strength of multi-modal data fusion for precise, reliable, and high-throughput wheat yield estimation.

4. Discussion

Recent studies have predominantly utilized high-resolution RGB images for plant phenotyping and yield estimation. In contrast, this study demonstrates that the YOLOv8 model can achieve strong performance using low-resolution pseudo-RGB images generated from hyperspectral data. This is particularly noteworthy, as it underscores the model's robustness and the potential for reducing reliance on high-resolution imaging systems, thereby making automated crop analysis more scalable and accessible by reducing computation time. The study presented an approach for wheat spike detection and yield prediction using YOLOv8, trained on pseudo-RGB images. The model effectively detected and segmented wheat spikes across a diverse set of images, supporting its application for spike counting and area estimation. These features were subsequently used for yield modelling, in combination with vegetation indices derived from the hyperspectral data cube. To assess the relative contribution of different features, three regression models were developed: one based solely on spike features, another on vegetation indices, and a third combining both. While the individual models showed reasonable predictive capacity, the combined model outperformed them, highlighting the synergistic value of integrating both morphological and spectral features. This fusion significantly improved yield prediction accuracy and demonstrated that spike morphology and vegetation indices contribute complementary information. Similar findings have been reported in previous studies (Al-Gaadi *et al.*, 2016; Xiong *et al.*, 2020). A comparison with existing literature further confirms the strength of our approach. Global Wheat Head Dataset (David *et al.*, 2020) used by Fang and Yang (2024) reported effective spike detection using high-resolution RGB data. Our model not only achieved comparable or better performance but did so using lower-resolution pseudo-RGB inputs. Similarly, our results surpassed the accuracy reported by Fang and Yang (2024); Wen *et al.* (2024), further validating the model's reliability. In the context of yield prediction, our approach showed notable improvements over Misra *et al.* (2022), where only RGB image-based spike traits were used.

This study demonstrated that combining hyperspectral imaging with deep learning provides a powerful framework for non-destructive wheat phenotyping and yield prediction. By integrating both spectral vegetation indices and spike-based morphological traits, the approach captures a holistic representation of plant health and productivity. Importantly, the findings are directly relevant to high-throughput phenomics facilities, where large numbers of genotypes are evaluated, as the proposed framework can substantially reduce the time and effort required for yield estimation. The present work was carried out under controlled conditions using pseudo-RGB images derived from hyperspectral data, with a limited dataset and without the complexity introduced by variable illumination, heterogeneous backgrounds, or environmental noise. While this represents a challenge for scaling the approach to large-scale field applications, it is not a limitation within phenomics facilities where standardized conditions prevail. In this context, the approach is immediately scalable for large-scale phenotyping and will be released as a software framework that can be seamlessly integrated into existing phenomics platforms.

Nonetheless, translating this framework to field-scale deployment will require addressing additional challenges, including illumination variability, canopy occlusion, soil–plant spectral interactions, and larger spatial heterogeneity. Illumination variability can be mitigated through radiometric calibration and deep learning approaches that learn illumination-

invariant features (Misra *et al.*, 2020). Soil and heterogeneous background effects can be reduced using soil-adjusted vegetation indices (such as SAVI, MSAVI, or TSAVI) or soil-line–based corrections, while topographic influences can be minimized with terrain-based corrections, including C-correction or Minnaert correction. Future work should therefore extend the methodology to field-based hyperspectral imaging under diverse agro-climatic conditions, incorporate auxiliary variables such as soil properties, weather data, and management practices, and integrate data from multiple sensing platforms. Coupled with advances in deep learning architectures and hardware efficiency, these efforts could enable robust, real-time, in-field monitoring of crop performance, thereby enhancing scalability and contributing significantly to precision agriculture.

5. Conclusion

This study developed and validated a YOLOv8-based framework for wheat spike detection and yield prediction using pseudo-RGB images derived from hyperspectral data. The model achieved high accuracy in spike detection and segmentation, demonstrating that reliable spike traits such as count and area can be extracted even from low-resolution pseudo-RGB inputs. When integrated with vegetation indices, these morphological features significantly enhanced yield prediction, with the combined regression model achieving superior performance ($R^2 = 0.902$), surpassing models based on individual feature sets. The framework was further strengthened through rigorous validation, including error analysis and LASSO regression, which confirmed the dominance of spike traits while highlighting the complementary role of vegetation indices. By deriving both morphological and physiological traits from hyperspectral data, the approach reduces dependency on high-resolution RGB systems and provides a cost-effective, scalable solution for high-throughput phenotyping. Importantly, this framework addresses key challenges in phenomics facilities, where large numbers of genotypes are routinely evaluated, by enabling rapid and automated yield estimation. Beyond its immediate utility, the approach offers a pathway toward accelerating breeding programs, improving genetic gain, and contributing to the broader goal of sustainable wheat production.

Acknowledgements

The facilities provided by the Indian Council of Agricultural Research - Indian Agricultural Statistics Research Institute (ICAR- IASRI), New Delhi, and the funding granted to the first author by ICAR in the form of an ICAR-SRF fellowship are duly acknowledged for carrying out this study, which is a part of his doctoral research being pursued at ICAR-IASRI. In addition, thanks to the NDPPC, ICAR-IARI, New Delhi, for conducting the experiment and providing the data for the study.

Conflict of interest

The authors do not have any financial or non-financial conflict of interest to declare for the research work included in this article.

Code and data availability

The processed datasets (pseudo-RGB images, vegetation index values, and yield records) and the YOLOv8 training scripts with analysis codes are available from the corresponding author upon reasonable request.

References

Aboelghar, M., Ali, A.-R., and Arafat, S. (2014). Spectral wheat yield prediction modeling using spot satellite imagery and leaf area index. *Arabian Journal of Geosciences*, **7**, 465–474.

Al-Gaadi, K. A., Hassaballa, A. A., Tola, E., Kayad, A. G., Madugundu, R., Alblewi, B., and Assiri, F. (2016). Prediction of potato crop yield using precision agriculture techniques. *PLOS ONE*, **11**, e0162219.

Arora, A., Misra, T., Kumar, M., Marwaha, S., Kumar, S., and Chinnusamy, V. (2023). Computer vision approaches for plant phenotypic parameter determination. In *Digital Ecosystem for Innovation in Agriculture*, 263–270. Springer Nature Singapore, Singapore.

Baret, F. and Guyot, G. (1991). Potentials and limits of vegetation indices for lai and apar assessment. *Remote Sensing of Environment*, **35**, 161–173.

Crippen, R. E. (1990). Calculating the vegetation index faster. *Remote Sensing of Environment*, **34**, 71–73.

Curtis, T. and Halford, N. G. (2014). Food security: the challenge of increasing wheat yield and the importance of not compromising food safety. *Annals of Applied Biology*, **164**, 354–372.

Dagar, P., Arora, A., Ray, M., Kumar, S., Chourasia, H., Kumar, M., and Chinnusamy, V. (2024). High-resolution reconstruction of images for estimation of plant height in wheat using rgb-d camera and machine learning approaches. *Current Science*, **127**, 1440–1446.

David, E., Madec, S., Sadeghi-Tehran, P., Aasen, H., Zheng, B., Liu, S., Kirchgessner, N., Ishikawa, G., Nagasawa, K., Badhon, M. A., Pozniak, C., de Solan, B., Hund, A., Chapman, S. C., Baret, F., Stavness, I., and Guo, W. (2020). Global wheat head detection (gwhd) dataset: A large and diverse dataset of high-resolution rgb-labelled images to develop and benchmark wheat head detection methods. *Plant Phenomics*, 3521852.

Dutta, A. and Zisserman, A. (2019). The via annotation software for images, audio and video. In *Proceedings of the 27th ACM International Conference on Multimedia*, 2276–2279.

Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., and Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, **88**, 303–338.

Fang, C. and Yang, X. (2024). Lightweight YOLOv8 for wheat head detection. *IEEE Access*, **12**, 66214–66222.

Hasan, M. M., Chopin, J. P., Laga, H., and Miklavcic, S. J. (2018). Detection and analysis of wheat spikes using convolutional neural networks. *Plant Methods*, **14**, 100.

Huete, A. R. (1988). A soil-adjusted vegetation index (savi). *Remote Sensing of Environment*, **25**, 295–309.

Jocher, G., Chaurasia, A., Qiu, J., and Stoken, A. (2023). YOLO by ultralytics. GitHub repository.

Jordan, C. F. (1969). Derivation of leaf-area index from quality of light on the forest floor. *Ecology*, **50**, 663–666.

Khaki, S. and Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in Plant Science*, **10**, 621.

Kumar, M., Arora, A., Marwaha, S., Chinnusamy, V., Kumar, S., Jain, R., and Pal, S. (2024). Machine learning based approach for wheat plant senescence quantification. *Plant Physiology Reports*, **29**, 823–835.

Kumar, S., Raju, D., Sahoo, R. N., and Chinnusamy, V. (2016). Phenomics: unlocking the hidden genetic variation for breaking the barriers in yield and stress tolerance. *Indian Journal of Plant Physiology*, **21**, 409–419.

Lin, T. Y., Maire, M., Belongie, J., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Lecture Notes in Computer Science*, volume 8693, 740–755.

Liu, Y., Sun, L., Liu, B., Wu, Y., Ma, J., Zhang, W., Wang, B., and Chen, Z. (2023). Estimation of winter wheat yield using multiple temporal vegetation indices derived from uav-based multispectral and hyperspectral imagery. *Remote Sensing*, **15**, 4800.

Misra, T., Arora, A., Marwaha, S., Chinnusamy, V., Rao, A. R., Jain, R., Sahoo, R. N., Ray, M., Kumar, S., Raju, D., Jha, R. R., Nigam, A., and Goel, S. (2020). Spikesegnet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging. *Plant Methods*, **16**, 40.

Misra, T., Arora, A., Marwaha, S., Ranjan Jha, R., Ray, M., Kumar, S., Kumar, S., and Chinnusamy, V. (2022). Yield-spikesegnet: An extension of spikesegnet deep-learning approach for the yield estimation in the wheat using visual images. *Applied Artificial Intelligence*, **36**, 2137642.

Muruganantham, P., Wibowo, S., Grandhi, S., Samrat, N. H., and Islam, N. (2022). A systematic literature review on crop yield prediction with deep learning and remote sensing. *Remote Sensing*, **14**, 1990.

Patrignani, A. and Ochsner, T. E. (2015). Canopeo: A powerful new tool for measuring fractional green canopy cover. *Agronomy Journal*, **107**, 2312–2320.

Qiongyan, L., Cai, J., Berger, B., Okamoto, M., and Miklavcic, S. J. (2017). Detecting spikes of wheat plants using neural networks with laws texture energy. *Plant Methods*, **13**, 83.

Ren, J., Chen, Z., Zhou, Q., and Tang, H. (2008). Regional yield estimation for winter wheat with modis-ndvi data in shandong, china. *International Journal of Applied Earth Observation and Geoinformation*, **10**, 403–413.

Richardson, A. J. and Everitt, J. H. (1992). Using spectral vegetation indices to estimate rangeland productivity. *Geocarto International*, **7**, 63–69.

Sishodia, R. P., Ray, R. L., and Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, **12**, 3136.

Su, X., Wang, J., Ding, L., Lu, J., Zhang, J., Yao, X., Cheng, T., Zhu, Y., Cao, W., and Tian, Y. (2023). Grain yield prediction using multi-temporal UAV-based multispectral vegetation indices and endmember abundance in rice. *Field Crops Research*, **299**, 108992.

Sultana, S. R., Ali, A., Ahmad, A., Mubeen, M., Zia-Ul-Haq, M., Ahmad, S., and Jaafar, H. Z. (2014). Normalized difference vegetation index as a tool for wheat yield estimation: A case study from faisalabad, pakistan. *The Scientific World Journal*, 725326.

Varghese, R. and Sambath, M. (2024). YOLOv8: A novel object detection algorithm with enhanced performance and robustness. In *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, 1–6.

Wen, C., Ma, Z., Ren, J., Zhang, T., Zhang, L., Chen, H., Su, H., Yang, C., Chen, H., and Guo, W. (2024). A generalized model for accurate wheat spike detection and counting in complex scenarios. *Scientific Reports*, **14**, 24189.

Xie, Y., Wang, C., Yang, W., Feng, M., Qiao, X., and Song, J. (2020). Canopy hyperspectral characteristics and yield estimation of winter wheat (triticum aestivum) under low temperature injury. *Scientific Reports*, **10**, 244.

Xiong, C., Fan, C., and Huang, X. (2020). Multipath exploitation with time reversal waveform covariance matrix for SNR maximization. *Remote Sensing*, **12**, 3565.

Xue, J. and Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 1–17.

Zang, H., Wang, Y., Ru, L., Zhou, M., Chen, D., Zhao, Q., Zhang, J., Li, G., and Zheng, G. (2022). Detection method of wheat spike improved YOLOv5s based on the attention mechanism. *Frontiers in Plant Science*, **13**, 993244.

Zhou, H., Yang, J., Lou, W., Sheng, L., Li, D., and Hu, H. (2023). Improving grain yield prediction through fusion of multi-temporal spectral features and agronomic trait parameters derived from UAV imagery. *Frontiers in Plant Science*, **14**, 1217448.

ANNEXURE

Table 5: Plant varieties and genotypes used for imaging

S.No.	Variety/Genotype	S.No.	Variety/Genotype	S.No.	Variety/Genotype
1	BarhamRF	23	DBW51	45	LOKBOLD
2	C518RF	24	DL1266-2	46	LSP3043
3	CG1029	25	Drysdale	47	MACS6222
4	CHIRIYA3	26	HD3271	48	MP1358
5	Coreset110	27	HD2967	49	NIAW34
6	Coreset112	28	HD3086	50	NP4
7	Coreset128	29	HD308673RF	51	NP852
8	Coreset129	30	HD3237	52	NW1014
9	Coreset141	31	HD3271	53	PBW343
10	Coreset17	32	HD2329	54	PBW502
11	Coreset32	33	HD2687	55	PBW644
12	Coreset38	34	HD2833	56	PBW681
13	Coreset41	35	HD2864	57	PBW825
14	Coreset42	36	HD3171	58	PISSI-LOCAL
15	Coreset53	37	HI1500RF	59	RAJ4229
16	Coreset58	38	HI1544	60	Raj3077
17	Coreset66	39	HS611	61	Raj4037
18	Coreset7	40	IHD-29672	62	UAS446
19	Coreset74	41	K9465	63	UASBW10453
20	Coreset93	42	KACHU#1	64	UP2425
21	DBW110	43	LOK1	65	UP262
22	DBW222	44	LOK67	66	VL892