



OPEN ACCESS

EDITED BY

Lifeng Xu,
Zhejiang University of Technology, China

REVIEWED BY

Vadim Volkov,
London Metropolitan University,
United Kingdom
Lihong Zhu,
Changsha University of Science and
Technology, China

*CORRESPONDENCE

James C. Schnable
✉ schnable@unl.edu
Patrick S. Schnable
✉ schnable@iastate.edu
Baskar Ganapathysubramanian
✉ baskarg@iastate.edu

RECEIVED 25 April 2025

ACCEPTED 31 August 2025

PUBLISHED 30 September 2025

CITATION

Powadi AA, Jubery TZ, Tross M, Shrestha N,
Coffey L, Schnable JC, Schnable PS and
Ganapathysubramanian B (2025) Enhancing
yield prediction from plot-level satellite
imagery through genotype and
environment feature disentanglement.
Front. Plant Sci. 16:1617831.
doi: 10.3389/fpls.2025.1617831

COPYRIGHT

© 2025 Powadi, Jubery, Tross, Shrestha,
Coffey, Schnable, Schnable and
Ganapathysubramanian. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Enhancing yield prediction from plot-level satellite imagery through genotype and environment feature disentanglement

Anirudha A. Powadi¹, Talukder Z. Jubery², Michael Tross³,
Nikee Shrestha³, Lisa Coffey⁴, James C. Schnable^{3*},
Patrick S. Schnable^{4,5*} and Baskar Ganapathysubramanian^{5,6*}

¹Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, United States,

²Translational AI Research and Education Center, Iowa State University, Ames, IA, United States,

³Department of Agronomy, University of Nebraska, Lincoln, NE, United States, ⁴Department of Agronomy, Iowa State University, Ames, IA, United States, ⁵Plant Sciences Institute, Iowa State University, Ames, IA, United States, ⁶Department of Mechanical Engineering, Iowa State University, Ames, IA, United States

Accurately predicting yield during the growing season enables improved crop management and better resource allocation for both breeders and growers. Existing yield prediction models for an entire field or individual plots are based on satellite-derived vegetation indices (VIs) and widely used machine learning-based feature extraction models, including principal component analysis (PCA) and autoencoders (AE). Here, we significantly enhance pre-harvest yield prediction at plot-scale using Compositional Autoencoders (CAE) — a deep-learning-based feature extraction approach designed to disentangle genotype (G) and environment (E) features — on high-resolution, plot-level satellite imagery. Our approach uses a dataset of approximately 4,000 satellite images collected from replicated plots of 84 hybrid maize varieties grown at five distinct locations across the U.S. Corn Belt. By deploying the CAE model, we improve the separation of genotype and environment effects, enabling more accurate incorporation of genotype-by-environment (G×E) interactions for downstream prediction tasks. Results show that the CAE-based features improve early-stage yield predictions by up to 10% compared to traditional autoencoder-based features and outperform vegetation indices (VIs) by 9% across various growth stages. The CAE model also excels in separating environmental factors, achieving a high silhouette score of 0.919, indicating effective clustering of environmental features. Moreover, the CAE consistently outperforms standard models in unseen environments and unseen genotypes yield predictions, demonstrating strong generalizability. This study demonstrates the value of disentangling G and E effects for providing more accurate and early yield predictions that support informed decision-making in precision agriculture and plant breeding.

KEYWORDS

representation learning, genotype × environment interactions, crop yield prediction, satellite data, latent feature extraction

1 Introduction

Maize (*Zea mays* L.) is one of the world's most important cereal crops, with nearly one billion tons of maize produced on about 200 million planted hectares annually (Erenstein et al., 2022). The productivity of maize is highly variable, largely due to differences in environmental conditions (Jägermeyr et al., 2021; Yang et al., 2024), crop management practices (Lobell et al., 2014; Ort and Long, 2014; Lobell et al., 2020), and genetic factors (Gamble, 1962; Troyer, 1990; Butrón et al., 2012; Nolan and Santos, 2012). These three components interact in complex ways, complicating the task of accurately predicting yields for specific hybrids under particular conditions and crop management strategies.

Yield forecasting is crucial for multiple stakeholders in agriculture. Farmers benefit from reliable early predictions, as these help optimize management practices, maximizing the use of limited resources like water and fertilizers. Meanwhile, breeders pursue stable genotypes that perform well in diverse environments, which requires extensive testing of thousands of genotypes in dozens of locations to identify those most suitable for target regions (Cooper et al., 2014; Li et al., 2024; Tarekegne et al., 2024). Thus, improving yield prediction systems directly supports agricultural resilience and the long-term stability of food systems.

Remote sensing technologies have been extensively used to tackle the challenges of yield prediction. Historically, low-resolution field- or county-level satellite imagery has been extensively used for predictive plant phenotyping, including yield estimation for dry beans, rice, maize, wheat, soybean (Hamar et al., 1996; Yang et al., 2009; Wang et al., 2010; Noureldin et al., 2013; Johnson et al., 2016b; Jin et al., 2017; Schwalbert et al., 2018; Kamir et al., 2020; Schwalbert R. et al., 2020; Schwalbert RA. et al., 2020; Shendryk et al., 2021; Roznik et al., 2022; Joshi et al., 2023). This imagery allows researchers to gain a broader perspective on crop performance across large regions. UAVs (unmanned aerial vehicles) have emerged as a complementary tool, offering plot-level high-resolution data (López-Granados et al., 2016; Kanning et al., 2018; Han et al., 2019; Du et al., 2022), yet satellite imagery retains certain advantages, particularly for large-scale studies. It allows data collection over extensive areas with considerably less logistical effort, while also delivering preprocessed data and reducing user workload. This makes it the preferred choice for multi-location analyses. Despite this benefit, the utility of satellite data for plot-level prediction is often hindered by lower spatial resolution, which requires specialized acquisition strategies to obtain more detailed imagery (Phang et al., 2023). High-resolution satellite imagery is becoming available, although applications still face obstacles such as restricted satellite availability, limited onboard storage capacity, and cloud cover (over the target area), making high-resolution data

acquisition challenging. Despite the challenges, the advantages of collecting high-resolution satellite data instead of UAV data for breeding applications may outweigh the limitations. Evaluating the large numbers of varieties (called hybrids) across many different environments becomes easier with satellite data. The most important and most common metric of evaluating hybrids is yield. The ability to forecast yield earlier in the growing season enhances breeding decisions, thus accelerating the selection process. Lowering the cost of collecting trait data from more plots in diverse locations allows for more extensive evaluations. This increases the accuracy of genetic yield potential estimates and accelerates genetic gain per breeding cycle.

Over time, a variety of methods have been developed for yield forecasting using satellite data. Many studies combine satellite data with crop growth model estimations (de Wit and van Diepen, 2008; Jeong et al., 2018; Zhao et al., 2020; Zare et al., 2022; Luo et al., 2023), incorporate additional environmental factors like weather conditions (Bai et al., 2010; Schwalbert R. et al., 2020; Schwalbert RA. et al., 2020; Nieto et al., 2021; Lang et al., 2022), soil data (Broms et al., 2023; Mahalakshmi et al., 2025), and terrain information (Li et al., 2022; Sahbeni et al., 2023). Machine learning has been widely applied, using handcrafted features derived from raw images, including histogram-based representation of time-series data (You et al., 2017), and vegetation indices, particularly the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) (Yang et al., 2006; Yang et al., 2013; Johnson et al., 2016b; Peralta et al., 2016; Schwalbert et al., 2018; Cai et al., 2019; Mateo-Sanchis et al., 2019; Kamir et al., 2020; Khalil and Abdullaev, 2021; Meroni et al., 2021). While these indices have been effective, they are inherently limited by the scope of human expertise, restricting the ability to capture the full complexity of crop conditions (Feldmann et al., 2021).

Representation learning has emerged as a promising alternative approach for overcoming these limitations (Gage et al., 2019; Ubbens et al., 2020; Tross et al., 2023). Conventionally, various machine learning techniques, including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), t-distributed Stochastic Neighbor Embedding (t-SNE), and autoencoders, have been applied to derive 'latent representations' from high-dimensional datasets (Zhong et al., 2016; Kopf and Claassen, 2021; Alexander et al., 2022; Gomari et al., 2022; Iwasaki et al., 2023; Song et al., 2023). Autoencoders stand out for their ability to identify non-linear patterns. Through the encoding of data into a reduced latent space followed by reconstruction of the input, these models generate a concise and meaningful representation that plays a key role in phenotyping (Gage et al., 2019; Ubbens et al., 2020; Tross et al., 2023). Despite their value, representations produced by autoencoders commonly struggle to differentiate between genetic and environmental effects, creating 'entangled' latent spaces in which specific plant traits—like 'leaf number,' 'height,' and 'chlorophyll concentration'—are blended rather than distinctly isolated. Separating these traits in the latent space could enhance the interpretability of the resulting latent factors. Unlike traditional feature engineering, representation

Abbreviations: CAE, Compositional autoencoder; VIs, vegetation indices; AE, Vanilla autoencoder; PC, Principal component; RMSE, Root mean square error; zg, genotype component in the latent space generated by the CAE; ze, macro-environment component in the latent space generated by the CAE; zp, micro-environment component in the latent space generated by the CAE.

learning can automatically extract complex, high-dimensional features from raw input data without predefined formulas, potentially providing a richer understanding of crop traits and often outperforming traditional indices (Okada et al., 2024). This ability makes representation learning well-suited for capturing characteristics that influence crop performance. Recent studies have applied representation learning techniques to satellite data for various agricultural applications, such as classifying crop traits, mapping floods, and monitoring land use (Dumeur et al., 2024; Goyal et al., 2024; Nakayama and Su, 2024). While these methods have demonstrated success in large-scale yield estimation, they have yet to be broadly adopted for localized, plot-level predictions, where high precision is crucial for actionable insights.

Additionally, most existing approaches have not adequately addressed the complex interactions between genotype (G) and environment (E)—commonly referred to as GxE interactions—in their forecasting models. Incorporating GxE interactions is particularly important, as these interactions are a major source of variability in crop performance and the primary reason that the top-performing hybrids in one environment will often rank lower in relative performance in a second environment (Kusmiec et al., 2018). A more nuanced understanding of how specific genotypes respond to different environmental conditions could significantly enhance prediction precision, especially at smaller spatial scales (e.g., plot-level).

In our recent work, we introduced a model, the Compositional Autoencoder (CAE), specifically designed to address these challenges (Powadi et al., 2024). The CAE integrates GxE interactions into the yield prediction framework, allowing for a more comprehensive representation of high-dimensional hyperspectral ground-based sensor data as GxE components (Figure 1). By capturing both genetic and environmental influences, the CAE outperformed traditional methods, demonstrating its potential as a valuable tool for improving the accuracy of yield predictions.

Building on this foundation, the present study aims to extend the CAE framework to satellite data at the plot level. This extension enables more precise yield forecasting and allows for deeper insights into the interactions between genetic and environmental factors. Our objectives in this work are to (a) evaluate CAE performance on disentangling genotype and environment features from satellite

imagery collected at the plot scale to better understand their contributions to yield (Figure 2), and (b) to improve yield prediction accuracy by leveraging these disentangled features, as evidenced by the CAE's superior performance over other representations, including traditional vegetation indices.

2 Materials and methods

2.1 Satellite images and yield data

2.1.1 Locations and plants

Maize (*Zea mays L.*) field experiments were conducted at five locations: Scottsbluff, NE (41.85°N, -103.70°W), Lincoln, NE (40.86°N, -96.61°W), Missouri Valley, IA (41.67°N, -95.94°W), Ames, IA (42.01°N, -93.73°W), and Crawfordsville, IA (41.19°N, -91.48°W) in 2022. A map of these locations is provided in the Supplementary Figure S1. Depending on local weather conditions, planting occurred between April 29 and May 23, 2022. This work utilizes 84 hybrid genotypes planted in two replicated plots at each of these five locations.

The maize plants were cultivated under rain-fed conditions at four locations: Lincoln, Missouri Valley, Ames, and Crawfordsville. In contrast, the Scottsbluff site was irrigated, receiving a total of 16.9 inches (429.26 mm) of water over the growing season. All locations except Missouri Valley had one rate of nitrogen fertilization treatment of 150 lbs/acre. Nitrogen fertilization treatment for Missouri Valley was 175 lbs/acre. More details on the experimental design and data collection are available in Shrestha et al. (Shrestha et al., 2025). For one of the evaluation experiments of the CAE (experiment 3 in section 3.2), we additionally used satellite data of 45 common hybrids from Ames, Lincoln, and Missouri Valley from 2023.

2.1.2 Satellite imagery and plot extraction

For this study, we utilized the Pléiades Neo satellite constellation to capture images at four different time points (TP), of which two were early in the season (July), referred to as TP1 and TP2, capturing the late vegetative stage and early flowering stages. Two others were late in the season (September), referred to as TP3 and TP4, capturing mature and harvest-ready plants. The growth-

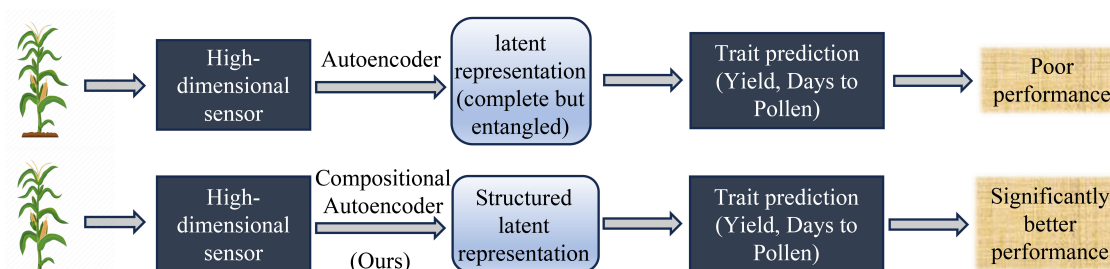
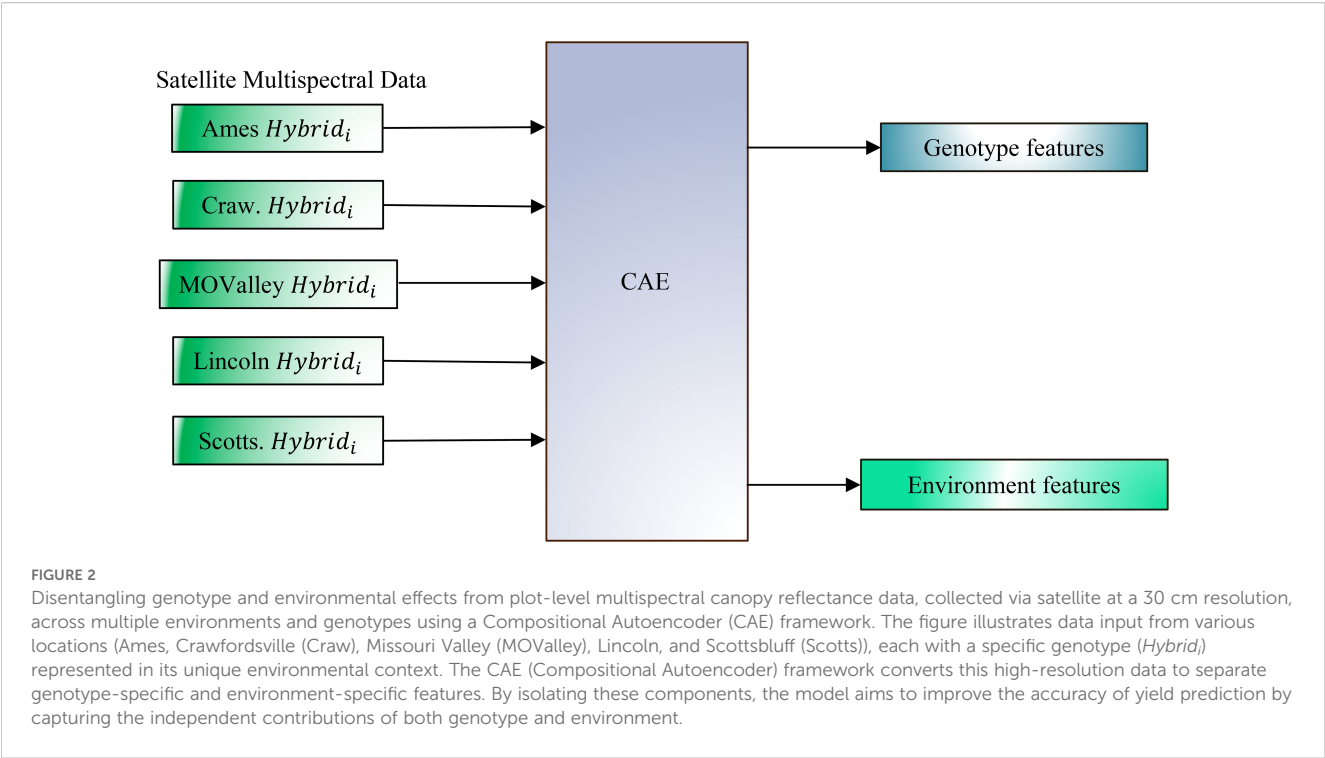


FIGURE 1

Overview of trait prediction from high-dimensional plant-level hyperspectral sensor data workflow and performance of Compositional Autoencoder (CAE) representation with other representation learning as shown in (Powadi et al., 2024).



stage timeline for the maize plants was recorded for the plants grown in Lincoln is shown in the [Supplementary Figure S2](#). [Table 1](#) provides the specifications of the Pléiades Neo satellite constellation used in our study. The six bands in the satellite multispectral images cover the following spectral ranges: Red (620–690 nm), Green (530–590 nm), Blue (450–520 nm), Near-infrared (NIR, 770–880 nm), Red Edge (700–750 nm), Deep Blue (400–450 nm).

TABLE 1 Satellite specifications for the Pléiades Neo satellite constellation used in our study.

Parameter	Description
Orbit	Sun-synchronous, 10:30 a.m. descending node, 620 km altitude
Number of Satellites	2 identical satellites in constellation
Sensor Bands	Panchromatic (450–800 nm)
	Multispectral: • Red (620–690 nm) • Green (530–590 nm) • Blue (450–520 nm) • Near-infrared (770–880 nm) • Red Edge (700–750 nm) • Deep Blue (400–450 nm)
Sensor Resolution	30 cm (panchromatic band), 1.2 m (multi-spectral bands)
Dynamic Range	12-bits per pixel
Swath Width	At nadir: 14 km
Revisit Frequency	Daily anywhere (30° off-nadir) or twice daily anywhere (45° off-nadir)
Acquisition Capacity	Up to 2 million km ² per day

In addition to multispectral images, a single-band panchromatic raster file with a wide bandwidth of approximately 450–800 nm was generated. Each image covered a total area of 100 km × 100 km per location, encompassing the entire experimental field at each location. The final 16-bit GeoTIFF satellite images with 30-cm resolution were generated after panchromatic sharpening (pansharpening), manual orthorectification, and atmospheric correction by the service provider prior to the delivery of imagery. The plot-level images were extracted from the satellite data as described in our previous paper ([Shrestha et al., 2025](#)). The extracted plot images are of the shape 11 x 22 x 6 (Width x Height x Channels) pixels. [Figure 3](#) shows samples of satellite image data for the 4 time points for all the locations we capture. The plots measure approximately 3.3 meters by 6.6 meters (shown in [Figure 3](#) on the TP4 sample for the Mo Valley), covering an area of 21.78 square meters. With the Pléiades Neo satellite’s native spatial resolution of 30 cm (0.3 meters) per pixel, each pixel represents a ground area of 0.09 square meters. A single plot would thus be sampled by approximately 11 pixels along the width and 22 pixels along the length, for a total of about 242 pixels. This limited pixel count per plot results in the expected pixelated appearance in the imagery, as fine details smaller than the resolution cannot be resolved. The plot boundaries and extraction were performed using the ArcPy Jupyter Notebook environment implemented in ArcGIS Pro V3.2.0.

2.1.3 Preprocessing of plot images

The preprocessing of the plot images was divided into two stages: outlier removal and normalization. For each spectral channel, outliers were removed by clipping values outside 3 standard deviations from the mean. [Figure 4](#) illustrates the distribution of pixel values across channels after outlier removal.

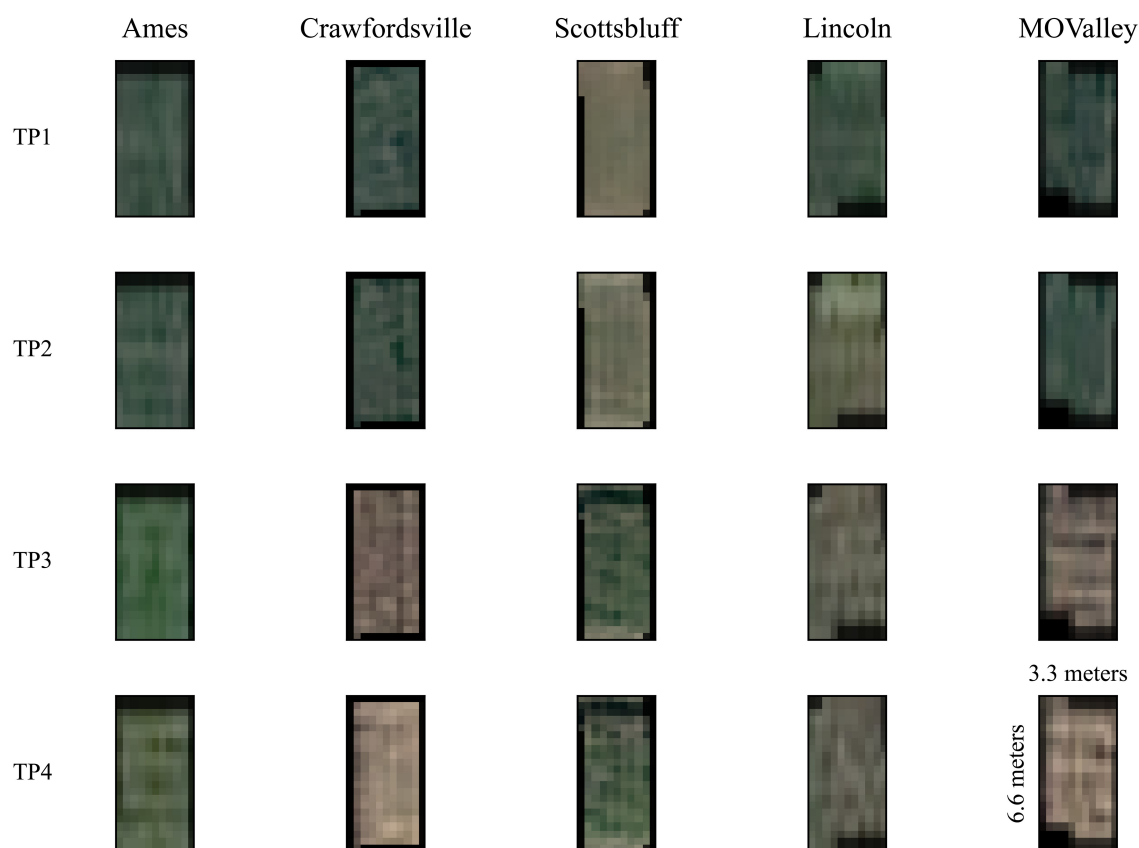


FIGURE 3

Satellite data samples from an arbitrary plot at each location are shown across four timepoints (TP), with the first two collected early in the growing season (July) and the last two collected toward the end of the season (September). Only the R, G, and B channels are visualized. Each plot is 11 by 22 pixels, representing 3.3 meters by 6.6 meters (as shown in TP4, MOValley). The distinctive variations among images at each time point are primarily due to differences in planting dates, which ranged from April 29 to May 23, 2022, depending on local weather conditions. Notably, images from Scottsbluff exhibit visible differences, as they were captured from a field planted later than the others.

The 6-channel data were then normalized on a per-channel basis using min-max scaling. After normalization, each plot image was flattened into a 1 x 1452 vector before being fed into the neural network.

2.1.4 Yield data

Yield data is recorded at the time of the harvest in the field. Plants were mechanically harvested from the two middle rows of each four-row plot using a plot combine. At the time of harvesting, the combine recorded percent moisture content, on a fresh weight basis, and the fresh weight of the grain in pounds. To compare across locations, yield in bushels per acre at 15.5% moisture on a fresh weight basis and 56 pounds per bushel was calculated based on these values and the total area occupied by the middle two rows of each plot at each location.

The yield data were recorded for these 84 hybrids, which were used for evaluating the predictive capabilities of the CAE as well as the degree of disentanglement of environmental features. Figure 5 shows the yield distribution for different locations. As can be expected, genotypes respond differently to different environments. Clearly, Crawfordsville shows the highest yield overall, while Lincoln shows the lowest yield distribution. While different yields

in different environments are expected due to the differences in soil, weather, and elevation conditions, Lincoln produced the lowest yield due to intense weed pressure. A deeper analysis of these yields and experiments can be found in the previous work (Shrestha et al., 2025).

2.2 Autoencoder

An autoencoder is a type of neural network designed for representation learning, enabling the extraction of efficient and compressed representations of input data. It consists of two main components: an encoder and a decoder. The encoder compresses the input data into a latent representation, while the decoder aims to reconstruct the original input from this latent space. Through this process, the autoencoder effectively captures the essential characteristics of the input data, thereby learning meaningful representations. Figure 6 shows our implementation, where the encoder and decoder are built using multiple fully connected layers, each utilizing the SeLU activation function. The encoder compresses the 6-channel satellite input data into a lower-dimensional latent space, which retains key features of the input

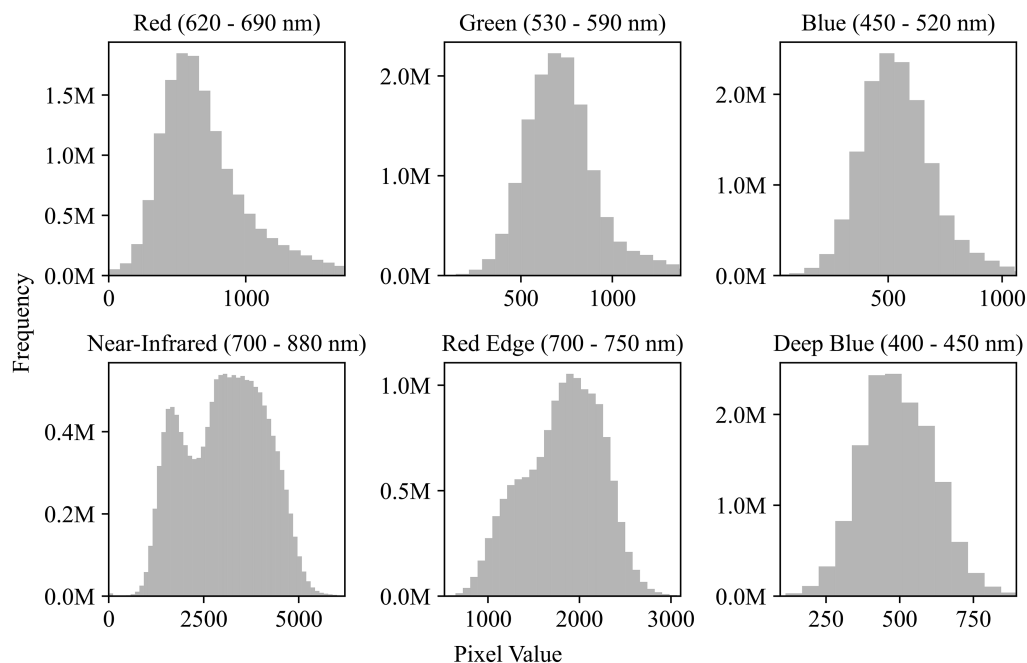


FIGURE 4

Pixel value distributions across spectral channels. Histograms show the frequency of pixel values for six satellite image channels: Red (620–690 nm), Green (530–590 nm), Blue (450–520 nm), Near-Infrared (700–880 nm), Red Edge (700–750 nm), and Deep Blue (400–450 nm). Although the image data is 16-bit (range: 0–65,536), most pixel values fall within 0–5000, indicating a limited dynamic range in practice.

while being highly informative. The decoder then reconstructs the original input from this compressed latent representation. [Tables 2a, b](#) provide detailed information about the specific layers used in the encoder and decoder. This encoder-decoder architecture forms a latent space that is both compact and representative of the input data, providing added benefits of noise reduction and feature enhancement. This “vanilla autoencoder” serves as our baseline representation learning model.

2.3 Compositional autoencoder

The Compositional Autoencoder (CAE) builds upon the standard autoencoder to disentangle the latent space into multiple components that reflect various influential factors in the data. By learning structured representations, the CAE captures general, environmental, and specific attributes, making it significantly more informative than a conventional autoencoder. The CAE is composed of three primary components: an encoder, a fusion block, and a decoder.

2.3.1 Architecture overview

The CAE architecture retains the traditional encoder-decoder structure, with the addition of a fusion block to combine and disentangle encoded data from multiple images. Satellite data of the same genotype across all the locations are grouped together in 84 groups (84 being the number of genotypes). Each group will have 5 images (for 5 locations), which will be sequentially passed to the

encoder, resulting in a separate latent representation for each image in the group. The encoder used here is identical to the one implemented in the standard autoencoder, ensuring consistency in feature extraction across both approaches.

The latent features derived from these images are then aggregated using a fusion layer. This fusion layer operates by combining the encoded information, resulting in a unified latent vector that integrates features from all input images. A detailed description of this fusion mechanism is provided in [Table 3](#).

Once fused, the latent representation is divided into three distinct components. The general features capture shared characteristics present in all images. The environmental features represent attributes specific to a given environment, shared across images from similar conditions. Finally, specific features represent unique aspects that differ across each image, such as particular objects or anomalies. This structured separation ensures that the CAE can provide a richer, more nuanced representation of the input data.

After this separation, the decoder reconstructs each original image by reassembling the general, environmental, and specific features. The decoder, identical to that used in the basic autoencoder, leverages the latent space to reconstruct the input, thereby enabling an in-depth understanding of the underlying factors driving variations in the satellite data. [Figure 7](#) presents an overview of the CAE architecture, showing how different components contribute to reconstructing the input images. The detailed structure of the fusion layer is provided in [Table 3](#), and [Table 4](#) demonstrates the disentangled latent representation for each individual image.

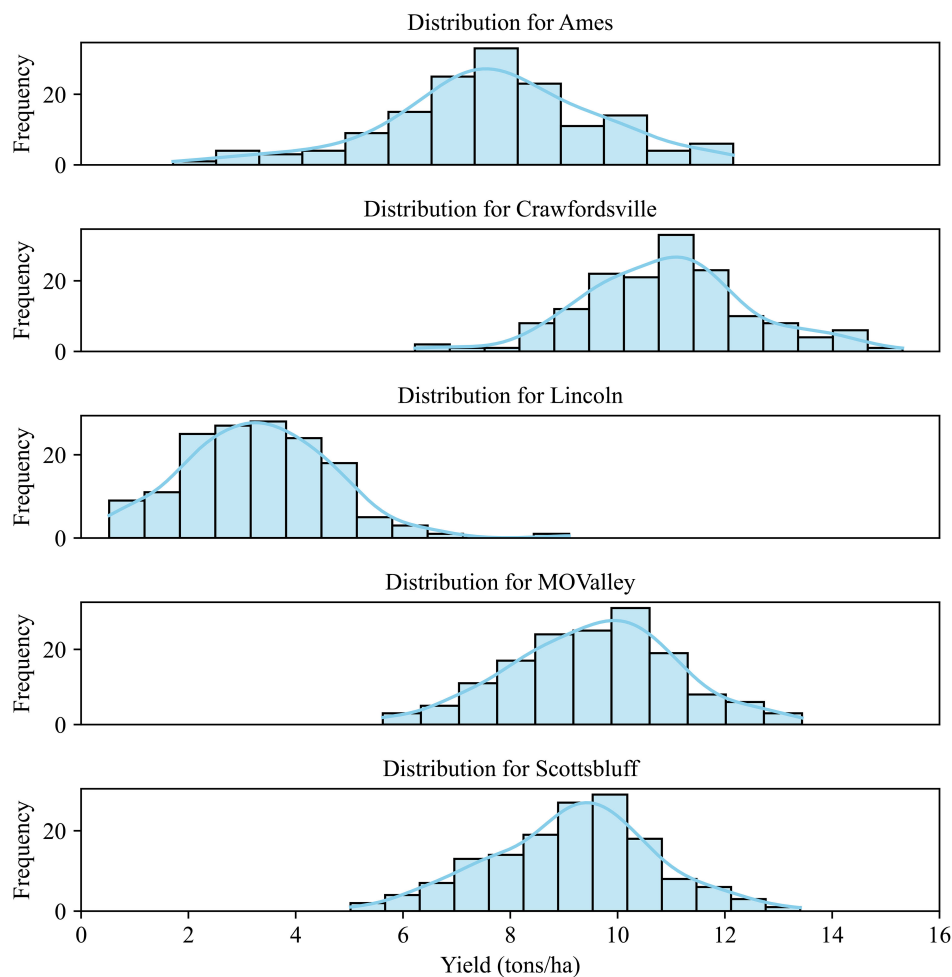


FIGURE 5

Yield distribution for the 84 hybrids across different environments (locations). The plants were grown under rain-fed conditions in Lincoln, Missouri Valley, Ames, and Crawfordsville. The experimental field in Scottsbluff was irrigated.

2.3.2 Loss function

To effectively train the CAE, a two-part loss function is used, consisting of a reconstruction loss and a correlation loss.

The reconstruction loss, computed as the mean squared error (MSE), ensures that the network can accurately reconstruct the original satellite images from the learned latent space. This loss encourages the CAE to capture the most relevant features for high-quality reconstruction, ensuring that the latent space is both informative and efficient.

The correlation loss is applied to maintain the independence of the disentangled components — general, environmental, and specific features — within the latent space. This loss penalizes any correlation between different parts of the latent space, thus enforcing their distinctiveness and improving the interpretability of the model. The correlation loss is expressed as:

$$\text{Correlation Loss} = \sum_{i=1}^N \sum_{j=i}^N |\text{CorrMat}_{ij}| - I_{ij} \quad (1)$$

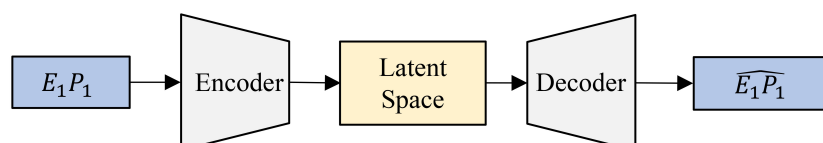


FIGURE 6

An autoencoder processes a single image (for example, $E1P1$ (Environment 1 - Replicate 1)) through an encoder to compress it into a lower-dimensional latent space, capturing its essential features. The decoder then reconstructs the input from this compressed representation.

TABLE 2 Configuration details for encoder and decoder.

Layer type	Dimensions	Activation
(a) Encoder Configuration		
Linear	input shape → 2200	SELU
Linear	2200 → 2000	SELU
Linear	2000 → 1000	SELU
Linear	1000 → zg + ze + zp	None
(b) Decoder Configuration		
Linear	zg + ze + zp → 2024	SELU
Linear	2024 → 3000	SELU
Linear	3000 → 2200	SELU
Linear	2200 → input_shape	Sigmoid

input shape = 1 x 1452, zg = dimensions for genotype features, ze = dimensions for macro-environment features, zp = dimensions for microenvironment features.

In this equation, CorrMat_{ij} represents the correlation coefficient between the latent space dimensions i and j , and N denotes the size of the square correlation matrix. The term I_{ij} refers to the identity matrix, which ensures that the diagonal elements (where $i = j$) do not contribute to the loss, thus emphasizing only the off-diagonal correlations.

The combination of reconstruction and correlation loss allows CAE to learn latent representations that are informative, distinct, and suitable for downstream tasks, such as yield prediction and analysis of genotype-by-environment interactions. The correlation coefficient used here is the Pearson correlation coefficient (r), which measures the linear relationship between two variables. It is defined by the following equation:

$$r = \frac{\sum_{i=1}^n (p_i - \bar{p})(k_i - \bar{k})}{\sqrt{\sum_{i=1}^n (p_i - \bar{p})^2 \sum_{i=1}^n (k_i - \bar{k})^2}}$$

(2)

In this equation (Equation 2), n represents the number of data points. p_i and k_i are elements from different dimensions of the latent space. $\bar{p}$ and $\bar{k}$ are the means of the p^{th} and k^{th} dimensions, respectively. Our objective is to achieve zero correlation between the latent space features that represent general, environmental, and specific variations. By enforcing this condition through the correlation loss function (Equation 1), the model ensures that each disentangled component captures its respective factor independently, leading to a more interpretable and effective representation of the underlying data.

TABLE 3 Fusion layer details.

Layer type	Dimensions	Activation
Linear	$N(zg + ze + zp) \rightarrow zg + E(ze) + N(zp)$	None

'N' = number of replicates per image, 'E' = number of environments, 'zg' = dimensions allocated to capture general features, 'ze' = dimensions allocated to capture environmental features, 'zp' = dimensions allocated to specific features.

3 Results and discussion

3.1 Environment disentanglement

We first assess the degree of disentanglement achieved by the Compositional Autoencoder (CAE) across different environments in our dataset.

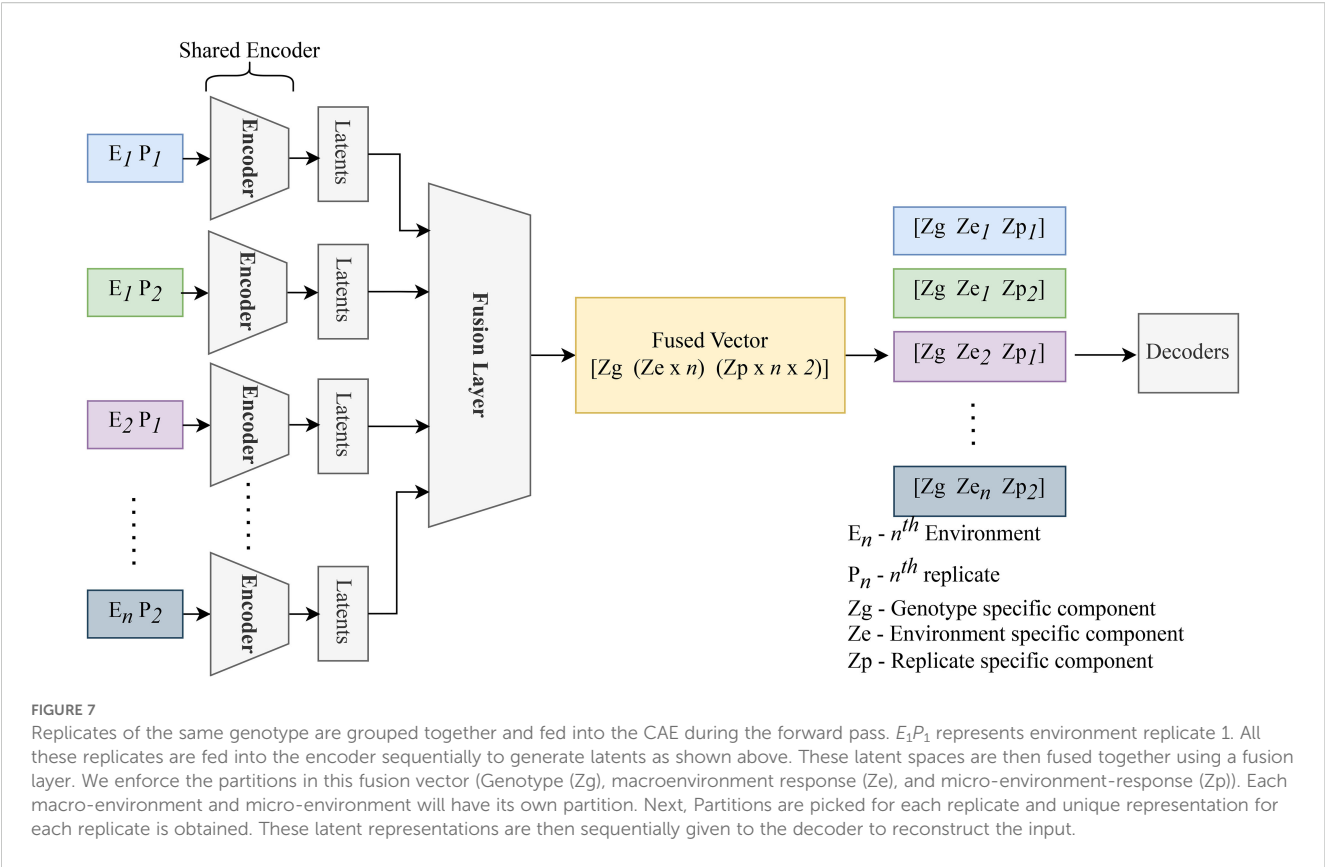
Our dataset comprises information from five distinct locations. To establish a baseline, we first analyzed how well the raw satellite data differentiates among these locations. We applied Principal Component Analysis (PCA) to the raw data, selecting the first three principal components. These components were then visualized in a 3D plot, with data points color-coded by location. To quantify the distinctness of the resulting clusters, we calculated their silhouette scores. The silhouette score measures how well data points in a cluster are grouped by comparing the average intra-cluster distance (how close points are within the same cluster) to the nearest inter-cluster distance (how far they are from points in the nearest cluster). Scores range from -1 to 1, where values close to 1 indicate well-separated, compact clusters, and values near or below 0 suggest overlapping or poorly defined clusters. Figure 8A illustrates this visualization of the raw data. Notably, the clusters representing different locations show significant overlap, resulting in a low silhouette score of only 0.293. This indicates that the raw satellite data alone does not effectively distinguish among environments.

To evaluate the environmental disentanglement achieved by the CAE, we analyzed its latent space. The CAE's latent space is divided into 'genotype', 'macro-environment', and 'micro-environment' components. For this analysis, we consider the 'macro-environment' component (and average over the 'genotype' and 'micro-environment' components) and perform the same clustering and visualization process as with the raw data. Figure 8B displays the clusters for all five locations generated by the CAE. Strikingly, the silhouette score for these clusters is 0.919, a substantial improvement over the raw data. This high score indicates that the CAE has successfully disentangled macroenvironmental factors, creating much more distinct clusters for each location.

We also investigated the CAE's ability to disentangle micro-environmental factors. As shown in Figure 9, the CAE demonstrates considerable success in this aspect as well, further highlighting its robustness in environmental disentanglement. These results suggest the CAE's potential to separate environmental factors detected via satellite imagery, which could lead to a more nuanced understanding of how different environments may contribute to plant characteristics and performance.

3.2 Performance on yield prediction

We next conducted multiple experiments to assess the performance of the latent vectors constructed by the Compositional Autoencoder (CAE) in predicting maize yield. Each experiment consists of using the latent representation as input to train a machine-learning model that predicts maize yield

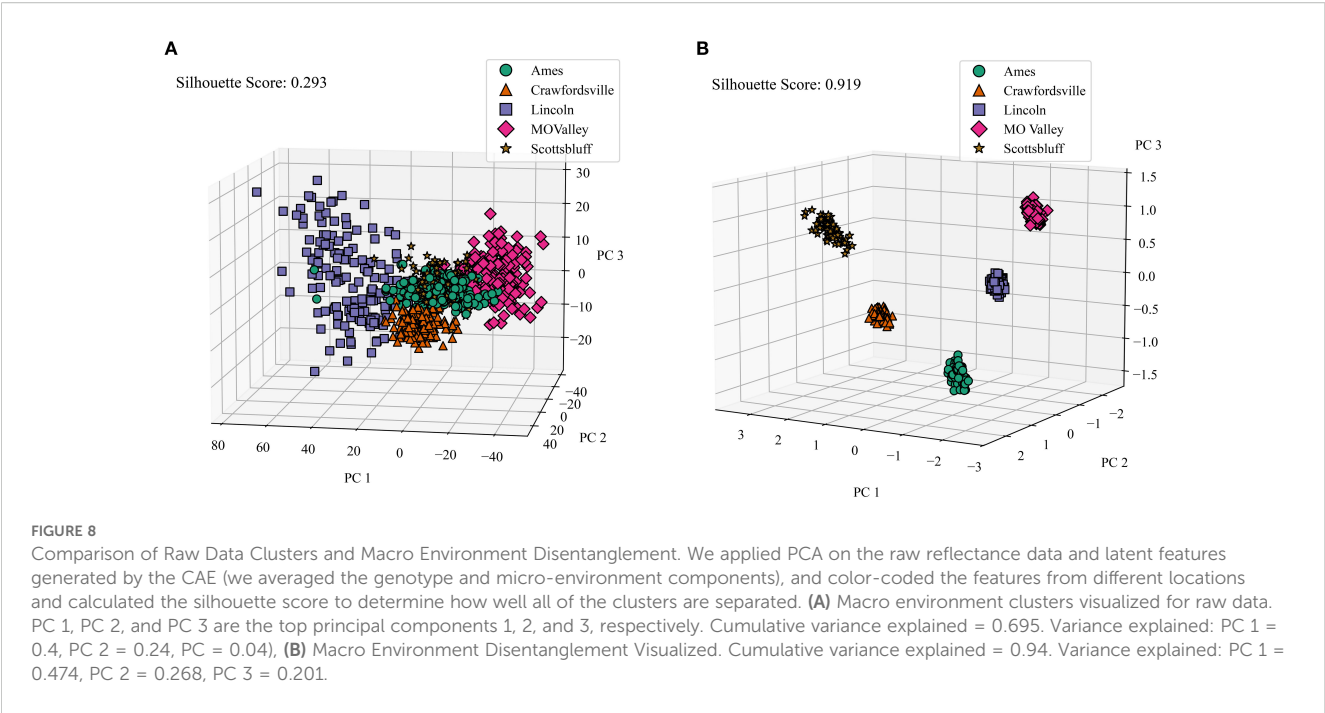


in the plots. The latent configuration used for all these downstream predictions shown in this section is ‘6-6-1’ (zg-ze-zp) which was selected based on the sensitivity analysis provided in [Supplementary Table S2](#). For yield prediction, we used the XGBoost model, which consistently outperformed other machine learning algorithms in our preliminary evaluation. Comparative results with classical models such as PLSR, Ridge regression, and Random Forest are provided in [Supplementary Tables S3 and S4](#). The hyperparameter settings used for each model configuration are summarized in

[Supplementary Tables S2](#). The three experiments were designed to evaluate model performance across different generalization scenarios. The corresponding data splits and evaluation protocols are described in [Supplementary Section 6](#). In experiment 1, we compare the predictive capability of the CAE latent vectors against two baselines: latent vectors generated by a vanilla autoencoder (AE) and vegetation indices (VIs). We perform cross-validation across genotypes. In experiment 2, we evaluate the ability of the CAE-based latent vectors to rank order genotypes according to their

TABLE 4 Disentangled latent-space representation of each image.

Image	Representation
E_1P_1	{{(Zg) genotype features, (Ze) macro-env. features [1], (Zp) micro-env. features [1]}}
E_1P_2	{{(Zg) genotype features, (Ze) macro-env. features [1], (Zp) micro-env. features [2]}}
E_2P_1	{{(Zg) genotype features, (Ze) macro-env. features [2], (Zp) micro-env. features [3]}}
E_2P_2	{{(Zg) genotype features, (Ze) macro-env. features [2], (Zp) micro-env. features [4]}}
E_3P_1	{{(Zg) genotype features, (Ze) macro-env. features [3], (Zp) micro-env. features [5]}}
E_3P_2	{{(Zg) genotype features, (Ze) macro-env. features [3], (Zp) micro-env. features [6]}}
E_4P_1	{{(Zg) genotype features, (Ze) macro-env. features [4], (Zp) micro-env. features [7]}}
E_4P_2	{{(Zg) genotype features, (Ze) macro-env. features [4], (Zp) micro-env. features [8]}}
E_5P_1	{{(Zg) genotype features, (Ze) macro-env. features [5], (Zp) micro-env. features [9]}}
E_5P_2	{{(Zg) genotype features, (Ze) macro-env. features [5], (Zp) micro-env. features [10]}}



predicted yield. Early and accurate rank ordering is a critical functionality in plant breeding, where early and accurate ranking can guide selection decisions. Finally, in experiment 3, we evaluate the performance of the latent CAE representation to predict the yield for a set of genotypes in an unseen environment (using data collected in Ames in the next year, 2023).

3.2.1 Experiment 1: CAE vs AE vs VIs yield prediction comparison, with 5-fold crossvalidation

We compared the yield prediction performance of the CAE latent vectors with a vanilla autoencoder (AE) latent vectors and vegetation indices (VIs) using 5-fold cross-validation across four key time points. Table 5 presents a summary of the results,

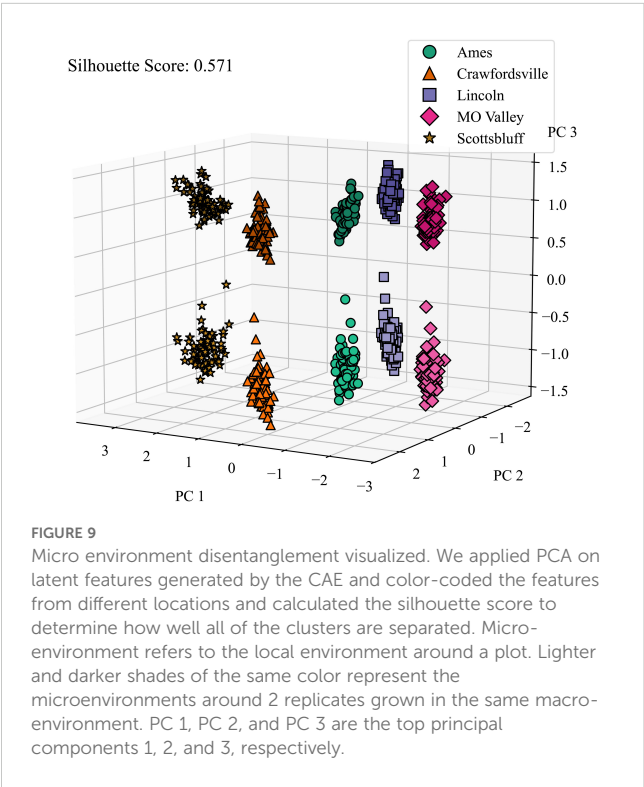


TABLE 5 Prediction performance of CAE (compositional autoencoder), AE (vanilla autoencoder), and VIs (vegetation indices) across timepoints 1, 2, 3, and 4, covering both early growth stages and near-harvest stages of the growing season.

Model	Timepoint	R2	RMSE (tons/ha)
CAE	1	0.79 (0.05)	1.36 (0.16)
AE	1	0.67 (0.04)	1.69 (0.13)
VIs	1	0.75 (0.03)	1.48 (0.12)
CAE	2	0.79 (0.043)	1.34 (0.14)
AE	2	0.72 (0.057)	1.56 (0.16)
VIs	2	0.80 (0.04)	1.30 (0.15)
CAE	3	0.77 (0.07)	1.41 (0.14)
AE	3	0.27 (0.1)	2.51 (0.21)
VIs	3	0.77 (0.07)	1.41 (0.20)
CAE	4	0.76 (0.04)	1.45 (0.14)
AE	4	0.45 (0.06)	1.51 (0.12)
VIs	4	0.70 (0.07)	1.41 (0.18)

Results are reported for yield prediction. Each entry shows the mean value, with the standard deviation from 5-fold cross-validation in parentheses. Bold values indicate the best performance for a given timepoint. R^2 is a unitless indicator of model fit, whereas RMSE is expressed in yield units (tons/ha) and represents the absolute prediction error.

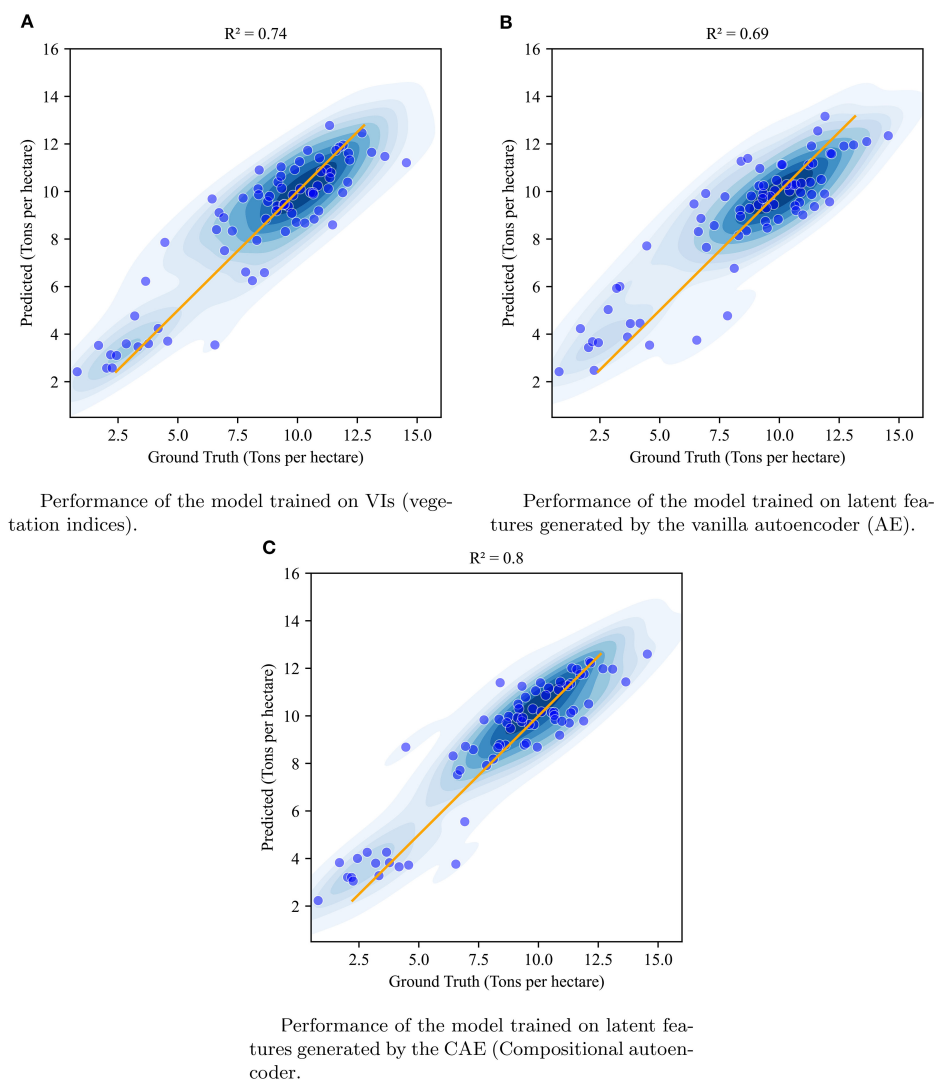


FIGURE 10

Visualization of the performance of XgBoost model when using vegetation indices (A), vanilla autoencoder generated latent features (B), and compositional autoencoder generated latent features (C) at timepoint 1.

including R^2 values and Root Mean Square Error (RMSE) in tons per hectare, which offer robust estimates of model performance across different timepoints. A visualization of these fits is given in Figure 10 with absolute values of the yield (tons/ha) for timepoint 1 (TP1). As the k-folds were grouped based on genotypes, the results here reflect the accuracy with which the performance of unseen genotypes could be predicted in observed environments.

The results clearly demonstrate that the CAE latent vector based predictor consistently outperforms the AE latent vector based predictor across all growth stages. Notably, the CAE latent vector based predictor shows higher R^2 values and lower RMSE compared to the AE latent vector based predictor at every time point. The performance difference is particularly evident at time point 3, where the CAE latent vector based model maintains a relatively strong performance (R^2 of 0.766) while the AE's performance drops significantly (R^2 of 0.27). Overall, the CAE latent vector based predictor demonstrates more consistent and robust performance

throughout the growing season. Additionally, when compared with a model that uses vegetation indices (VIs) (Table 5), the CAE latent vector based predictor consistently outperforms them at timepoints 1 and 4 and shows comparable performance at time points 2 and 3. This comparable performance during timepoint 2 can be attributed to the fact that plant growth at this stage is near its peak, which correlates well with final yield. The vegetation indices used in this paper are given in the Table 6. Their description is provided in the Supplementary Materials. The CAE latent vector based predictor demonstrates particular strength in early-season predictions, with improvements (over AE) of approximately 11% and 7% in R^2 values for the first two time points, respectively. Moreover, the CAE latent vector based predictor maintains more consistent performance across all time points, illustrating its robustness throughout the growing season.

The improved yield prediction accuracy demonstrated by the Compositional Autoencoder (CAE) latent vector based predictor suggests utility for both farmers and plant breeders. For example, in

TABLE 6 Vegetation indices (VIs) used in this study, along with their full names and source references.

VI	Full form	Citation
GLI	Green Leaf Index	(Mounir Louhaichi and Johnson, 2001)
NGRDI	Normalized Green-Red Difference Index	(Tucker, 1979)
VARI	Visible Atmospherically Resistant Index	(Gitelson et al., 2002)
VEG	Vegetation Index	(Hague et al., 2006)
RGBVI	Red-Green-Blue Vegetation Index	(Bendig et al., 2015)
ExG	Excess Green Index	(Woebbecke D et al., 1995)
ExR	Excess Red Index	(Meyer and Neto, 2008)
NDVI	Normalized Difference Vegetation Index	(Rouse et al., 1974)
GNDVI	Green Normalized Difference Vegetation Index	(Gitelson et al., 1996)
EVI	Enhanced Vegetation Index	(Huete et al., 2002)
SAVI	Soil-Adjusted Vegetation Index	(Huete, 1988)
NDRE	Normalized Difference Red Edge	(Gitelson and Merzlyak, 1994)
RDVI	Renormalized Difference Vegetation Index	(Roujean and Breon, 1995)

the early season, one month after planting, the CAE achieves an R^2 of 0.785 compared to the traditional autoencoder’s 0.67, representing an 11% improvement and around 4% improvement over vegetation indices-driven predictions. This accuracy increases further at 1.5 months post-planting, with the CAE latent vector based predictor reaching an R^2 of 0.793. In the late-season, the final timepoint, close to harvest, the CAE maintains a high R^2 of 0.757, significantly outperforming the traditional autoencoder’s 0.45 and also those of vegetation indices by 6%. To statistically validate these performance gains, we conducted pairwise ANOVA tests across all timepoints. The results show statistically significant differences in model performance for time points 1 and 4, with CAE significantly outperforming the baseline methods (see [Supplementary Tables S5–S8](#)). Such early predictions could help growers make timely decisions about resource allocation, potentially improving crop management, resource use efficiency, and yield.

3.2.2 Experiment 2: rank ordering of top yielding genotypes

Breeders often rank order varieties to make breeding selections. We evaluated the ability of the CAE latent vector-based predictor to rank order the top 25% and top 50% performing genotypes. We consider time-point 2. This is around the time all the maize plants are near the end of the vegetative stage. We trained the same XGBoost model as before (which uses the disentangled latents as inputs and the yield as outputs) and performed a leave-one-location-out evaluation. That is, for evaluating the performance of a location, we trained the model on all the other locations and tested the model on that location.

TABLE 7 Performance on predicting top 25% and top 50% of the genotype for different locations was obtained via a model trained from CAE (compositional autoencoder) latents, VIs (vegetation indices), and CAE + VIs.

Location	VIs	CAE	CAE+VIs
(a) Top 25%			
Ames	44.44	50.0	27.77
Crawfordsville	38.33	50.0	50.0
Lincoln	50.0	55.55	50.0
MOValley	27.78	55.55	33.33
Scottsbluff	44.33	33.33	38.88
(b) Top 50%			
Ames	56.75	56.75	67.56
Crawfordsville	62.16	64.86	70.27
Lincoln	37.84	59.45	56.75
MOValley	45.94	75.67	78.37
Scottsbluff	40.54	56.75	51.35

For testing the model for a location, we trained the model on data from all other locations and then predicted yield for all the samples from that location. Then, we rank ordered the predictions and compared them against the ground truth rank ordering of the genotypes. The performance was evaluated by calculating the percentage of overlap between the top 25% or top 50% of the ground truth rank order list and the predicted rank order list. All the numbers given below are in percentages. The numbers in bold represent the best performance observed for a particular location.

For example, for evaluating ‘Ames’, we trained the model on the data from all the other locations and then tested its performance on that of ‘Ames’. To understand how well a model is able to capture the top $n\%$ highest-performing genotypes, we consider the list of top-yielding genotypes for each location and then calculate the percentage overlap of the top-yielding genotypes from the respective predictions. [Tables 7a, b](#) show the performance of CAE features and vegetation indices for the top 25% and top 50%. We observe that the model trained on CAE features consistently outperforms the model trained using vegetation indices, with one exception for the top 25% (Scottsbluff). We can also see that using both of these feature sets together does not seem to improve the performance for most cases. To better understand the outcome, we examined the cross-correlation between features from CAE and VIs as shown in [Figure 11](#). The plot shows relatively high correlations between several VIs and the z_e components from the CAE. This overlap may be introducing redundancy, which could be affecting the combined model’s performance. We also note that the CAE currently processes images from different locations based on calendar timepoints. Aligning the images by growth stages, rather than dates, might help improve the consistency of the encoded information and could potentially lead to better model performance in future work.

3.2.3 Experiment 3: performance on an unseen environment

In the third experiment, we evaluated the CAE’s performance in predicting yield in a completely new environment. To do this, we used data from Ames during the 2023 growing season to test the

Cross-Correlations Between CAE latent features (CAE) and Vegetative indices (VIs)

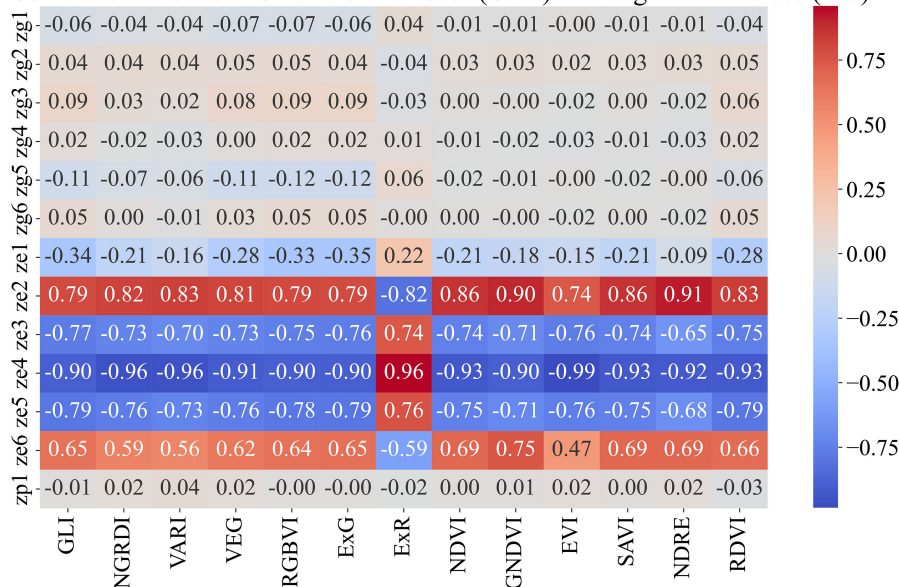


FIGURE 11

Feature correlations for combined feature space (CAE+VIs). We observe that the VIs show high correlation with the 'ze' components generated by CAE. High correlations between these feature sets will introduce redundant information, which could negatively affect the model's performance.

downstream prediction accuracy of a model that uses the CAE latent representation as inputs. Specifically, a disentangled latent representation was first generated via the CAE on all available (unlabeled) data – three from 2022 and three from 2023 (Ames, Lincoln, and MO Valley). These disentangled latent features were then used to train a downstream XGBoost model to predict yield. The downstream model was trained on all the data except Ames 2023. We then tested this model on Ames 2023 data and observed an RMSE of 27.815, demonstrating an improvement over traditional methods such as linear mixed models (RMSE = 29.91) and vegetation indices (RMSE = 45.80). The CAE also outperformed large pre-trained models like ResNet-18, which had an RMSE of 48. This highlights the ability of the CAE to generalize well to new environmental conditions, even when models like ResNet-18 (He et al., 2016), with its 11.7 million parameters and deeper architecture, struggled.

4 Conclusion

This study extends our previous work on disentangling Genotype x Environment (GxE) features by applying a Compositional Autoencoder (CAE) to create disentangled latent representations of satellite imagery. These disentangled latent representations produced improved yield prediction. Our results demonstrate the CAE's effectiveness in separating environmental factors and improving yield predictions at various growth stages. The CAE outperformed raw satellite data in distinguishing macro-environmental factors, as evidenced by the improvement in silhouette scores from 0.293 to 0.919. This enhanced separation of environmental features suggests the CAE's potential for more

precise modeling of environmental influences on crop performance. The model also showed promising results in disentangling micro-environmental factors.

In terms of yield prediction, regression (XGBoost) models trained on CAE-based latent representations consistently outperformed regression models trained on latent representations from vanilla autoencoders across all time points and beat models trained on VIs for yield prediction for the vegetative stage and post-flowering stage, with competitive performance during the flowering stage. This enhanced early-stage prediction capability could provide breeders with valuable insights for resource management throughout the growing season.

The genotype features extracted by the CAE could also be valuable for genome-wide association studies, offering a new way to link genetic markers to complex traits. Additionally, exploring the CAE's use with other sensing modalities and applying it to time-series data may further improve its predictive capabilities and reveal new biological insights. Combining the CAE's disentangled latent representations with other data sources, such as crop models or physiological measurements, could lead to enhanced end-of-season trait prediction models. While this study focused on high-resolution (30 cm) Pleiades Neo imagery to enable fine-grained plot-level yield prediction, an important direction for future work is to evaluate the performance and generalizability of the CAE framework across coarser resolutions (e.g., 1 m, 3 m, and 10 m) to understand the trade-offs between spatial resolution, model accuracy, and operational scalability for broader deployment. Another important direction is to examine how pixel size relates to maize plant size, incorporating factors such as planting density and pixel-level variability, and to assess how these spatial effects correlate with trait variability among maize hybrids.

Data availability statement

All code is available at bitbucket at https://bitbucket.org/baskargroup/cae_sat/src/main/. The satellite plot-level images from multiple time points and locations of experimental fields, along with ground truth data, are accessible at <https://doi.org/doi:10.5061/dryad.905qfttm>.

Author contributions

AP: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. TJ: Writing – original draft, Writing – review & editing, Formal analysis, Methodology, Software, Validation. MT: Writing – review & editing, Validation. NS: Writing – review & editing, Data curation, Validation. LC: Writing – review & editing, Data curation. JS: Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Resources, Supervision. PS: Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Resources, Supervision. BG: Writing – review & editing, Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This research was supported in part by the U.S. Department of Agriculture, National Institute of Food and Agriculture under award numbers 2020-67021–31528 and 2020-68013–30934 and the AI Research Institutes program supported by NSF and USDA-NIFA under the AI Institute for Resilient Agriculture, Award No. 2021-67021-35329.

References

- Alexander, T. A., Irizarry, R. A., and Bravo, H. C. (2022). Capturing discrete latent structures: choose LDs over PCs. *Biostatistics* 24, 1–16. doi: 10.1093/biostatistics/kxab030
- Bai, J., Chen, X., Dobermann, A., Yang, H., Cassman, K. G., and Zhang, F. (2010). Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. *Agron. J.* 102, 9–16. doi: 10.2134/agronj2009.0085
- Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (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 Observ. Geoinform.* 39, 79–87. doi: 10.1016/j.jag.2015.02.012
- Broms, C., Nilsson, M., Oxenstierna, A., Sopasakis, A., and Åström, K. (2023). Combined analysis of satellite and ground data for winter wheat yield forecasting. *Smart Agric. Technol.* 3, 100107. doi: 10.1016/j.atech.2022.100107
- Butrón, A., Romay, M. C., Peña-Asín, J., Alvarez, A., and Malvar, R. A. (2012). Genetic relationship between maize resistance to corn borer attack and yield. *Crop Sci.* 52, 1176–1180. doi: 10.2135/cropsci2011.11.0584
- Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., et al. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agric. For. Meteorol.* 274, 144–159. doi: 10.1016/j.agrformet.2019.03.010
- Cooper, M., Gho, C., Leafgren, R., Tang, T., and Messina, C. (2014). Breeding drought-tolerant maize hybrids for the US corn-belt: discovery to product. *J. Exp. Bot.* 65, 6191–204. doi: 10.1093/jxb/eru064
- de Wit, A. J. W., and van Diepen, C. A. (2008). Crop growth modelling and crop yield forecasting using satellite-derived meteorological inputs. *Int. J. Appl. Earth Obs. Geoinf.* 10, 414–425. doi: 10.1016/j.jag.2007.10.004
- Du, L., Yang, H., Song, X., Wei, N., Yu, C., Wang, W., et al. (2022). Estimating leaf area index of maize using UAV-based digital imagery and machine learning methods. *Sci. Rep.* 12, 15937. doi: 10.1038/s41598-022-20299-0
- Dumey, I., Valero, S., and Inglada, J. (2024). Self-supervised spatio-temporal representation learning of satellite image time series. *IEEE J. Select. Topics Appl. Earth Observ. Remote Sens.* PP, 1–18. doi: 10.1109/JSTARS.2024.3358066

Conflict of interest

JS has equity interests in Data2Bio, LLC, and Dryland Genetics Inc., and has performed paid work for Alphabet. He is a member of the scientific advisory boards of GeneSeek and Aflo Sensors. PS is a co-founder, managing partner, and/or CEO of Data2Bio, LLC, and Dryland Genetics, Inc. He is also a member of the scientific advisory boards of Kemin Industries and Centro de Tecnologia Canavieira.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1617831/full#supplementary-material>

- Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., and Prasanna, B. (2022). Global maize production, consumption and trade: Trends and R&D implications. *Food Secur.* 14, 1295–1319. doi: 10.1007/s12571-022-01288-7
- Feldmann, M. J., Gage, J. L., Turner-Hissong, S. D., and Ubbens, J. R. (2021). Images carried before the fire: The power, promise, and responsibility of latent phenotyping in plants. *Plant Phenome J.* 4, e20023. doi: 10.1002/ppj2.20023
- Gage, J. L., Richards, E., Lepak, N., Kaczmar, N., Soman, C., Chowdhary, G., et al. (2019). In-field whole-plant maize architecture characterized by subcanopy rovers and latent space phenotyping. *Plant Phenome J.* 2, 190011. doi: 10.2135/tppj2019.07.0011
- Gamble, E. E. (1962). Gene effects in corn (*Zea mays* L.): I. Separation and relative importance of gene effects for yield. *Can. J. Plant Sci.* 42, 339–348. doi: 10.4141/cjps62-048
- Gitelson, A. A., Kaufman, Y. J., and Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298. doi: 10.1016/S0034-4257(96)00072-7
- Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sens. Environ.* 80, 76–87. doi: 10.1016/S0034-4257(01)00289-9
- Gitelson, A., and Merzlyak, M. N. (1994). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B: Biol.* 22, 247–252. doi: 10.1016/1011-1344(93)06963-4
- Gomari, D. P., Schweickart, A., Cerchietti, L., Paietta, E., Fernandez, H., Al-Amin, H., et al. (2022). Variational autoencoders learn transferrable representations of metabolomics data. *Commun. Biol.* 5, 645. doi: 10.1038/s42003-022-03579-3
- Goyal, P., Kaur, A., Ram, A., and Goyal, N. (2024). “Efficient representation learning of satellite image time series and their fusion for spatiotemporal applications,” in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-24)*, (Washington, DC, USA: AAAI Press), Vol. 38. 8436–8444.
- Hague, T., Tillett, N. D., and Wheeler, H. (2006). Automated crop and weed monitoring in widely spaced cereals. *Precis. Agric.* 7, 21–32. doi: 10.1007/s11119-005-6787-1
- Hamar, D., Ferencz, C., Lichtenberger, J., Tarcas, G., and Ferencz-Arkos, I. (1996). Yield estimation for corn and wheat in the Hungarian Great Plain using Landsat MSS data. *Int. J. Remote Sens.* 17, 1689–1699. doi: 10.1080/01431169608948732
- Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., et al. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* 15, 10. doi: 10.1186/s13007-019-0394-z
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016)*, (Las Vegas, NV, USA: IEEE Computer Society), 770–778. doi: 10.1109/CVPR.2016.90
- Huete, A. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25, 295–309. doi: 10.1016/0034-4257(88)90106-X
- Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. The Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring. doi: 10.1016/0034-4257(88)90106-X
- Iwasaki, D., Cooray, S., and Takeuchi, T. T. (2023). *Extracting an Informative Latent Representation of High-Dimensional Galaxy Spectra* (arXiv). 2311.17414[astro-ph.GA].
- Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., et al. (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nat. Food* 2, 873–885. doi: 10.1038/s43016-021-00400-y
- Jeong, S., Ko, J., and Yeom, J. M. (2018). Nationwide projection of rice yield using a crop model integrated with geostationary satellite imagery: A case study in South Korea. *Remote Sens. (Basel)* 10, 1665. doi: 10.3390/rs10101665
- Jin, Z., Azzari, G., and Lobell, D. B. (2017). Improving the accuracy of satellite-based high-resolution yield estimation: A test of multiple scalable approaches. *Agric. For. Meteorol.* 247, 207–220. doi: 10.1016/j.agrformet.2017.08.001
- Johnson, M. D., Hsieh, W. W., Cannon, A. J., Davidson, A., and Bédard, F. (2016b). Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. *Agric. For. Meteorol.* 218–219, 74–84. doi: 10.1016/j.agrformet.2015.11.003
- Joshi, D. R., Clay, S. A., Sharma, P., Rekabdarkolae, H. M., Kharel, T., Rizzo, D. M., et al. (2023). Artificial intelligence and satellite-based remote sensing can be used to predict soybean (Glycine max) yield. *Agron. J.* 260, 107984. doi: 10.1002/agj2.21473
- Kamir, E., Waldner, F., and Hochman, Z. (2020). Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. *ISPRS J. Photogrammetry Remote Sens.* 160, 124–135. doi: 10.1016/j.isprsjprs.2019.11.008
- Kanning, M., Kühling, I., Trautz, D., and Jarmer, T. (2018). High-resolution UAV-based hyperspectral imagery for LAI and chlorophyll estimations from wheat for yield prediction. *Remote Sens.* 10, 2000. doi: 10.3390/rs10122000
- Khalil, Z., and Abdullaev, S. (2021). Neural network for grain yield predicting based multispectral satellite imagery: comparative study. *Proc. Comput. Sci.* 186, 269–278. 14th International Symposium “Intelligent Systems. doi: 10.1016/j.procs.2021.04.146
- Kopf, A., and Claassen, M. (2021). Latent representation learning in biology and translational medicine. *Patterns (N. Y.)* 2, 100198. doi: 10.1016/j.patter.2021.100198
- Kusmec, A., de, L. N., and Schnable, P. S. (2018). Harnessing phenotypic plasticity to improve maize yields. *Front. Plant Sci.* 9, 1377. doi: 10.3389/fpls.2018.01377
- Lang, P., Zhang, L., Huang, C., Chen, J., Kang, X., Zhang, Z., et al. (2022). Integrating environmental and satellite data to estimate county-level cotton yield in Xinjiang Province. *Front. Plant Sci.* 13, 1048479. doi: 10.3389/fpls.2022.1048479
- Li, Z., Ding, L., and Xu, D. (2022). Exploring the potential role of environmental and multi-source satellite data in crop yield prediction across Northeast China. *Sci. Total Environ.* 815, 152880. doi: 10.1016/j.scitotenv.2021.152880
- Li, Z., Zhang, J., and Song, X. (2024). Breeding maize hybrids with improved drought tolerance using genetic transformation. *Int. J. Mol. Sci.* 25, 10630. doi: 10.3390/ijms251910630
- Lobell, D. B., Deines, J. M., and Tommaso, S. D. (2020). Changes in the drought sensitivity of US maize yields. *Nat. Food* 1, 729–735. doi: 10.1038/s43016-020-00165-w
- Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., et al. (2014). Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science* 344, 516–519. doi: 10.1126/science.1251423
- López-Granados, F., Torres-Sánchez, J., De Castro, A. I., Serrano-Pérez, A., Mesas-Carrascosa, F. J., and Peña, J. M. (2016). Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery. *Agron. Sustain. Dev.* 36, 67. doi: 10.1007/s13593-016-0405-7
- Luo, L., Sun, S., Xue, J., Gao, Z., Zhao, J., Yin, Y., et al. (2023). Crop yield estimation based on assimilation of crop models and remote sensing data: A systematic evaluation. *Agric. Syst.* 210, 103711. doi: 10.1016/j.agry.2023.103711
- Mahalakshmi, S., Jose Anand, A., and Partheeban, P. (2025). Soil and crop interaction analysis for yield prediction with satellite imagery and deep learning techniques for the coastal regions. *J. Environ. Manage.* 380, 125095. doi: 10.1016/j.jenvman.2025.125095
- Mateo-Sanchis, A., Piles, M., Muñoz-Marí, J., Adsua, J. E., Pérez-Suay, A., and Camps-Valls, G. (2019). Synergistic integration of optical and microwave satellite data for crop yield estimation. *Remote Sens. Environ.* 234, 111460. doi: 10.1016/j.rse.2019.111460
- Meroni, M., Waldner, F., Seguí, L., Kerdiles, H., and Rembold, F. (2021). Yield forecasting with machine learning and small data: What gains for grains? *Agric. For. Meteorol.* 308309, 108555. doi: 10.1016/j.agrformet.2021.108555
- Meyer, G. E., and Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* 63, 282–293. doi: 10.1016/j.compag.2008.03.009
- Mounir Louhaichi, M. M. B., and Johnson, D. E. (2001). Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto Int.* 16, 65–70. doi: 10.1080/10106040108542184
- Nakayama, Y., and Su, J. (2024). *Spatio-Temporal SwinMAE: A Swin Transformer based Multiscale Representation Learner for Temporal Satellite Imagery* (ArXiv). abs/2405.02512.
- Nieto, L., Schwalbert, R., Prasad, P. V. V., Olson, B. J. S. C., and Ciampitti, I. A. (2021). An integrated approach of field, weather, and satellite data for monitoring maize phenology. *Sci. Rep.* 11, 15711. doi: 10.1038/s41598-021-95253-7
- Nolan, E., and Santos, P. (2012). The contribution of genetic modification to changes in corn yield in the United States. *Am. J. Agric. Econ* 94, 1171–1188. doi: 10.1093/ajae/aas069
- Nourelidin, N., Aboelghar, M., Saudy, H., and Ali, A. (2013). Rice yield forecasting models using satellite imagery in Egypt. *Egyptian J. Remote Sens. Space Sci.* 16, 125–131. doi: 10.1016/j.ejrs.2013.04.005
- Okada, M., Barras, C., Toda, Y., Hamazaki, K., Ohmori, Y., Yamasaki, Y., et al. (2024). High-throughput phenotyping of soybean biomass: conventional trait estimation and novel latent feature extraction using UAV remote sensing and deep learning models. *Plant Phenomics* 6, 0244. doi: 10.34133/plantphenomics.0244
- Ort, D. R., and Long, S. P. (2014). Limits on yields in the corn belt. *Science* 344, 484–485. doi: 10.1126/science.1253884
- Peralta, N. R., Assefa, Y., Du, J., Barden, C. J., and Ciampitti, I. A. (2016). Mid-season high-resolution satellite imagery for forecasting site-specific corn yield. *Remote Sens.* 8, 848. doi: 10.3390/rs8100848
- Phang, S. K., Chiang, T. H. A., Happonen, A., and Chang, M. M. L. (2023). From satellite to UAV-based remote sensing: A review on precision agriculture. *IEEE Access* 11, 127057–127076. doi: 10.1109/ACCESS.2023.3330886
- Powadi, A., Jubery, T. Z., Tross, M. C., Schnable, J. C., and Ganapathysubramanian, B. (2024). Disentangling genotype and environment specific latent features for improved trait prediction using a compositional autoencoder. *Front. Plant Sci.* 15, 1476070. doi: 10.3389/fpls.2024.1476070
- Roujean, J. L., and Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* 51, 375–384. doi: 10.1016/0034-4257(94)00114-3
- Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Spec. Publ.* 351, 309. Available online at: <https://ui.adsabs.harvard.edu/abs/1974NASSP.351..309R>.
- Roznik, M., Boyd, M., and Porth, L. (2022). Improving crop yield estimation by applying higher resolution satellite NDVI imagery and high-resolution cropland masks. *Remote Sens. Appl.: Soc. Environ.* 25, 100693. doi: 10.1016/j.rsae.2022.100693

- Sahbeni, G., Sz'ekely, B., Musyimi, P. K., Tim'ar, G., and Sahajpal, R. (2023). Crop yield estimation using sentinel-3 SLSTR, soil data, and topographic features combined with machine learning modeling: A case study of Nepal. *AgriEngineering* 5, 1766–1788. doi: 10.3390/agriengineering5040109
- Schwalbert, R., Amado, T., Nieto, L., Corassa, G., Rice, C., Peralta, N., et al. (2020). Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. *Agric. For. Meteorol.* 284, 107886. doi: 10.1016/j.agrformet.2019.107886
- Schwalbert, R. A., Amado, T. J. C., Nieto, L., Varela, S., Corassa, G. M., Horbe, T. A. N., et al. (2018). Forecasting maize yield at field scale based on high-resolution satellite imagery. *Biosyst. Eng.* 171, 179–192. doi: 10.1016/j.biosystemseng.2018.04.020
- Schwalbert, R., Amado, T., Nieto, L., Corassa, G., Rice, C., Peralta, N., et al. (2020). Mid-season county-level corn yield forecast for US Corn Belt integrating satellite imagery and weather variables. *Crop Sci.* 60, 739–750. doi: 10.1002/csc2.20053
- Shendryk, Y., Davy, R., and Thorburn, P. (2021). Integrating satellite imagery and environmental data to predict field-level cane and sugar yields in Australia using machine learning. *Field Crops Res.* 260, 107984. doi: 10.1016/j.fcr.2020.107984
- Shrestha, N., Powadi, A., Davis, J., Ayanlade, T. T., Liu, H., Tross, M. C., et al. (2025). Plot-level satellite imagery can substitute for UAVs in assessing maize phenotypes across multistate field trials. *Plants People Planet* 7, 1011–1026. doi: 10.1002/ppp3.10613
- Song, M. K., Niaz, A., and Choi, K. N. (2023). "Image generation model applying PCA on latent space," in *Proceedings of the 2nd Asia Conference on Algorithms, Computing and Machine Learning (ACML 2023)*, (Shanghai, China: ACM), 419–423.
- Tarekne, A., Wegary, D., Cairns, J. E., Zaman-Allah, M., Beyene, Y., Negera, D., et al. (2024). Genetic gains in early maturing maize hybrids developed by the International Maize and Wheat Improvement Center in Southern Africa during 2000–2018. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1321308
- Tross, M. C., Grzybowski, M. W., Jubery, T. Z., Grove, R. J., Nishimwe, A. V., Torres-Rodriguez, J. V., et al. (2023). Data driven discovery and quantification of hyperspectral leaf reflectance phenotypes across a maize diversity panel. *bioRxiv*. 15, 1476070. doi: 10.1002/ppj2.20106
- Troyer, A. F. (1990). A retrospective view of corn genetic resources. *J. Heredity* 81, 17–24. doi: 10.1093/oxfordjournals.jhered.a110920
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150. doi: 10.1016/0034-4257(79)90013-0
- Ubbens, J., Cieslak, M., Prusinkiewicz, P., Parkin, I., Ebersbach, J., and Stavness, I. (2020). Latent space phenotyping: automatic image-based phenotyping for treatment studies. *Plant Phenomics* 2020, 13. doi: 10.34133/2020/5801869
- Wang, Y. P., Chang, K. W., Chen, R. K., Lo, J. C., and Shen, Y. (2010). Large-area rice yield forecasting using satellite imageries. *Int. J. Appl. Earth Observ. Geoinform.* 12, 27–35. doi: 10.1016/j.jag.2009.09.009
- Woebbecke D, M., Meyer G, E., Von Bargen, K., and Mortensen, D. (1995). Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASAE* 38, 259–269. doi: 10.13031/2013.27838
- Yang, C., Everitt, J. H., and Bradford, J. M. (2006). Comparison of quickBird satellite imagery and airborne imagery for mapping grain sorghum yield patterns. *Precis. Agric.* 7, 33–44. doi: 10.1007/s11119-005-6788-0
- Yang, C., Everitt, J., and Bradford, J. (2009). Evaluating high resolution SPOT 5 satellite imagery to estimate crop yield. *Precis. Agric.* 10, 292–303. doi: 10.1007/s11119-009-9120-6
- Yang, C., Everitt, J. H., Du, Q., Luo, B., and Chanussot, J. (2013). Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture. *Proc. IEEE* 101, 582–592. doi: 10.1109/JPROC.2012.2196249
- Yang, Y., Tilman, D., Jin, Z., Smith, P., Barrett, C. B., Zhu, Y. G., et al. (2024). Climate change exacerbates the environmental impacts of agriculture. *Science* 385, eadn3747. doi: 10.1126/science.adn3747
- You, J., Li, X., Low, M., Lobell, D., and Ermon, S. (2017). "Deep gaussian process for crop yield prediction based on remote sensing data," in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI-17)*, (Washington, DC, USA: AAAI Press). Vol. 31.
- Zare, H., Weber, T. K. D., Ingwersen, J., Nowak, W., Gayler, S., and Streck, T. (2022). Combining crop modeling with remote sensing data using a particle filtering technique to produce real-time forecasts of winter wheat yields under uncertain boundary conditions. *Remote Sens. (Basel)* 14, 1360. doi: 10.3390/rs14061360
- Zhao, Y., Potgieter, A. B., Zhang, M., Wu, B., and Hammer, G. L. (2020). Predicting wheat yield at the field scale by combining high-resolution Sentinel-2 satellite imagery and crop modelling. *Remote Sens. (Basel)* 12, 1024. doi: 10.3390/rs12061024
- Zhong, G., Wang, L. N., Ling, X., and Dong, J. (2016). An overview on data representation learning: From traditional feature learning to recent deep learning. *J. Finance Data Sci.* 2, 265–278. doi: 10.1016/j.jfids.2017.05.001