

Context-Aware Deep Learning Model for Yield Prediction in Potato Using Time-Series UAS Multispectral Data

Suraj A. Yadav^{ID}, Xin Zhang^{ID}, Nuwan K. Wijewardane^{ID}, Max Feldman, Ruijun Qin, Yanbo Huang^{ID}, Sathishkumar Samiappan^{ID}, Wyatt Young^{ID}, and Francisco G. Tapia^{ID}

Abstract—The study demonstrated the efficacy of integrating time-series uncrewed aerial system (UAS) multispectral imaging with data-driven deep learning methodologies to systematically and precisely predict field-scale crop yield throughout the growing seasons. A UAS equipped with a micasense rededge MX+ sensor was used for data acquisition at the Hermiston Agricultural Research and Extension Center, Oregon State University. The data were collected throughout the potato (*Solanum tuberosum L.*) growing seasons under varied nitrogen (N)-rates ranging from 0 to 639 kg/ha. The raw data were preprocessed using Pix4Dmapper and the quantum geographic information system. A linear unmixing model followed by Otsu-based adaptive autosegmentation was implemented to generate soil-masked spatio-spectral fusion maps for accurate vegetation feature extraction. The proposed feature engineering and prediction model followed a two-fold approach: first, adoption of partial least squares regression (PLSR) algorithm to extract features relevant to yield, and second, a novel context-aware attention and residual connection convolution-bidirectional gated recurrent unit bidirectional long short-term memory-network (CAR Conv1D-BiGRU-BiLSTM-Net) to exploit time-series multifeatures information to predict final yield. On integrating the PLSR-derived robust features, the proposed model demonstrated an increase in predictive capability from emergence (T1) to bulking (T4) growth stage by effectively capturing the temporal dynamics of physiological and biological traits. Overall, using multifeatures such as

Received 13 September 2024; revised 27 January 2025; accepted 3 February 2025. Date of publication 5 February 2025; date of current version 27 February 2025. This research is based upon the work supported by the United States Department of Agriculture – Agricultural Research Service (USDA-ARS) under agreement number 58-6064-3-007 and 58-6064-3-009. (Corresponding author: Suraj A. Yadav.)

Suraj A. Yadav, Xin Zhang, and Nuwan K. Wijewardane are with the Department of Agricultural and Biological Engineering, Mississippi State University, Mississippi State, MS 39762 USA (e-mail: say70@msstate.edu; xzhang@abe.msstate.edu; nuwanw@abe.msstate.edu).

Max Feldman is with the Temperate Tree Fruit and Vegetable Research Unit, USDA-ARS, Wenatchee, WA 98851 USA (e-mail: max.feldman@usda.gov).

Ruijun Qin is with the Hermiston Agricultural Research and Extension Center, Oregon State University, Hermiston, OR 97838 USA (e-mail: ruijun.qin@oregonstate.edu).

Yanbo Huang and Wyatt Young are with the Genetics and Sustainable Agriculture Research Unit, USDA-ARS, Mississippi State, MS 39762 USA (e-mail: yanbo.huang@usda.gov; wyatt.young@usda.gov).

Sathishkumar Samiappan is with the Department of Biosystems Engineering and Soil Sciences, University of Tennessee, Knoxville, TN 37996 USA (e-mail: sathish@utk.edu).

Francisco G. Tapia is with the Forage Seed and Cereal Research Unit, USDA-ARS, Corvallis, OR 97331 USA (e-mail: paco.gonzalez@usda.gov).

The source code is available at <https://github.com/SAY70/CAR-CNN-BiRNN>.

This article has supplementary downloadable material available at <https://doi.org/10.1109/JSTARS.2025.3539217>, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2025.3539217

simple ratio, Chlorophyll Green, modified anthocyanin reflectance index, vegetation fraction (V_f), and N-rate from T1–T4 growth stage resulted in predictive accuracy with high $R^2 = 0.775$ and low root mean square error of 16.4%, outperforming other deep learning models.

Index Terms—Context-aware attention mechanism and residual connection based one-dimensional convolution-bidirectional gated recurrent unit-bidirectional long short-term memory-network (CAR Conv1D-BiGRU-BiLSTM-Net), feature engineering, multispectral imaging, PLSR, segmentation.

I. INTRODUCTION

AIRBORNE or spaceborne imaging spectroscopy, commonly known as multispectral/hyperspectral remote sensing, has emerged as a vital technology with diverse applications, encompassing environmental monitoring with agriculture [1], [2], forestry [3], oceanography [4], Earth observation [5], glaciology [6], and national security [7]. This technology has gained significant attention due to the proliferation of drone industries and both existing and forthcoming missions dedicated to imaging spectroscopy, such as NASA's airborne visible/infrared imaging spectrometer [8], hyperspectral thermal emission spectrometer [9], surface biology and geology [10], Copernicus hyperspectral imaging mission for the environment [11], hyperspectral imager suite [12], AggieAir [13], precursore iperspettrale della missione applicativa [14], and environmental mapping and analysis program [15]. Collectively, these missions have capabilities to comprehensively investigate and characterize terrestrial and agroecosystem at multiple scales, ranging from plot/field, county, and state levels to global extents.

The advent of these imaging technologies has greatly encouraged scientists and researchers to harness the power of uncrewed aerial system (UAS)-based imaging spectroscopy to support endeavors at field scale. UAS's can provide the flexibility and agility needed to capture high spatial, spectral, and temporal resolution data at field scales [13]. The field scale in situ observations, such as soil characteristics, weather information, biophysical, and biochemical parameters holds significant potential in the development of more robust and generalized multispectral/hyperspectral imaging applications. In addition, the temporal data obtained from UAS imaging from crops under diverse nitrogen (N)-rate can lead to accurate and efficient time-series approaches in precision agriculture, such as early-stage

disease/anomaly and weed detection for spot spraying, classification of N-rich soil for cost-saving fertilization, and vegetation index or chlorophyll map for assessing health and prediction analysis [1], [2], [16], [17].

Potatoes are globally cultivated and consumed, serving as a staple food rich in essential nutrients such as carbohydrates, vitamin C, potassium, and fiber [18]. From an economic perspective, potatoes represent a major source of income for farmers and a key export commodity for many countries. In 2023, potato production in the United States was estimated at 418 million hundredweight units (i.e., 19 942.4 million kilograms), reflecting a 9% increase compared to the previous year [19]. This growth highlights the importance of accurate yield estimation, which is crucial for optimizing resource allocation, planning for market demands, and stabilizing global trade revenues. The growth and development of potato plants are highly influenced by N management, influencing key physiological traits, such as chlorophyll content, leaf area index, and biomass accumulation, which directly impact yield outcomes [20]. Studies have shown that the N-rate variations within a field influences the patterns of spectral and spatial features detectable by UAS imaging platform [21], [22]. Therefore, incorporating N-rates along with spectral and spatial features in precise yield prediction models are essential, as it accounts for the wide range of variability caused by N availability. However, extracting these features across different growth stages and varying N-rates is challenging due to changes in plant growth, soil properties, and mixed pixels. In addition, UAS data are often captured under diverse lighting and atmospheric conditions, leading to noise and inconsistencies in the data, which can affect the accuracy of vegetation features. Thus, data preprocessing and computer vision techniques play crucial roles in addressing these challenges, ensuring precise and accurate data extraction for yield prediction. Preprocessing steps such as radiometric, atmospheric, and geometric corrections are essential for obtaining meaningful data from UAS imagery and computer vision frameworks, including linear unmixing (LUM) [23], nonnegative matrix factorization [24], and universal pattern decomposition matrix [25], can enhance feature extraction by classifying vegetation abundances at different growth stages. Integrating these frameworks with autosegmentation algorithms improves the process of accurately extracting spatial and spectral features of vegetation, particularly, during challenging growth stages where auto distinguishing threshold for soil and vegetation pixels is difficult [26].

Various machine learning (ML) methodologies have also demonstrated remarkable efficacy in addressing complex nonlinear tasks, such as feature extraction, classification, regression, and predictive modeling utilizing time-series data mining from UAS data [16], [27], [28], [29], [30], [31], [32]. For instance, Maimaitijiang et al. [16] highlighted the potential of a generic ML models utilizing UAS-based multisource data for soybean yield prediction. Results demonstrated the enhanced accuracy of yield prediction and adaptability to spatial variations with good correlation and root mean square error (RMSE) ($R^2 = 0.720$ and RMSE% = 15.9%). Ma et al. [30] applied a Bayesian neural network (BNN) model on corn crop, utilizing large set of variable including remotely sensed vegetation indices, climate data, soil property, and corn yield, which achieved an average

$R^2 = 0.77$ for late-season prediction across the U.S. Corn Belt. The study enveloped over 84% of observed yield records within its predictive uncertainty interval, emphasizing the importance of understanding environmental stress effects on crop yield estimation. Ali et al. [31] used four ML algorithms, namely, decision tree, random forest, gradient boosting (GB), and extreme GB to predict wheat grain yield utilizing high-throughput UAS multispectral imaging with field scale precision. Zhao et al. [33] developed a robust spectral–spatial agricultural crop mapping method based on conditional random fields to address challenges in heterogeneous crop identification, demonstrating significant improvements in classification performance using data from UAS platforms. Lin et al. [34] introduced a target-less framework for temporal UAS-based light detection and ranging (LiDAR) data crop characterization and quality control, using terrain patches and row/alley locations to evaluate point cloud accuracy. The proposed method effectively identifies rows and alleys in various field conditions, achieving vertical and planimetric accuracy of ± 3 cm and ± 8 cm, respectively, thereby facilitating automated phenotyping. Verma et al. [32] showed the importance of partial least squares regression (PLSR) algorithm to find optimum vegetation features for biophysical and biochemical parameter retrieval. Sun et al. [35] utilized a combined recurrent and convolutional neural network model that can integrate vegetation spatial and temporal features, weather, and soil data to predict corn yield at the county level using satellite data in U.S. Corn Belt states from 2013 to 2016. The analysis showed that phenological factors significantly influence yield estimation.

Neural network architectures, CNN's, BNN's, and RNN's, have become the predominant methods for feature extraction in classification and prediction tasks. These architectures excel at capturing spatial hierarchies and learning intricate patterns from the data. In the context of time-series predictive analysis, recently, RNN's, particularly gated recurrent unit (GRU) and long short-term memory (LSTM) model, have gained significant interest due to their capability of mitigating the challenges of gradient instability, enabling the capture of short term and long-term dependencies within sequential data [36]. However, the quality of GRU/LSTM-based predictive model's may be constrained as they do not consider the backward propagation of temporal vegetation change information (i.e., future vegetation information). In view of the mentioned problem, an enhanced version known as bidirectional LSTM (BiLSTM) and bidirectional GRU (BiGRU), has been introduced by the authors in [37] and [38]. Bidirectional model leverages both forward and backward temporal states through its bidirectional structure, which integrates two layers (i.e., forward and backward). Thus, it can handle sequential data and capture temporal dependencies and, hence, suitable for prediction tasks [17], [39].

In addition to these advancements, the attention mechanism has emerged as a pivotal component in enhancing the performance of deep neural networks architectures. The attention mechanism, initially developed for translation tasks, allows the network to dynamically focus on the most relevant parts of the input data while minimizing the impact of less coherent or noisy data. This mechanism significantly improves the network's ability to learn and generalize from complex datasets by

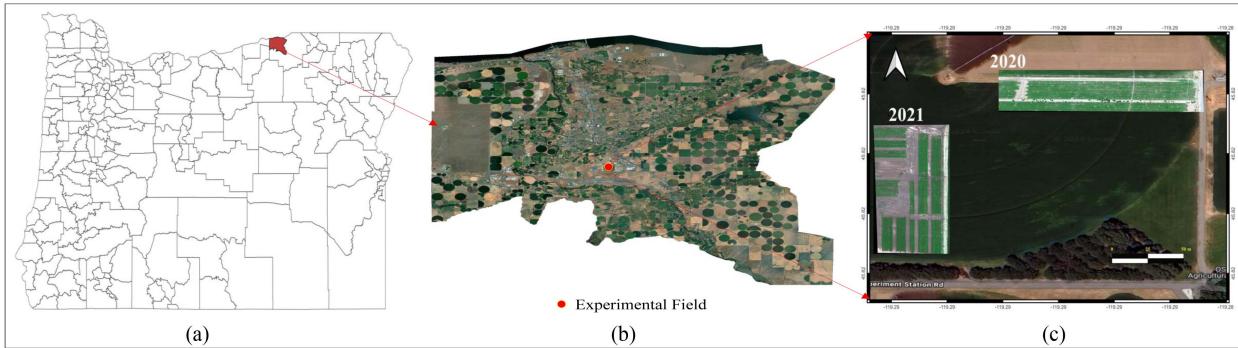


Fig. 1. Geographic illustration of experimental site located in Hermiston-Utmila county, Oregon, USA, highlighting Hermiston Agricultural Research and Extension Center. (a) Counties within Oregon state, USA with Hermiston-Utmila county highlighted. (b) Closer look of Hermiston-Utmila county overlayed on hybrid google satellite view. (c) Exact geographical location of the experimental field at the Oregon State University-Hermiston Agricultural Research and Extension Center (OSU-HAREC) where UAS data acquisition took place.

emphasizing critical features and relationships within the data [40]. Moreover, the integration of attention mechanisms into neural networks still requires substantial improvements to fully harness their potential in agricultural applications. Specifically, there is a notable scarcity of research dedicated to exploring the combined capacity of deep learning network architectures and attention mechanisms by integrating time-series crop features gathered from UAS data at field scale to predict yields comprehensively. Most of the predictive model's have been trained using single-stage multisource data at regional or county scales [30], [35], [41]. These models often lack detailed insights into the dynamic growth processes of crops, potentially leading to biased predictions. For instance, the use of regional scale data fails to capture microvariations in crop health and environmental conditions that are critical for precise yield prediction. To address gaps in yield forecasting with finer spatial and temporal resolution including vegetation growth dynamics, we propose a detailed, field scale new multispectral image processing pipeline for UAS-collected data. The pipeline integrates residual connected deep learning with context-aware attention mechanisms to effectively extract features importance and predict crop yield. As we have not found any existing research on this specific topic, our study is novel and contributes new knowledge to the scientific field of remote sensing using multispectral imagery.

The overarching goal of this study are to evaluate the impact of varied N-rates on the tuber yield of different potato cultivars precisely and to demonstrate the efficacy of integrating multispectral image processing pipeline with novel data-driven deep learning methodologies for yield forecasting at field scale precision with respect to different crop growth stages. The specific scientific objectives are as follows.

- 1) Design accurate experiment to investigate the yield responses of selected potato cultivars to varied N-rates and gather UAS multispectral imaging data that captures temporal changes in crop growth characteristics.
- 2) To develop a computer vision framework that integrates an LUM with a nonbinary adaptive semantic autosegmentation algorithm, utilizing multispectral images gathered by UAS. This approach aims to produce the spatio-spectral fusion map, allowing for precise segmentation and the extraction of spectral and spatial features.

- 3) Employ feature engineering utilizing the partial least squares regression (PLSR) algorithm on LUM maps to investigate feature importance and its relevance to yield prediction across different growth stages.
- 4) Implement a novel context-aware attention mechanism and residual connection (CAR)-based deep learning model to predict final yield, aiming to optimize agronomic practices and improve prediction accuracy.

II. STUDY AREA: FIELD DESIGN AND EXPERIMENTS

The field experiments were conducted in 2020 and 2021 on potato crop cultivation at Hermiston Agricultural Research and Extension Center, Oregon State University (HAREC-OSU) in Hermiston-Utmila county, Oregon, USA, as shown in Fig. 1(a). The experimental field is located at latitude 45° 8' 43.9548" N, longitude 119° 29' 33.5076" W, with an elevation of 140 m above sea level, as shown in Fig. 1(b). The trials were carried out within a smaller segment of the HAREC-OSU agricultural plot, encompassing approximately 8 acres, as illustrated in Fig. 1(c). The experimentation fields were facilitated by a center-pivot irrigation system. The split-plot design for allocating different N-rates is illustrated in Figs. 2(a) and 3(a), for 2020 and 2021, respectively. The allocation of different potato cultivars is depicted in Figs. 2(b) and 3(b), with three cultivars in 2020 [Russet Norkotah (RN), Rainier Russet (RR; newly released), and Russet Burbank (RB)] and four cultivars in 2021 [Alturas (AL), RN, RR, and RB]. The traditional cultivar RB, released in 1902 as May's Netted Gem, is known for its long tuber dormancy [18]. Another well-known cultivar is the RN, released in 1987, which has been studied for its response to drought stress, showing varying degrees of reduction in tuber yield under stress conditions [42]. RR is a modern potato cultivar introduced in 2004, known for its promising yield performance, quality, and resistance to diseases and pests, making it a sustainable choice for agriculture. This cultivar is characterized by its uniform tuber shape and size, smooth skin, and excellent processing qualities. Notably, RR has demonstrated the ability to maintain high yields even under varying N-rates, indicating its adaptability to diverse growing conditions [43]. The AL cultivar was released in 2002 by the USDA Agricultural Research Service and the Agricultural

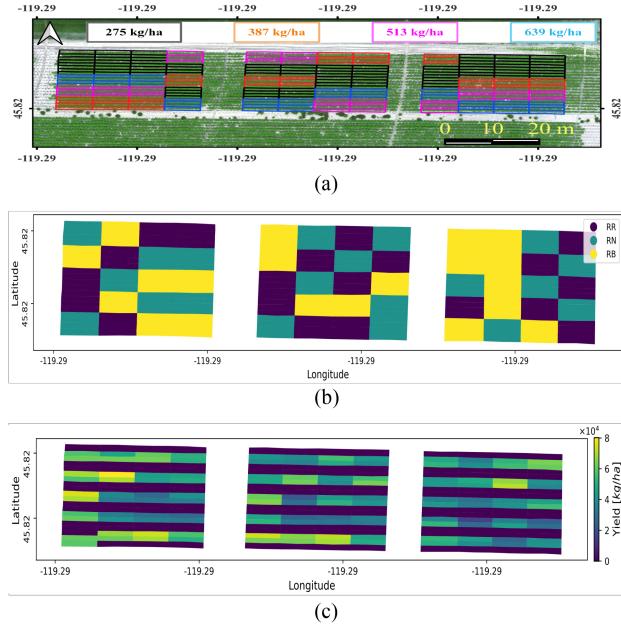


Fig. 2. Split-plot design of the year 2020. (a) Subplots shape file overlayed on the experimental field, and color coded to represent N-rate. (b) Primary plot that belong to potato cultivars (RR: Rainier Russet, RN: Russet Norkotah, and RB: Russet Burbank) in the experimental field design. (c) Yield plot of two inner rows that are reserved for cultivars under analysis. The subplot with no yield values represents buffer rows.

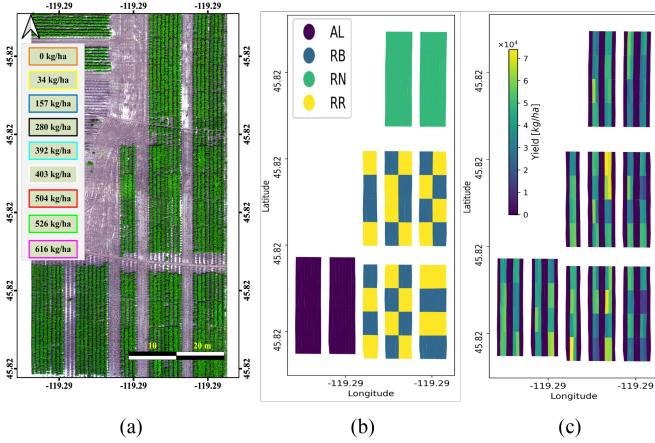


Fig. 3. Split-plot design of the year 2021. (a) Subplots shape file overlayed on the experimental field, and color coded to represent N-rate. (b) Primary plot that belong to potato cultivars (RR: Rainier Russet, RN: Russet Norkotah, RB: Russet Burbank, and AL: Alturas) in the experimental field design. (c) Yield plot of two inner rows that are reserved for cultivars under analysis. The subplot with no yield values represents buffer columns.

Experiment Stations of Idaho, Oregon, and Washington. It is known for its high yield potential and good processing qualities, especially for dehydration [44]. The Certified seed tubers of these cultivars were hand-cut to approximately 43–57 g per piece, treated with fludioxonil (Maxim4FS), and suberized for 2 weeks at around 95% relative humidity.

The planting of these seed tubers took place in late April of 2020 and 2021, on the same date, utilizing an assisted-feed system with specific row and column spacing. In experimental setup, each plot comprises two outer rows designated as buffer

rows, as shown in Figs. 2(c) and 3(c), while the two inner rows are reserved for the cultivars under investigation. Notably, all the tested cultivars exhibited moderate vine size, preventing overshadowing during the crop season. This arrangement allows us to focus our analysis solely on the inner rows to predict yield accurately. By concentrating on the inner rows, we aim to minimize any potential influence of the border rows on N mixing related tuber yield outcomes, which ensures the validity of our findings. The wide range of N-rates were applied to the crop ranging from 0 to 639 kg/ha over the span of two years. Treatment 0 denoted the absence of N-rates, indicating a control group with no fertilizers application for 2021 experimental field design [refer Fig. 3(a)]. The cultivars were chosen because of their relevance to local agricultural practices and market demand. Standard agricultural practices, including weed, insect, and disease management, were consistently applied throughout the field experiments in accordance with local growers protocols to ensure the integrity and reliability of the experimental data.

The soil composition at HAREC-OSU consists primarily of Adkins fine sandy loam soil, which exhibited moderate acidity levels throughout the 2020 and 2021 period. To gather environmental data, monthly mean air temperature and precipitation measurements were obtained from the AgriMet (<https://www.usbr.gov/pn/agrimet/graphs.html>) Columbia-Pacific Northwest Region automated weather station situated at HAREC-OSU. Specifically, in 2021, the mean daily air temperature during the growing season, spanning from May to August, was slightly elevated at 22.51 °C compared to the corresponding period in 2020, which recorded a slightly lower temperature of 20.80 °C. In addition, 2020 experienced slightly higher ≈ 1/4th (i.e., 26.57%) of the total rainfall observed in 2021, indicating significant variations in precipitation patterns between the two years. Irrigation practices were adjusted based on the evapotranspiration of potatoes crop, with irrigation beginning in early June and continuing until harvest. From early June to mid-August, irrigation was applied daily, and from mid-August to harvest, it was applied three times per week, reflecting changes in response to varying environmental conditions during the 2020 and 2021 growing seasons. These environmental parameters provide valuable insights into the climatic conditions and water management practices affecting crop growth and development at HAREC-OSU.

III. DATASETS AND IMAGE PREPROCESSING

A. UAS Multispectral Imaging Datasets

UAS multispectral data were obtained using a 5-band MicaSense RedEdge MX+ sensor (MicaSense, Inc., Seattle, WA, USA) mounted on a DJI Matrice 600 Pro hexacopter (SZ DJI Technology Company, Ltd, Shenzhen, China). The multispectral data acquisitions were conducted under calm wind and clear weather conditions, typically within a 3-h (11:00 A.M. to 2:00 P.M.) window of solar noon, on a weekly basis throughout the growing seasons of 2020 and 2021, except in cases where technical issues or unfavored weather conditions hindered data collection. The multispectral data acquisition coinciding with

TABLE I
MULTISPECTRAL IMAGE ACQUISITION DATE

Year	Date of acquisition							
	28/5	03/6	10/6	15/6	23/6	16/7	21/7	29/7
2020	26/5	01/6	09/6	14/6	23/6	14/7	23/7	30/7
Growth stages	Emergence (T1)	Hilling (T2)		Tuberization (T3)	Bulking (T4)		Maturity (T5)	

different vegetative growth stages of potato crop for both the years are listed in Table I.

During data collection, the UAS maintained an 80% of front and side overlap, with an above-ground flight altitude of 30 m and a cruise speed of 5 m/s. This configuration resulted in the spatial resolution or ground sample distance of approximately 1.8 cm/pixel. Flight mission planning was executed using the MicaSense Atlas Flight iOS application in both years. Flight parameters and sensors setting remained consistent across both years for data collection. The camera settings including, “padding” was set to “ON” and “skypoint” was set to “OFF.” Before and after each flight, images of a MicaSense Calibrated Reflectance Panel (CRP; Model RP04) were captured for calibration purposes. More comprehensive protocol detailing of the data collection process can be found at <https://www.protocols.io/view/usda-ars-potato-genetics-lab-drone-data-collection-bp2l6148dvqe/v1>.

B. Preprocessing

1) *Image Stitching and Calibration:* In the image preprocessing workflow, the individual flight paths were captured by the multispectral imaging system and stitched together to create a single, geo-rectified, seamless composite image known as an orthomosaic. This was achieved using a customized “MicaSense multispectral” module in Pix4Dmapper 4.8.4 software (Pix4D SA, Lausanne, Switzerland). This process involved aligning and blending individual image tiles to create a single and continuous image of each band, which is photogrammetrically ortho-rectified to cover the entire field of interest at certain time of the day. It attempted to eliminate any seams or artifacts that may occur due to the image overlapping. The sensors characteristics and potential distortions are corrected during the calibration process. It involved removing lens distortion, radiometric calibration, and accounting for variations in sensor sensitivity across different spectral bands. In the module, the digital number values of the multispectral images were converted to reflectance values using the calibration reflectance panel (CRP) images taken before and after the flight mission. Accurate calibration is crucial for reliable and quantifiable analysis of the imagery.

2) *Shape-File Creation:* The shape file, created using quantum geographic information system (QGIS) software (<https://www.qgis.org/>) for both 2020 and 2021, delineated the field layout by assigning attributes to each subplot within the plot. These attributes included subplot locations, N-rates, specific potato cultivar indicators for border row subplots, and replication numbers. The unique identifier for each experimental plot, referred as “Plot IDs,” were formulated from a combination of attributes, providing a specific reference to individual

plots within the experiment. These identifiers were coupled with geospatial information, defining the polygonal shape of each plot. The comprehensive shape file information facilitates both spectral and spatial analyses by enabling the isolation of specific geographic locations (i.e., buffer rows/subplots) within the experimental design.

C. Yield Datasets

In early October, tubers from the inner rows under investigation were mechanically harvested using a single-row harvester. In total, there were 264 inner row subplots, with 120 for the year 2020 and 144 for the year 2021. Upon harvest, potatoes were promptly graded to assess both tuber yield and quality. Specifically, tubers from each plot were weighed and classified into two size categories based on the minimum U.S. grade and carton standards: those weighing less than 113 grams (classified as culls) and those weighing 113 grams or more. Any tubers displaying external defects, such as rot, knobs, growth cracks, misshapen, or irregular skin appearance, were downgraded to the cull category based on the severity of the defect as outlined in the USDA grading standards. For more information, refer to [https://www.ams.usda.gov/sites/default/files/media/Potato_Standard\[1\].pdf](https://www.ams.usda.gov/sites/default/files/media/Potato_Standard[1].pdf). The weight of tubers falling into both the categories were recorded. Tubers weighing more than 113 grams were classified as U.S. No. 1 grade and deemed marketable as per USDA standards (<https://www.ams.usda.gov/sites/default/files/media/USConsumerStandardsPotatoes.pdf>), and is finally used as total yield by converting it to kg/ha.

IV. METHODOLOGY

In the context of our research goal, we proposed two-step image postprocessing framework that is designed as a paradigm of computer vision within precision agriculture, aiming to extract precise and plausible vegetation information. The framework is developed to establish a robust pipeline for automatic segmentation from preprocessed UAS multispectral imaging maps. Particularly, linear unmixing (LUM) model is considered as an initial step. The LUM technique separates the image into two key components: endmember and abundance matrix [23]. The endmember matrix captures the spectral characteristics of target elements within the agricultural landscape, while the abundance matrix represents the spatial distribution of these elements. Furthermore, Otsu-based nonbinary adaptive autosegmentation technique is proposed for masking soil coordinates. While Otsu methods are effective in controlled or simpler scenarios, their accuracy can be low when applied to raw multispectral data collected over an entire crop cycle with diverse N-rates and across large areas. This is due to the significant impact of diverse N-rates on crops, resulting in higher variations in band reflectivity that impact segmentation accuracy. LUM addresses this limitation by decomposing the spectral signatures of each pixel into fractional contributions from given endmembers. This enhances both spatial and spectral features of the end members, facilitating autosegmentation during challenging growth stages [45]. Therefore, in our case, LUM is considered essential

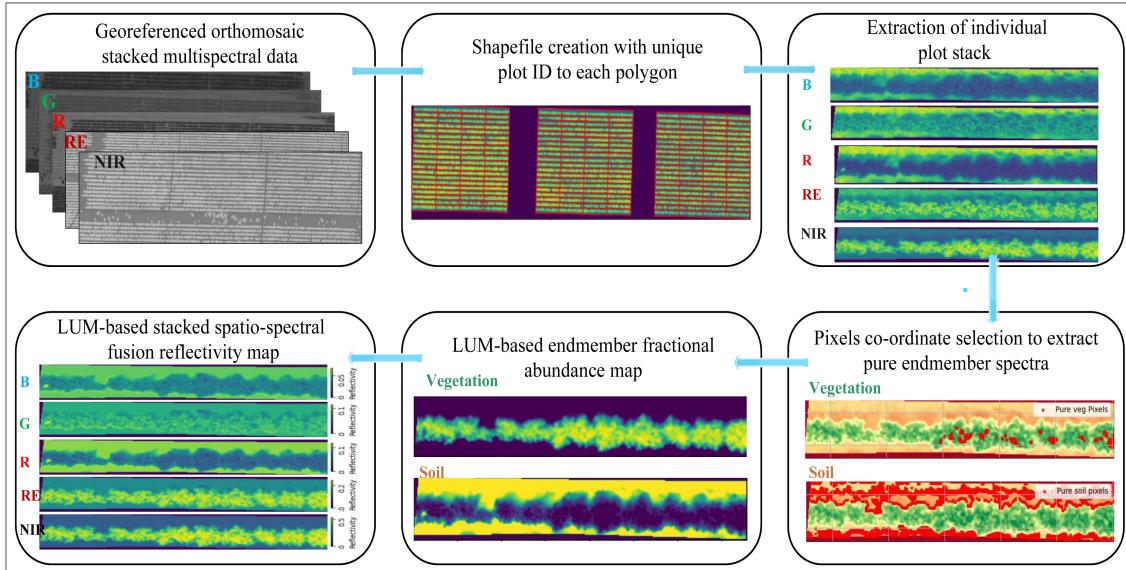


Fig. 4. Postprocessing step-1: Generating spatio-spectral fusion map from georeferenced seamless orthomosaic raw multispectral data. The process involved clipping each subplot from the primary plot and finding pure pixel endmembers for each subplot which facilitates the creation on endmembers abundance map utilizing linear unmixing(LUM) model. The utilization of abundance map and endmembers were further used to create spatio-spectral fusion map.

in overcoming the limitations of threshold-based segmentation on raw spectral data. The masked LUM fusion map is then utilized to compute precise spatial and spectral features. Moreover, the robust features for predicting yield were identified by effectively extracting latent variables from all spatial and spectral features using the partial least squares regression (PLSR) algorithm. Finally, the proposed CAR deep learning model uses temporal responses of the robust features systematically by increasing the vegetative growth stages from the emergence (T1) to hillling (T2), and tuberization (T3), bulking (T4), and maturity (T5) stage. As vegetation spatial and spectral indices can capture environmental influences and nutrient levels, this approach ensures a more comprehensive understanding of the underlying temporal vegetation dynamics influencing crop yield. The details of the image processing framework and deep learning operation are discussed in the following sections.

A. Linear Unmixing (LUM) Model

LUM assumes that the observed spectrum is a linear combination of the individual endmember spectra captured within the multispectral image, each weighted by its abundance fraction. The observed spectrum can be denoted as $\mathbf{Y} \in \mathbb{R}^{L \times N}$, where L and N represent the number of spectral bands and pixels, respectively. The endmembers spectra are denoted by $\mathbf{E} \in \mathbb{R}^{L \times M}$, with M endmembers, such that $M < L$ is satisfied. The goal is to estimate the matrix of fractional abundance (i.e., $\mathbf{A} \in \mathbb{R}^{M \times N}$). Thus, the spectral unmixing can be mathematically related using the following equation:

$$\mathbf{Y} = \mathbf{EA} + \mathbf{n} \quad (1)$$

where \mathbf{n} represents the residual matrix containing noise. The process of extracting pure endmembers from a multispectral image is carried out through a threshold-based pure pixel selection

approach. The endmembers matrix are derived by averaging the values of all valid pixels. Once the endmembers are identified, the spectral unmixing is treated as an optimization problem, where the least-squares method is employed to estimate the fractional abundance matrix (i.e., \mathbf{A}). Furthermore, abundance non-negativity constraint (ANC) and abundance sum-to-one constraints (ASC), as shown in (2) and (3), are applied to the matrix \mathbf{A} , indicating the spatial distributions of the end-members over desired image, ensuring the physically valid matrix

$$\text{ANC: } \sum_{n=1}^M A_n = 1 \quad (2)$$

$$\text{ASC: } A_n \geq 0; n = 1, \dots, M. \quad (3)$$

1) Spatial–Spectral Fusion Map: The complete process involved in generating spatio-spectral fusion map in this study for extracting time-series vegetation features is illustrated in Fig. 4. The spatio-spectral fusion map is widely known as composite image or fused image, which is created by combining both the spectral domain (i.e., multispectral or hyperspectral data) and the spatial domain (i.e., panchromatic and high-resolution data). Spatio-spectral fusion offers a distinct advantage by combining the detailed spatial information from high-resolution imagery with the spectral content of multispectral/hyperspectral data. Once the fractional abundance map was generated using LUM, the spatio-spectral fusion map (R_L) was created by multiplying the pure endmembers spectra with their respective fractional abundance map, as expressed in the following equation:

$$R_L = \sum_{n=1}^M A_n E_n. \quad (4)$$

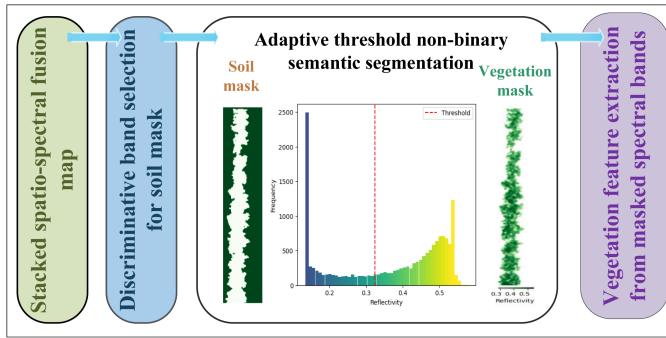


Fig. 5. Postprocessing step-2: Adaptive threshold non-binary autosegmentation using Otsu method. The process involved identify endmembers from most discriminating to identify soil pixel coordinates. The identified soil pixel coordinates were masked from all the bands to create masked multispectral map for spatio-spectral features extraction utilizing band math.

Thereafter, (4) produced the enhanced reflectivity maps corresponding to each spectral bands. It is worth emphasizing that the accuracy and quality of the results are closely tied to the precision of endmembers selection. This method is especially useful for isolating heterogeneous backgrounds or segmenting specific endmembers, particularly in cases where simple binary thresholding fails to classify all pixels accurately.

B. Adaptive Threshold Nonbinary Autosegmentation via Isolating Pixel Coordinates

Following the acquisition of the spatio-spectral fusion maps, the next step involved nonbinary adaptive autosegmentation, which classifies image into different region (i.e., foreground and background) based on auto-optimized threshold value from the discriminating band, as illustrated in Fig. 5. Several binary and nonbinary methods such as Otsu threshold [46], entropy histogram [47], RGB color-based indexing, and Gaussian based threshold [48] were used by researchers to detect homogeneity threshold based on index values. However, these methods for threshold selection for vegetation segmentation have several limitations when applied to raw multispectral data collected over an entire crop cycle with diverse N-rate and across large areas. One significant limitation is the difficulty in selecting appropriate time points and thresholds during the crop growth cycle to accurately map cropping patterns, which can impact the overall accuracy of the segmentation [49]. Some of them are often relied on index-based manual or fixed threshold values, making them sensitive to variations in other index and challenging lighting conditions collected over large area. Some methods assume uniform distribution of vegetation, which may not hold in different crop growth stages. RGB color-based indexing may not offer sufficient discriminating power, and methods based on Gaussian distributions may misrepresent the true color/value distribution. The threshold value over large area changes significantly based on biophysical and biochemical properties of the soil and vegetation. The estimation of phenological threshold value based on time-series data require careful consideration and calibration to ensure the reliability of segmentation results.

In light of the constraints posed by conventional threshold-based vegetation segmentation techniques, our proposed method

introduces an enhanced strategy. It uses Otsu-based adaptive autothreshold detection, which is complemented by LUM maps, and applied over time and subplot scales for more precise segmentation. The process utilizes most discriminating band within the spatio-spectral fusion map produced by LUM that can effectively identify soil pixel coordinates across large areas without compromising the accuracy of vegetation pixels. In addition, the identification of soil pixel coordinates facilitates the isolation of soil pixels across all bands by matching soil pixel coordinates. The LUM technique dynamically classifies the complex fluctuations in vegetation growth stages, specifically when it is difficult to classify the soil and vegetation pixel automatically. This approach offers improved robustness and accuracy in segmentation outcomes.

Let R_L represent the spatio-spectral fusion map with L number of band. The mathematical formulations of the process are expressed in (5)–(7).

1) *Band Identification:* Identify the discriminating band (R_L) by analyzing differential spectral characteristics of the available endmembers at given vegetative stage (t)

$$R_L = \arg \left\{ \max_{R_L} \{ \text{DiscriminativeMetric}(t, R_L) \} \right\} \quad (5)$$

where the $\text{DiscriminativeMetric}(t, R_L)$ quantifies the discriminating power of R_L band at growth stage t . This is crucial for accurate discrimination between different classes/endmembers. Instead of relying on a single fixed threshold value over whole plot, we implemented adaptive Otsu-based autotuning of the threshold T_t at particular growth stage t based on discriminating band R_L at subplot scale within the image.

2) *Soil Pixel Identification:* Identify the coordinates of soil pixels, S_t , within the selected band, R_L , using the Otsu-based autothreshold detection. This is crucial step in distinguishing areas of interest and enabling precise segmentation and can be expressed mathematically using the following equation:

$$S_t = \text{Otsu} \{ (x, y) \mid R_L(x, y, t) > T_t \} \quad (6)$$

where (x, y) represents pixel coordinates. Here, T_t represents adaptive threshold. This entails selecting a T_t , that is tailored to the specific vegetative stage being observed over all subplots. The optimum threshold process autoadapts to changes in vegetation appearance over time.

3) *Masking Soil Coordinates:* Apply the identified soil coordinates, S_t , as a mask to all bands available within band R_L to create a masked multispectral fusion image R'_L . The process is executed utilizing the following equation:

$$R'_L(x, y, t) = \begin{cases} R_L(x, y, t), & \text{if } (x, y) \notin S_t \\ 0, & \text{if } (x, y) \in S_t. \end{cases} \quad (7)$$

This can effectively and precisely remove the soil coordinates from all the subplots available within multispectral image, ensuring that the subsequent analysis focuses exclusively on the vegetation physiological and biological growth characteristics.

C. Feature Extraction Using Masked Multispectral Image

The LUM-based soil-masked multispectral data were then used for spatial and spectral feature extraction. A total of 240 in

TABLE II
SPECTRAL AND SPATIAL FEATURES EXTRACTED FROM THE MASKED
MULTISPECTRAL BANDS

Feature	Formula	Reference
NDVI	$\frac{B5-B3}{B5+B3}$	[50]
GNDVI	$\frac{B5-B2}{B5+B2}$	[51]
DVI	$B5 - B3$	[52]
EVI	$\frac{2.5(B5-B3)}{(B5+6B3-7.5B1+1)}$	[53]
EVI2	$\frac{2.5(B5-B3)}{(B5+2.4B3+1)}$	[53]
MARI	$(\frac{1}{B2} - \frac{1}{B4}) \times B5$	[54]
CHLRE	$\frac{B5}{B4} - 1$	[55]
CHLGR	$\frac{B5}{B2} - 1$	[55]
SAVI	$\frac{1.5(B5-B3)}{B5+B3+0.5}$	[56]
SAVI2	$\frac{B5}{B3+(b/a)}; a = 0.01, b = 1.43$	[56]
MSAVI	$\frac{2B5+1-\sqrt{4B5(B5-B3)+1}}{2}$	[57]
OSAVI	$(1+l)\frac{B5-B2}{B5+B2+l}; l = 0.16$	[58]
TSAVI	$\frac{a(B5-aB3-b)}{aB5+B3-ab+X(1+a^2)}$	[59]
MSR	$\frac{B5}{B3} - 1$	[60]
MCARI	$(B4 - B3 - 0.2(B4 - B2)) \frac{B4}{B3}$	[61]
MCARI OSAVI		[61]
PVI	$\frac{B5-aB3-b}{\sqrt{1+a^2}}$	[62]
SR	$\frac{B5}{B3}$	[60]
WDRI	$\frac{0.3B5-B3}{0.3B5+B3}$	[63]
V_f	$\frac{V_p}{T_p}$	

Bands: B1=Blue, B2=Green, B3=Red, B4=Red-Edge, and B5=Near-Infrared.
Spectral Features: NDVI = Normalized difference vegetation index, GNDVI = Green normalized difference vegetation index, DVI = Difference vegetation index, EVI = Enhanced vegetation index, EVI2 = Second-order enhanced vegetation index, MARI = Modified anthocyanin reflectance index, CHLRE = Chlorophyll red-edge index, CHLGR = Chlorophyll green index, SAVI = Soil-adjusted vegetation index, SAVI2 = Second-order soil-adjusted vegetation index, MSAVI = Modified soil-adjusted vegetation index, OSAVI = Optimized soil-adjusted vegetation index, TSAVI = Transformed soil-adjusted vegetation index, MSR = Modified simple ratio, MCARI = Modified chlorophyll absorption ratio index, PVI = Photochemical reflectance index, SR = Simple ratio, and WDRI = Wide dynamic range index.

Spatial Feature: V_f = Vegetation fraction, V_p = Total vegetation pixel, and T_p = Total pixel.

2020 and 288 subplots in 2021 were clipped from the primary plot design, including both buffer and inner rows under investigation. The plot IDs were used to exclude the buffer subplots, resulting in a total of 264 subplots (120 from 2020 and 144 from 2021). These 264 subplots were used for spatial and spectral feature extraction at each date of data acquisition, as mentioned in Table I.

The computation of diverse vegetation indices map (spectral features), for instance, ratio, difference, hybrid, normalized, modified vegetation indices were included in this study, which yield critical insights into vegetation health and vigor. In addition, we employed a spatial analysis technique for the estimation of vegetation fraction, a key parameter in understanding the distribution of its abundance at each subplot. A total of 20 features (19-spectral and 1-spatial) were extracted, as detailed in Table II, providing a comprehensive analysis of vegetation's physiological and biological growth characteristics. The average

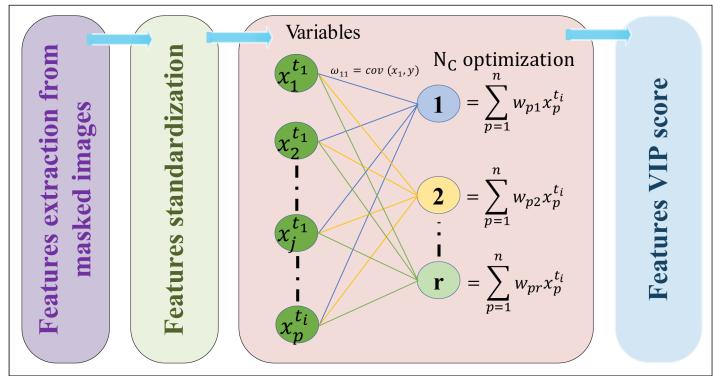


Fig. 6. PLSR network architecture for vegetation feature engineering. The spatio-spectral features extracted from the soil masked vegetation map were fed to the PLSR algorithm to identify robust features relevance to yield.

value of vegetation pixels within their respective subplots was used to compute the final spectral and spatial feature values. The importance of both spatial and spectral features, which are significant in predicting yield, were further analyzed using the partial least squares regression (PLSR) algorithm.

D. PLSR Model for Feature Engineering

With the large number of features/predictors, feature engineering is crucial to ensure deep learning model receives robust and high-quality inputs, which leads to more reliable and accurate predictions. Thus, PLSR algorithm were implemented to robustly analyze predictor variables comprising temporal responses of the spatial and spectral feature to final yield. PLSR is a multivariate statistical technique used for feature engineering (i.e., dimensionality reduction and the identification of important features) by building relationship between a set of extracted feature map (i.e., independent/predictor variables matrix (X)) and a yield map (i.e., dependent/response variable vector (Y)), while dealing with high collinearity data [64]. The steps involved in the PLSR algorithm to find robust predictors are shown in Fig. 6.

The PLSR process begins with features and yield map data preprocessing. Both input (i.e., spectral and spatial) predictor variables, X , and the response (i.e., yield) variable, Y , are standardized to have a mean of zero and a variance of one. Standardization ensures that all variables are on the same scale, facilitating the convergence of PLSR algorithm. The input data tensor is considered as time-series sequence of predictor variables, denoted as $X = [x_1, x_2, \dots, x_n]$ and output data tensor denoted as $Y = [y_1, y_2, \dots, y_n]$. Each vector has dimensions $[(n \times p)_t]$, where, n represents the number of samples, p is the number of predictor variables (i.e., input features), and t represents the number of time-step to be concatenated from T1 to T5, and $[]$ denote the concatenation operation.

1) *Hyperparameter Tuning*: The core of the PLSR algorithm lies the optimization of the number of principal components (N_c), which captures covariance ($Cov(x, y)$), weight (w), and variable importance in projection (VIP) among all the predictors. VIP scores indicate the relative importance of predictor

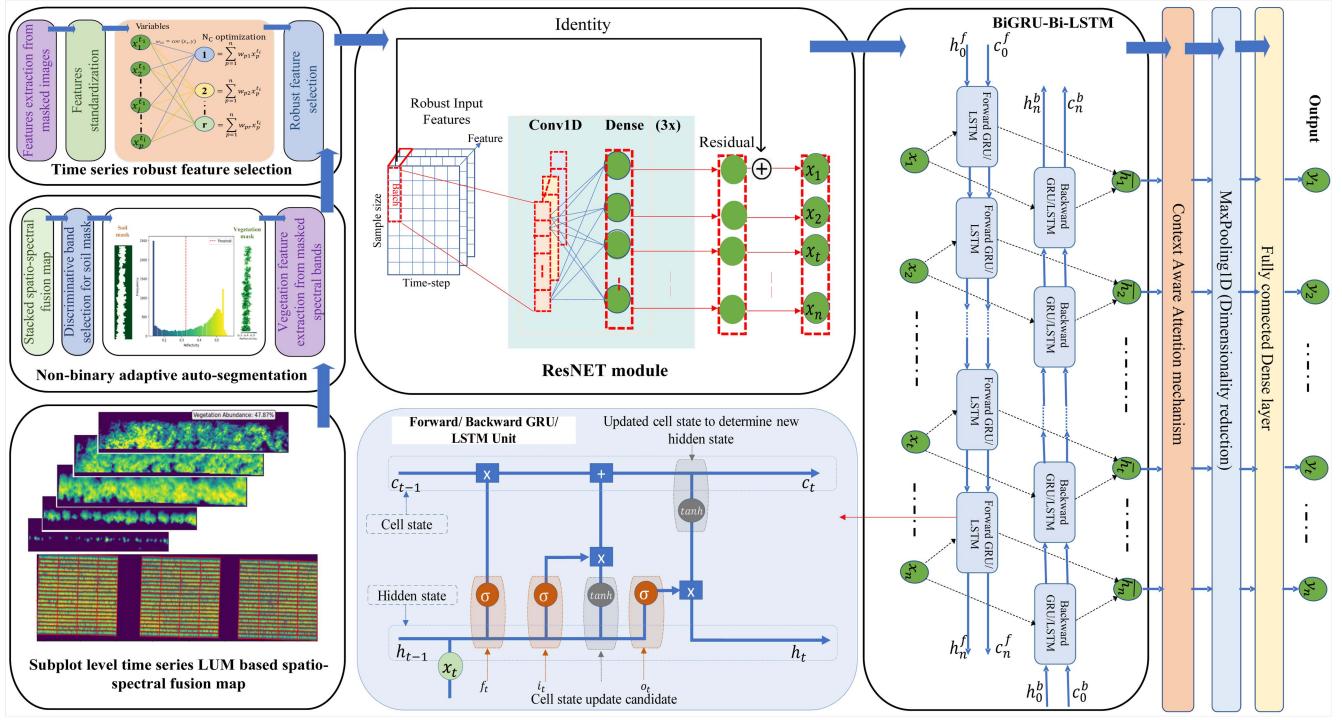


Fig. 7. Illustration of the overall architecture of the CAR Conv1D-BiGRU-BiLSTM-Net. The method starts with extraction of subplot level features from vegetation segmentation applied on LUM-based enhanced map. The model were implemented to take 3-dimensional input data across total number of sample (i.e., available subplot), multisource feature including spatial, agronomy, and robust spectral derived from PLSR algorithm, and their temporal variation to predict final yield.

variables. A random search was conducted to optimize the PLSR model hyperparameter, N_c by selecting a range from 1 to 5. The process included a maximum of 30 iteration, enabled scaling, and maintained a low tolerance values of 10^{-3} for convergence. The 5-fold cross-validation was used to determine the optimal number of components (N_c) within the PLSR model. The RandomizedSearchCV in scikit-learn was used to facilitate minimization during optimization, which typically uses negative value of scoring metric (i.e., root mean squared error) in PLSR algorithm [65]. Once the N_c were optimized, the VIP score for each predictor variable were calculated. The robust features identified for yield prediction were then fed into the proposed CAR-based deep learning model for end-season crop yield prediction.

E. Context-Aware Attention Mechanism and Residual Connection Based One-Dimensional Convolution-Bidirectional Gated Recurrent Unit-Bidirectional Long Short-Term Memory-Network (CAR Conv1D-BiGRU-BiLSTM-Net) Model

The results of PLSR-based robust features selection were then integrated to form three-dimensional input datasets, represented as $(n \times t \times p)$, which were used to feed the proposed deep learning model. Where n represents samples size, totaling 264 samples; t is vegetative growth stage, varying from T1 to T5; and p is number of robust predictors, ranging from 3 to 7 features. The output data were of dimension $(n \times 1)$, i.e., final yield \times 1 (represents the end-season stage). A total of 80% data were

used for training, while the remaining 20% were used for testing the model. Input data normalization was applied for smooth training of the proposed novel CAR Conv1D-BiGRU-BiLSTM-Net model. The overall step involved in the implementation of proposed model are shown in Fig. 7 and discussed in the following sections.

1) Residual Network Module: The proposed ResNet architecture initially applies three sequential Conv1D layers followed by dense layers to learn hierarchical features at different levels of abstraction from the normalized time series input datasets at each crop growth stage. Each Conv1D applied a set of learnable filters to extract input patterns. Then, the dense layers perform nonlinear transformations on the output of Conv1D layer. Each neuron in the dense layer is connected to every neuron in the previous layer, allowing the network to learn fine-grained relationships between inputs and output [39]. In addition, a residual connection was incorporated to improve gradient flow and mitigate the vanishing gradient problem by allowing the model to bypass certain layers. The residual addition preserve the original input information and improves the network's ability to learn complex hierarchical features.

2) BiGRU-BiLSTM Network: After applying ResNet module, a BiGRU network was introduced to capture short-term dependencies within time series data. It is designed to handle short-term dependencies in both past and future time steps, processing the input sequence in two directions. The GRU unit uses two gates, the update gate and the reset gate to control how information is updated and retained over time. Following the BiGRU layer, a BiLSTM layer was employed to capture

long-term dependencies in both past and future time steps [38]. The BiLSTM also processes input sequences bidirectionally but maintains a cell state governed by three gates: forget gate, input gate, and output gate. These gates control the flow of information within the LSTM cell [37].

3) *Context-Aware Attention Mechanism*: The context-aware attention mechanism was applied to enhance the model's ability to focus on relevant features within the residual-connected Conv1D-BiGRU-BiLSTM-Net. Rather than treating each feature vector in the sequence independently, the mechanism generates a context vector that highlights key features across the entire sequence. Each feature is assigned a weight based on its importance to the context, allowing the model to adaptively focus on different regions of the input. The outputs of the context-aware attention layer, following the pooling mechanism and fully connected dense layer, are depicted in the following equation:

$$Y_{\text{pred}} = \text{Linear} \left(\sum_j W_j \cdot h_t + b \right) \quad (8)$$

where h_t represents the feature vector obtained from the residual-connected Conv1D-BiGRU-BiLSTM layers, W_j , and b represent the weight matrix and bias vector of the output layer, respectively.

4) *Hyperparameter Tuning*: The model hyperparameter covered a range of crucial parameters that impact model performance. These parameters included the number of units, filters, kernel size, batch size, learning rate, exponential decay rates for the first and second moment estimates (β_1 and β_2), and epsilon (ϵ). The selected range of hyperparameter within the model are gives as: Number of units for Conv1D and BiGRU/BiLSTM layers: [32, 64, 128], Number of filters: [8, 16, 32], Kernel sizes: [1,2,3], Activation: [ReLU: Rectified linear unit, GeLU: Gaussian error linear unit, Tanh, Sigmoid], context aware attention mechanisms: [Softmax, Sigmoid, Additive], Optimizers: [SGD: Stochastic gradient descent, RMSprop: Root Mean Squared Propagation, Adam: Adaptive Moment Estimation], Batch size: [40, 50, 60], Learning rate: [0.001, 0.01, 0.1], β_1 : [0.9, 0.95, 0.99], β_2 : [0.99, 0.999, 0.9999], and ϵ : [10^{-4} , 10^{-5} , 10^{-6}]. The random search was found computationally less expensive compared to other methods, such as grid search or Bayesian optimization [66]. Thus, the randomly sampled hyperparameters combination from predefined ranges were efficiently used to explore the hyperparameter space and find the optimal model configuration.

5) *Model Evaluation Metrics*: The model's quantitative performance was assessed by calculating the coefficient of determination (R^2) and the RMSE between the actual yield (Y) and the predicted yield (Y_{pred}), expressed in (9) and (10)

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - Y_{\text{pred},i})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (9)$$

$$\% \text{RMSE} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - Y_{\text{pred},i})^2}}{\bar{Y}} \times 100 \quad (10)$$

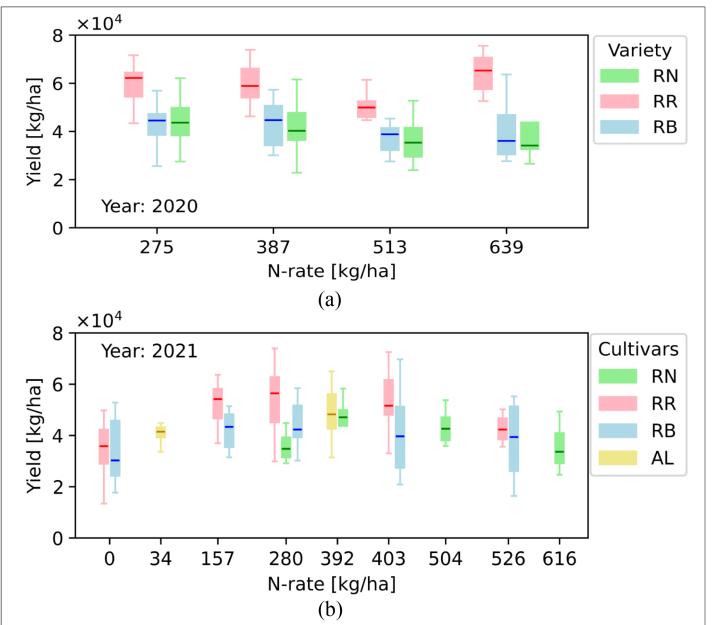


Fig. 8. Variation of in situ tuber yield among different potato cultivars (RR: Rainier Russet, RN: Russet Norkotah, RB: Russet Burbank, and AL: Alturas) across varied N-rate. (a) 2020. (b) 2021.

where n is the number of samples, Y_i is the actual yield for the i th sample, \bar{Y} is the mean of the actual yields, and $Y_{\text{pred},i}$ is the predicted yield for the i th sample.

V. RESULTS AND DISCUSSION

A. N-Rate Versus End Season Yield Analysis

The tuber yields were collected from two inner row subplots reserved for the cultivars under investigation. In total, there were 264 inner row subplots, with 120 subplots for the year 2020 and 144 subplots for the year 2021. The box plot variation of end-season tuber yield among different cultivars reserved under investigation across a range of applied N-rates for the years 2020 and 2021 are illustrated in Fig. 8(a) and (b). At specific N-rate, the average tuber yield for the RR (new variety) was found to be significantly higher compared to all other cultivars for both the years [43]. An increase in the N-rate generally resulted an increase in the average tuber yield of the RR cultivar, with the lower yield reported at 0, 513, and 526 kg/ha in 2020 and 2021, respectively [18]. The RB cultivars average tuber yield was found to be relatively stable for both the years, except at 0 kg/ha [18], [67]. The approximate variations of RN cultivar tuber yield were reported to be decreasing with N-rate, for year 2020. In 2021, the average RN tuber yield increased significantly, particularly from 280 to 392 kg/ha. However, it began to decrease at higher N-rates, starting from approximately 392 kg/ha onward [43], [68]. The AL cultivar (only in 2021) showed significant yield increases from 34 to 392 kg/ha, with variability at higher N-rates. No significant comments were made on AL cultivar because of the limited N-rate observations under investigation. The diverse cultivars, varied N-rates, and multiyear data introduces a broad range of spatial and spectral

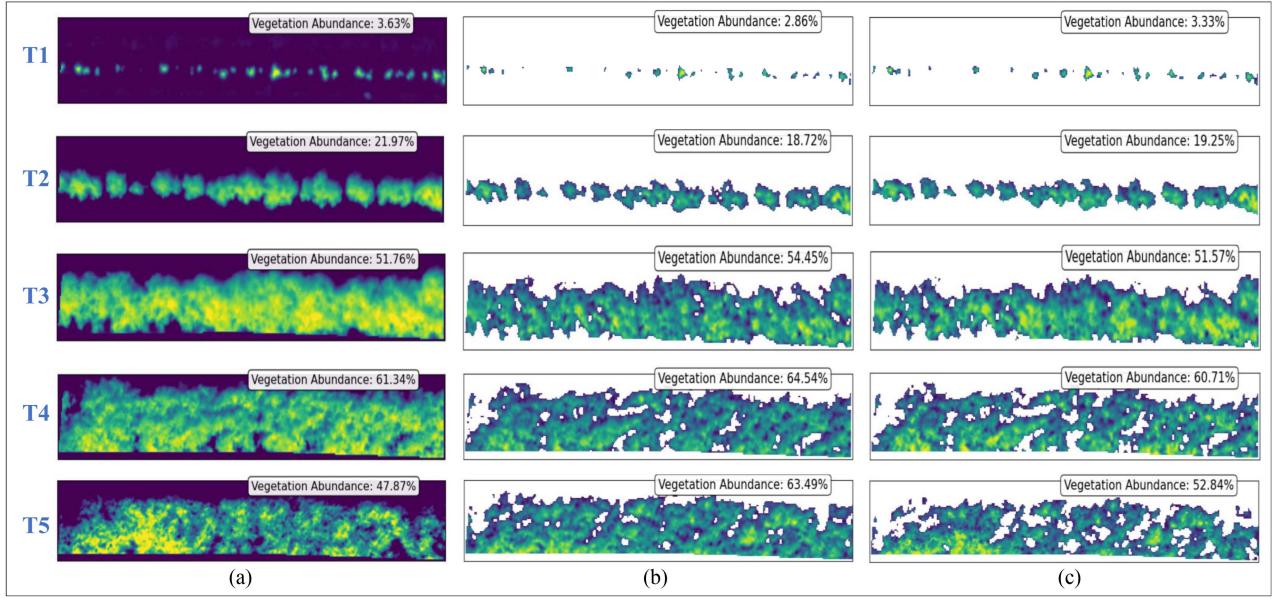


Fig. 9. Illustration of Otsu-based segmentation employed to a sample subplot throughout the growing season, varying from T1 to T5. The segmentation result were compared with the nonzero vegetation endmember abundance pixel obtained in LUM process from T1 to T5. In Column (a), nonzero vegetation endmember abundance pixel along with their percentage is depicted. Column (b) and (c) displayed vegetation abundance map from T1 to T5, when applied Otsu method on raw and LUM maps, respectively.

features that reflect plant health (e.g., chlorophyll levels, canopy structure, water stress), enabling a deep learning model to capture complex real-world variability. This approach enhances the model's capacity to generalize across different conditions and unbiased yield prediction by learning nonlinear relationships between features like N uptake, growth patterns, and environmental factors for more robust yield prediction.

B. Adaptive Autosegmentation Comparison on Raw and LUM Processed Data

In our study, the integration of Otsu-based autosegmentation method complemented by LUM has proven effective in making optimal threshold decisions. The raw multispectral data have diverse spectral variability caused by factors like, N-rate, soil type and moisture, growth stages, shadowing, and management practices. Applying an autosegmentation algorithm directly to raw data can produce inconsistent feature extraction outcomes. Moreover, using such algorithms on large areas may result in a single threshold that leads to inaccurate segmentation during critical growth stages, where soil pixels may be miss-classified as vegetation and vice versa. To bridge this gap for accurate spatial and spectral feature extraction, we performed a shape file clipping on a subplot level. The plot IDs were used to exclude the buffer subplots, resulting in a total of 120 subplots for 2020 and 144 subplots for 2021. The LUM technique was employed on each raw subplots to generate an enhanced reflectivity map, which offers accurate representation of spectral properties within each vegetation pixel. In addition, the spatial detail preserved by LUM enables high-resolution classification, allowing for the detection of subtle spectral differences between soil and vegetation at the pixel level, when compared to raw spectral differences at all crop growth stages. On capitalizing

LUM technique approach by minimizing artifacts that may be present in raw data, adaptive nonbinary Otsu-based autothreshold detection for soil mask were applied to most discriminating band (i.e., NIR band, in our case). The combination of Otsu-based autosegmentation with LUM-processed data offers a complementary advantage by enhancing the accuracy and reliability of soil segmentation throughout the crop growth cycle. Once the soil thresholds were identified, the soil pixel coordinates were determined from the discriminating georeferenced NIR band. These coordinates were then used to mask out the soil pixels from each of the remaining multispectral bands.

Fig. 9(a) represents nonzero vegetation endmember abundance obtained in LUM process as the growth progresses. At early stages, i.e., emergence (T1) and vegetative (T2), the vegetation is sparse, as indicated by the low abundance values, 3.63% and 21.97%, respectively. At Tuberization (T3), vegetation coverage becomes more prominent (51.76%), peaking during bulking (T4) with abundance of 61.34%, after slightly reduced at maturity (T5) with 47.87%. In the absence of ground truth for vegetation fraction, we compared the accuracy of the Otsu segmentation applied to both raw and LUM maps, as shown in Fig. 9(b) and (c), respectively, by comparing them with the nonzero vegetation endmember abundance obtained in the LUM process, as depicted in Fig. 9(a). Our result showed that vegetation abundance percentages matched better across all growth stages using LUM maps, when compared to raw maps.

In addition, the comparison between direct application of Otsu-based autosegmentation for threshold detection on raw and LUM data across the entire growth stages are also shown using histogram variation of NIR band pixel values in Figs. 10 and 11, respectively. As shown in Fig. 10, during T1 and T2 growth stages, the threshold range for all the subplots were found to be narrow and could be identified with similar

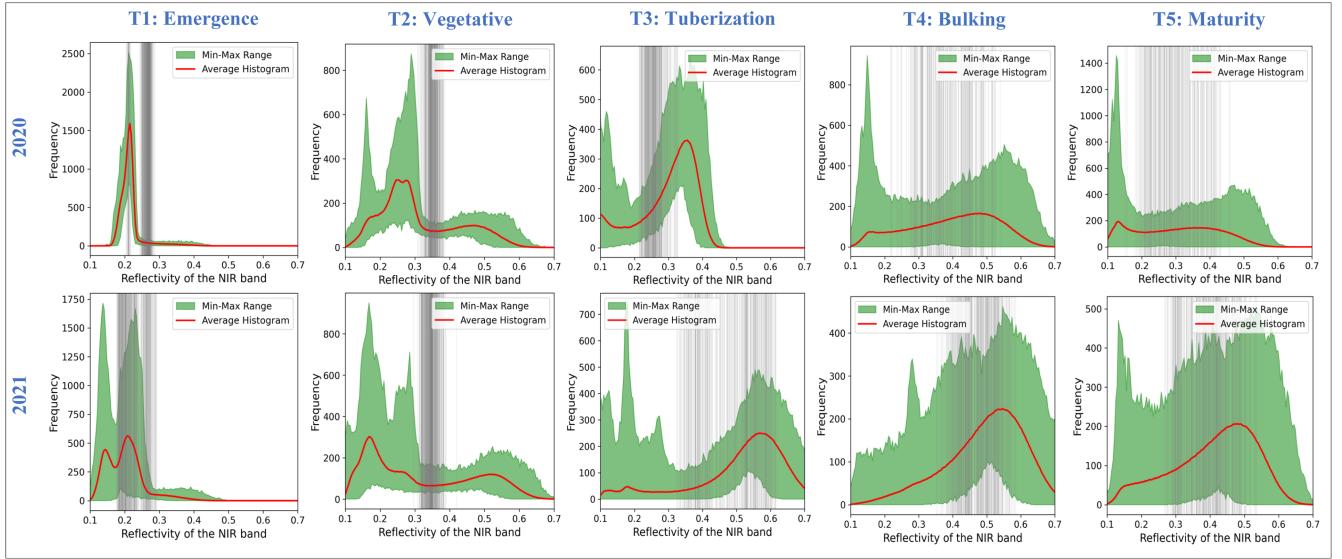


Fig. 10. Histogram plot showing the variation of average pixel value within NIR band of raw subplots for the years 2020 (top row) and 2021 (bottom row). The vertical gray line marks the threshold value determined by Otsu's adaptive semantic segmentation applied to the raw subplots.

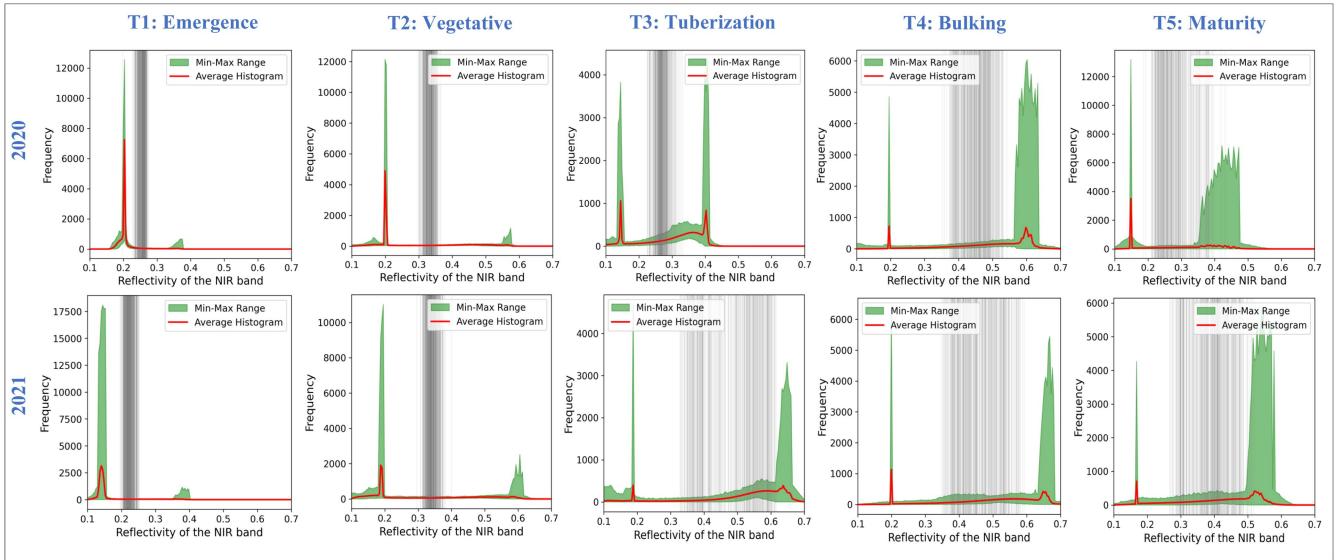


Fig. 11. Histogram plot showing the variation of average pixel value within NIR band of a linear unmixing (LUM) processed subplots for the years 2020 (top row) and 2021 (bottom row). The vertical gray line marks the threshold value determined by Otsu's adaptive semantic segmentation applied to the LUM processed subplots.

precision using Otsu segmentation on both raw and LUM-processed data. This is because the Otsu method determines the optimal threshold by maximizing the variance between classes (e.g., soil and vegetation), effectively separating pixels based on the frequency distribution of their intensities [46]. During these stages, the NIR band effectively distinguishes frequencies of soil and vegetation pixel intensities, enabling precise segmentation. However, during and after the T3 stage, the frequency of soil and vegetation pixel intensities in the NIR band histogram becomes more gradual and less distinct, as shown in Fig. 10, leading to less accurate segmentation, when applied to raw data. The gradual transitions, along with variability in spectral signatures due to N-rate variations, cause overlap between vegetation and soil pixel intensities in the NIR band. This led to a broader

threshold range for both years. The tuberization and bulking stages were found to be critical stages and have a high impact on yield estimation, as noted by several researchers [69], [70], [71]. The inconsistent segmentation outcomes at crucial stages may significantly affect the accurate extraction of spatial and spectral features and hence the yield prediction model.

To address this issue, we applied LUM techniques to the raw data, which effectively separates the spectral intensity of endmembers within the NIR band histogram by decomposing each pixel's spectral signature into the fractional abundances of soil and vegetation endmembers. The decomposition led to better spectral variability by accounting effect of varying N-rates, which facilitated the separation of vegetation and non-vegetation pixel frequency, as shown in Fig. 11. The improved

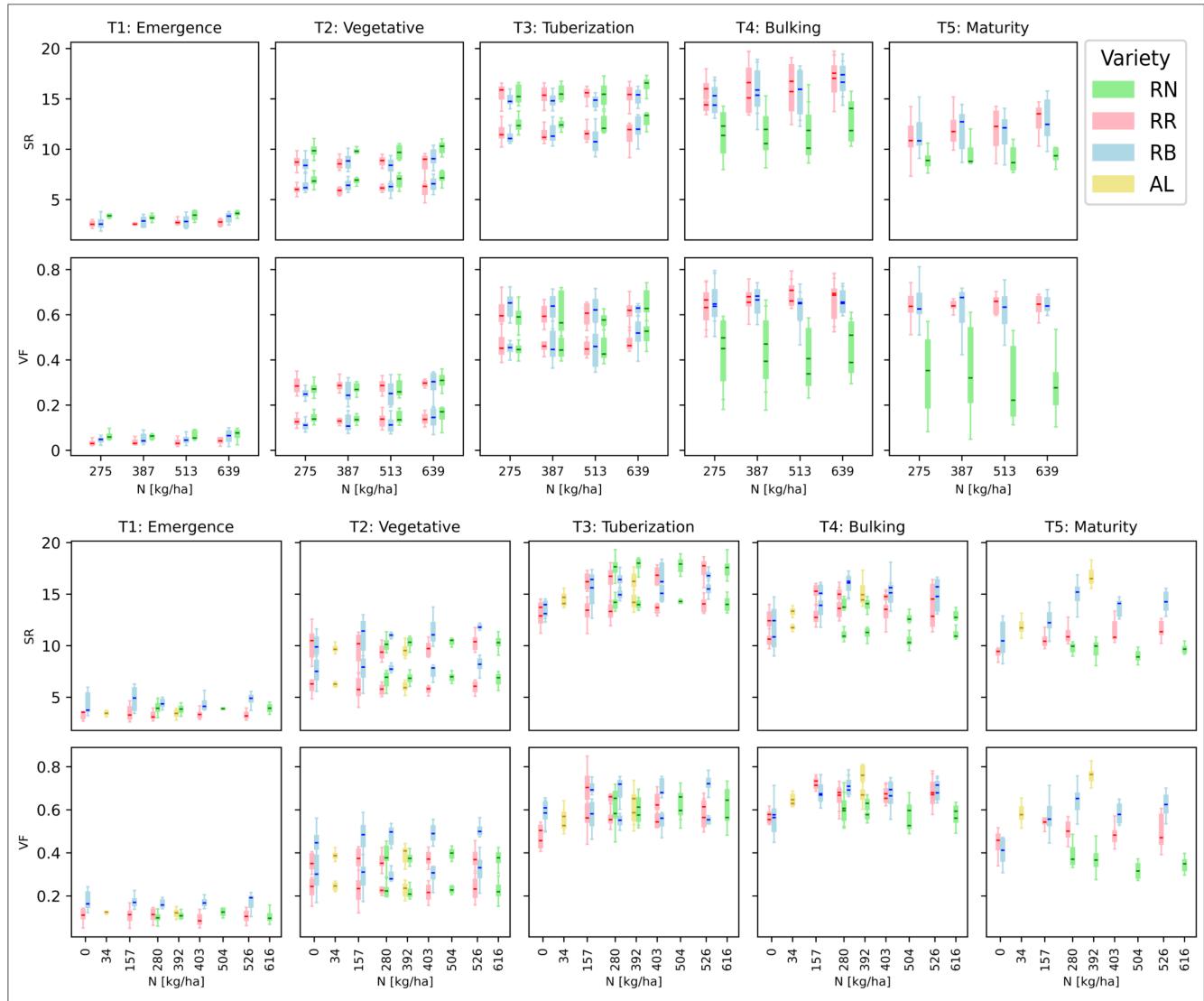


Fig. 12. Temporal variation of the spatial and spectral features extracted from the LUM processed raw orthomosaic data. These features, including 19-spectral and 1-spatial feature, were extracted for each cultivar and N rate under investigation. The box plot showing the variation of a single spectral (i.e., simple ratio (SR)) and a single spatial (i.e., vegetation fraction (VF)) feature for the years 2020 (top) and 2021 (bottom) to understand their temporal dynamics. The box plot variations for all features will be available in the supplementary datasets. RR, RN, RB, and AL represents Rainier Russet, Russet Norkotah, Russet Burbank, and Alturas potato cultivars type.

distinct spectral profiles (with nongradual histogram) of both endmembers, even during challenging growth stages allowed easy identification of threshold using Otsu method within discriminating bands. In addition, it preserved a broader threshold range, reflecting the impact of varying N-rates on spectral variability of soil and vegetation pixel values in the NIR band. Thus, by emphasizing the intrinsic spatial and spectral properties of each component, LUM-based segmentation provided more consistent and reliable thresholds for soil pixel masking and the extraction of spatial and spectral features.

C. Temporal Characteristics of Extracted Vegetation Features and Its Analysis

The temporal variations of all extracted features across different N-rates and cultivar types for both years were plotted and

included within the supplementary data in the Supplementary Material. To illustrate the temporal trends, we selected a single spectral (i.e., simple ratio (SR)) and a single spatial (i.e., vegetation fraction (VF)) feature variation for analysis, as depicted in Fig. 12, for both the years 2020 and 2021, respectively. The temporal variation of features, influenced by varying N-rates and different cultivars, presented in this study is divided into two stages: (1) pretuber stage and (2) post-tuber stage. A temporal analysis showed that the values of spectral and spatial feature increased until the late tuber stage, then decreased afterwards. In addition, an approximate increase in both spatial and spectral features was observed with the increase in applied N-rates. This increase is likely due to the greater leaf area that develops with higher N-rates, which enhances chlorophyll absorption and water uptake in the red-edge wavelengths, as well as increases the multiple scattering of photons in the NIR bands. However,

during and after the bulking stage of potato tubers, a reduction in leaf area and vegetation water content led to a corresponding decrease in these spatial and spectral features.

Moreover, the analysis was conducted at the cultivar level. Interestingly, a notable difference in vegetation spectral features were observed among all cultivars, despite receiving identical N-rates. The RN cultivar showed the highest variation in average spectral value until the late tuber stage, while the RR cultivar (i.e., new cultivar) had the lowest spectral value during this period when compared to the other cultivars. However, after the tuberization growth stage, the RN cultivar experienced the most notable decrease in average spectral values, while the RR cultivar still showed higher values compared to all other cultivars. Vegetation fraction follows a similar trend as spectral features, increasing until the tuber stage and subsequently decreasing after tuber stage. Notably, the RN cultivar showed high variation in vegetation fraction after the tuber stage. However, the RR, RB, and AL cultivars demonstrated low variation in vegetation fraction after the tuber stage for both the years. These unique responses under different N-rates, environmental factors, and growth stages can provide valuable insights into the real-world behavior of the physiological and biological traits of different cultivars that may contribute to differences in tuber yield potential. Capturing these time-series variation enhance yield prediction in deep learning model's by providing detailed insights into the dynamic growth patterns and health of crops, allowing the model to identify trends, detect early signs of stress, and adapt to seasonal changes for more accurate yield predictions. The study is further evaluated by developing a nonlinear complex relation between temporal variations in spatial and spectral features with yield utilizing PLSR algorithm for feature engineering.

D. Feature Engineering Using PLSR Algorithm

Using data from the T1 to T5, we depicted the variation of VIP score with optimal N_c at increment of vegetative growth stage in a step-by-step manner. In addition, the VIP score of predictor variables by concatenating datasets from T1 to T5 growth stages is also presented via histogram, as shown Fig. 13(a)–(e). The analysis shows that the VIP score of vegetation features by concatenating the temporal datasets until the tuberization stage (i.e., [T1, T2, T3]) exhibits a very low score (represents low importance in yield prediction). During these growth stages, which mostly involve physiological and structural changes, different sets of vegetation features with high VIP scores were found significant at each time-step of vegetative growth stages, as shown in Fig. 13(a)–(c). The dynamic range of top six robust vegetation features from T1 to T3 stage are found to be varying from ≈ 0.004 to 0.095 . Moreover, no significant common features were observed during these stages. The low significance (i.e., low VIP) of the predictors until tuberization stage might reflect its ability to capture key physiological and structural changes during the crop growth stages.

However, after concatenating the datasets of the bulking stage (i.e., [T1, T2, T3, T4]), the focus of the plant physiology shifts from vegetative growth to tuber formation and nutrient accumulation in potato tuber, led to the high VIP scores. In addition,

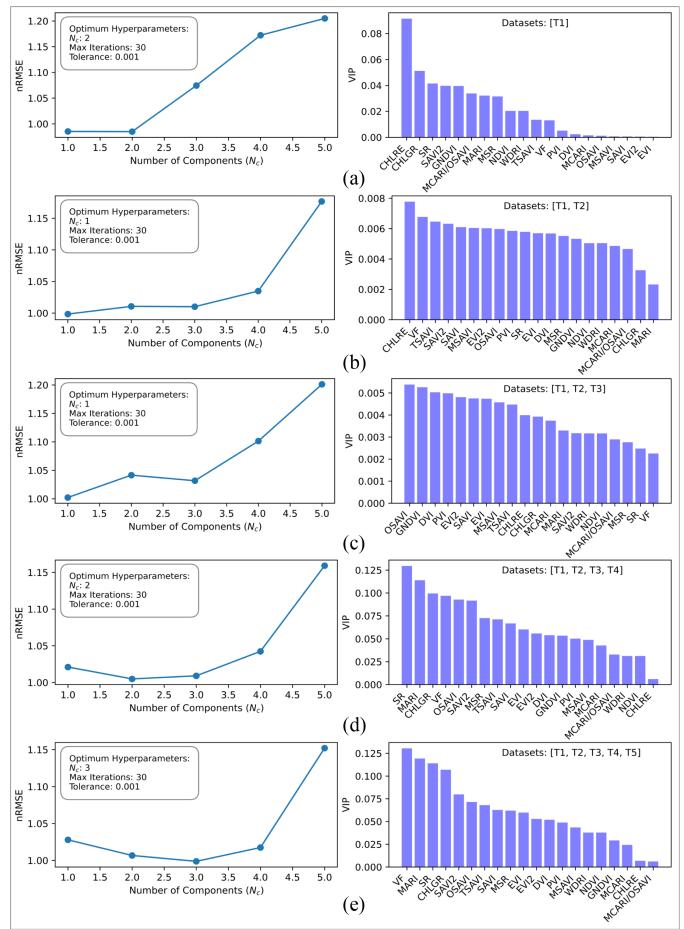


Fig. 13. (a)–(e) Variable of importance in projection (VIP) based feature engineering utilizing PLSR algorithm. The left side figures represent the optimization of a number of components (N_c), which captures significant covariance and weight to provide variable importance. The right side figures represent the VIP from highest to lowest. The VIP of variables is reported by concatenating datasets from T1 to T5 vegetative growth stage.

several common vegetation features with high VIP were also observed during the tuberization and maturity growth stages. The increase in VIP score of the top six features were observed, with the dynamic range of ≈ 0.076 – 0.126 , after tuberization stage, as shown seen in Fig. 13(d) and (e). These top six features, includes SR, MARI, CHLGR, VF, OSAVI, and SAVI2, remained consistent during the bulking and maturity stages. Interestingly, the PLSR algorithm showed that, in addition to the spectral features, the spatial feature, VF, also played significant importance in predicting crop yield, with high VIP score of ≈ 0.126 . PLSR suggested indices, namely, SR and CHLGR demonstrate sensitivity to changes in chlorophyll content and leaf area, which are closely associated with tuber development and nutrient assimilation processes [72]. In addition, indices like MARI, SAVI2, and OSAVI are known for their sensitivity to stress-induced alterations in plant physiology, making them valuable for predicting yield variations under diverse environmental conditions [73].

The shift in the importance of vegetation features from vegetative growth stages to the bulking stage highlights the dynamic nature of crop development and the various physiological processes that influence yield formation. Therefore, we

TABLE III
DESCRIPTION OF INPUT DIMENSION FOR MODEL IMPLEMENTATION

Growth stages	Date		Input dimension: $(n \times t \times p)$				
	2020	2021	3F	4F	5F	6F	7F
T1	28/5	26/5	(264, 1, 3)	(264, 1, 4)	(264, 1, 5)	(264, 1, 6)	(264, 1, 7)
[T1,T2]	03/6	01/6	(264, 2, 3)	(264, 2, 4)	(264, 2, 5)	(264, 2, 6)	(264, 2, 7)
	10/6	09/6	(264, 3, 3)	(264, 3, 4)	(264, 3, 5)	(264, 3, 6)	(264, 3, 7)
[T1,...,T3]	15/6	14/6	(264, 4, 3)	(264, 4, 4)	(264, 4, 5)	(264, 4, 6)	(264, 4, 7)
	23/6	23/6	(264, 5, 3)	(264, 5, 4)	(264, 5, 5)	(264, 5, 6)	(264, 5, 7)
[T1,...,T4]	16/7	14/7	(264, 6, 3)	(264, 6, 4)	(264, 6, 5)	(264, 6, 6)	(264, 6, 7)
	21/7	23/7	(264, 7, 3)	(264, 7, 4)	(264, 7, 5)	(264, 7, 6)	(264, 7, 7)
[T1,...,T5]	29/7	30/7	(264, 8, 3)	(264, 8, 4)	(264, 8, 5)	(264, 8, 6)	(264, 8, 7)

identified the top six vegetation features, as indicated by the high VIP score during and after the bulking stage, as robust indicators. These features include one spatial feature, VF, and five spectral features: SR, CHLGR, MARI, SAVI2, and OSAVI. These selected vegetation features, along with the N-rates, were used as inputs to the proposed CAR Conv1D-BiGRU-BiLSTM-Net model to optimize yield prediction. The decision to limit the analysis to six vegetation indices was made to ensure that the most relevant information from high VIP scores was captured while maintaining interpretability. This approach also aimed to improve computational efficiency by reducing dimensionality, thereby minimizing the risks of over fitting, under fitting, and misinterpretation in the proposed deep learning model.

E. Model Implementation

The model was implemented using the datasets ranging from $(264 \times 1 \times 3)$ to $(264 \times 8 \times 7)$, as detailed in Table III. A total of 264 samples were collected at the vegetative growth stages from T1 to T5, with 120 samples from the year 2020 and 144 from the year 2021. The data of same vegetative stage from the years 2020 and 2021 were combined to increase the training sample size that may lead better generalization of model performance. The model training started with 80% of the three predictor variables and gradually incorporating additional features as the training progressed. The first two predictor variables, i.e., N-rate and VF feature, were kept constant in all datasets. The N-rate remained unchanged throughout the growth stages, while temporal dependencies of spatial and spectral features from T1 to T5 were utilized across the input dimensions. PLSR suggested top 5-robust spectral features (SR, MARI, SAVI2, OSAVI, and CHLGR) were included into the input dimensions. The number of spectral feature predictors were systematically increased based on their VIP from highest to lowest, one at a time, along with its temporal dimensions to the input subset. This approach aimed to determine the optimal number of predictors and understanding the importance of vegetation temporal dynamics in predicting yield.

The random search optimization suggested the following optimal hyperparameter for the model: Three Conv1D layers with 64, 32, and 16 filters, each using a kernel size of 1 to capture localized temporal patterns while minimizing complexity. Following each Conv1D layer, a dense layer was applied with 64, 32, and 16 units, respectively, utilizing ReLU activation for nonlinearity and regularization. A residual connection preserved

original feature representations and mitigated gradient issues. Sequentially, a BiGRU layer with 32 units captured short-term dependencies, while a BiLSTM layer with 32 units captured long-term patterns. A context-aware attention mechanism, inspired by transformers, used to focus on relevant temporal features using softmax-activated attention weights. Finally, dense layer with unit 1 and linear activation function used to predict final yield. The model was compiled with the Adam optimizer (learning rate = 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-5}$) and trained for 100 epochs with a batch size of 50.

F. Model Performance

The model's performance metrics (i.e., R^2 and RMSE), as tabulated in Table IV, were computed by averaging an ensemble of five models, with bold values indicating the best performance metric. The performance metrics showed notable patterns on increasing input dimension across number of features and its temporal dynamics. A significant positive impact on R^2 and RMSE is evident as temporal information is incorporated from T1 to T5 growth stage, supporting importance of time-series analysis on yield prediction. Particularly, after the incorporating of late tuberization (T3) and bulking (T4) growth stages information in input datasets demonstrated improved R^2 and RMSE values on test datasets, when compared to early growth stages (T1, T2, and T3). However, the incorporation of maturity (T5) growth stage led to saturation in the value of R^2 and RMSE value. This confirmed model's ability to capture significant physiological changes and nutrient distribution during the late bulking and tuber formation growth stages, demonstrating that incorporating time-series features of these stages improved the yield predictive performance. Furthermore, the investigation into robust features subset from the model's performance underscored the importance of feature selection. The model showed higher R^2 and lower RMSE values as the input subsets dimension increase up to five features (5F) (i.e., SR, MARI, CHLGR, VF, and N-rate). The input subset dimension, both before and beyond 5F showed low R^2 and high RMSE. This decline in the model's performance suggests a potential risk of under fitting, over fitting, or saturation of the information captured.

The CAR Conv1D-BiGRU-BiLSTM-Net with 5F input subset exhibited strong predictive capability, achieving an average $R^2 = 0.78$ and average RMSE = 8.35×10^3 kg/ha (with % RMSE of 16.4%), when considering temporal information from T1 to T4. Once we obtained the model accuracy at optimum number of predictors, we compared its performance with other deep

TABLE IV
PROPOSED MODEL PERFORMANCE

Growth stage	Date		Features subset									
			3F		4F		5F		6F		7F	
	2020	2021	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]
[T1]	28/5	26/5	0.38	11.87 (23.3%)	0.42	11.67 (22.9%)	0.46	11.32 (22.2%)	0.44	11.72 (22.9%)	0.46	11.66 (22.9%)
[T1, T2]	03/6	01/6	0.42	11.73 (23.0%)	0.50	11.11 (21.8%)	0.53	11.03 (21.6%)	0.53	11.39 (22.3%)	0.52	11.22 (22.0%)
	10/6	09/6	0.45	11.50 (22.6%)	0.46	11.45 (22.5%)	0.48	11.23 (22.0%)	0.42	12.02 (23.5%)	0.47	11.61 (22.8%)
[T1, ..., T3]	15/6	14/6	0.42	11.61 (22.8%)	0.44	11.36 (22.3%)	0.48	11.07 (21.7%)	0.46	11.91 (23.4%)	0.45	11.59 (22.7%)
	23/6	23/6	0.48	11.50 (22.6%)	0.52	11.05 (21.7%)	0.55	10.96 (21.5%)	0.54	11.10 (21.7%)	0.52	10.99 (21.6%)
[T1, ..., T4]	16/7	14/7	0.65	9.77 (19.2%)	0.71	9.17 (17.9%)	0.72	8.99 (17.6%)	0.68	9.89 (19.4%)	0.65	10.12 (19.8%)
	21/7	23/7	0.69	9.47 (18.6%)	0.74	8.52 (16.7%)	0.78	8.35 (16.4%)	0.73	8.77 (17.2%)	0.70	9.56 (18.7%)
[T1, ..., T5]	29/7	30/7	0.70	9.38 (18.4%)	0.73	8.67 (17.0%)	0.77	8.40 (16.5%)	0.73	8.95 (17.6%)	0.73	9.16 (17.9%)

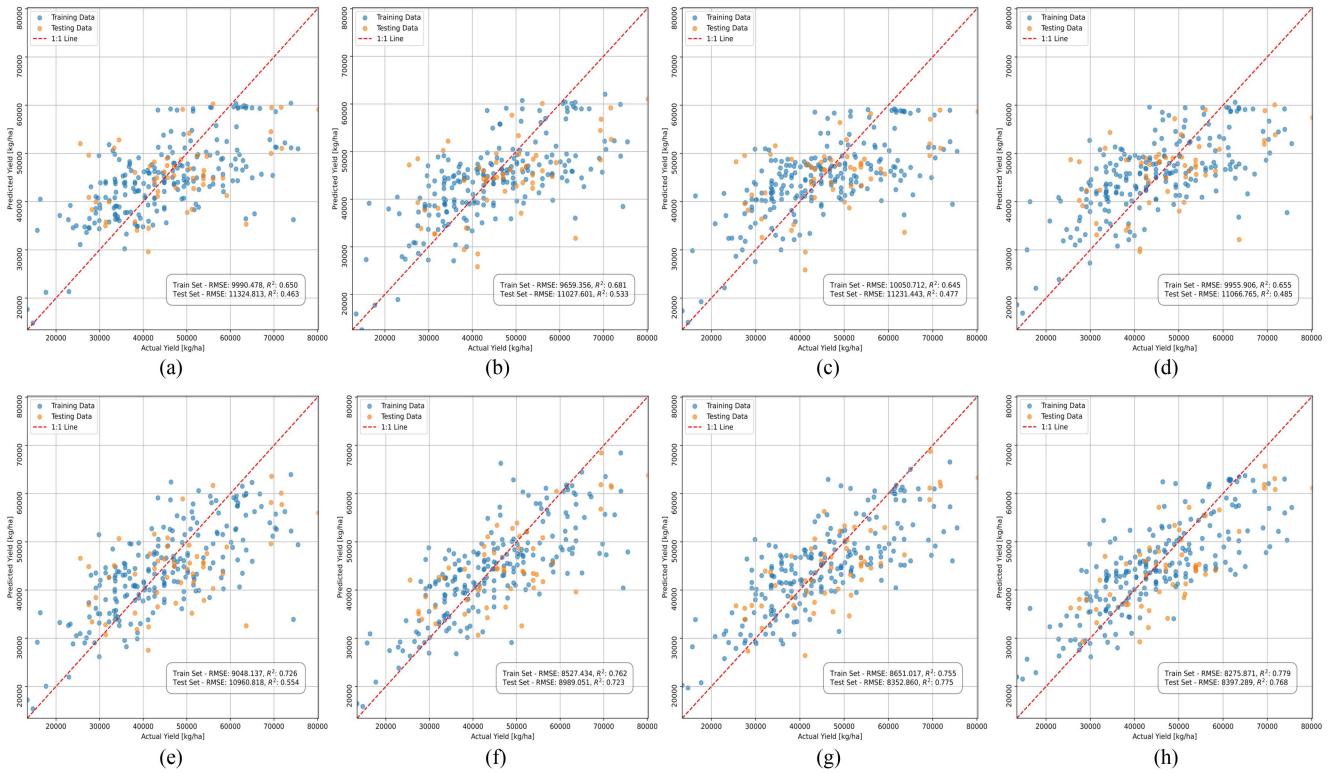


Fig. 14. (a)–(h) 5F-CAR Conv1D-BiGRU-BiLSTM-Net performance throughout the growth stages mention in Table IV. Where, 5F includes, SR, MARI, CHLGR, VF, and N-rate features. The performance metrics for train and testing datasets were recorded at integration of each vegetative growth stages. (a) T1. (b) and (c) T2. (d) and (e) T3. (f) and (g) T4, and (h) T5.

learning models, including Convolutional 1D (Conv1D), BiGRU, BiLSTM, and Conv1D-BiLSTM-Attention networks. The performance metrics of the proposed model outperformed all other deep learning models, except at the T1 vegetative stage, where Conv1D showed a better performance, as shown in Table V. The representation, as shown in Table VI, included the enumeration of 5F-CAR Conv1D-BiGRU-BiLSTM-Net trainable parameters intricately adjusted throughout the training phase to minimize the loss function. The model, characterized by 43218 trainable parameters, demonstrates superior predictive performance by integrating temporal dimension till T4 with optimum 5F subset (i.e., 5F: SR, MARI, CHLGR, VF, and N-rate). The prediction learning for 5F-CAR Conv1D-BiGRU-BiLSTM-Net at each time-step of vegetative growth stage are plotted in Fig. 14.

In summary, our proposed approach emphasized the significance of nutrient features and the temporal dynamics of spectral and spatial characteristics in optimizing predictive models for time-series yield prediction. With the given input sample size across all dimensions and their complexity, including various varieties and N-rates, we achieved an RMSE of 16.4%. The inclusion of substantial variability in agricultural systems adds significant real-world complexity to the model, influencing crop growth and yield prediction. This diversity provided a foundation for testing proposed model performance in complex scenarios, where numerous factors interact in nonlinear interactions. Although the proposed model is computationally intensive, it demonstrated the ability to better capture these complexities. Therefore, based on this study, we suggest that increasing the sample size by including yield responses, varieties, and N-rates

TABLE V
PROPOSED MODEL PERFORMANCE COMPARISON WITH OTHER DEEP LEARNING MODELS

Growth stage	Date		5F (SR, MARI, CHLGR, VF and N-rate)									
	2020	2021	Conv1D		BiGRU		BiLSTM		Conv1D-BiLSTM-Attention		Proposed	
			R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]	R ²	RMSE [10 ³ kg/ha][%]
[T1]	28/5	26/5	0.48	11.31 (22.2%)	0.36	12.21 (23.9%)	0.43	11.88 (23.3%)	0.45	11.67 (22.9%)	0.46	11.32 (22.2%)
[T1, T2]	03/6	01/6	0.49	11.24 (22.0%)	0.44	11.59 (22.7%)	0.46	11.50 (22.6%)	0.57	11.39 (23.9%)	0.53	11.03 (21.6%)
	10/6	09/6	0.52	10.86 (21.3%)	0.47	11.54 (22.6%)	0.44	11.27 (22.1%)	0.42	11.40 (22.3%)	0.48	11.23 (22.0%)
[T1, ..., T3]	15/6	14/6	0.47	11.36 (22.3%)	0.49	11.24 (22.0%)	0.42	11.59 (22.7%)	0.45	11.38 (22.3%)	0.48	11.07 (21.7%)
	23/6	23/6	0.55	10.64 (20.9%)	0.50	11.01 (21.6%)	0.45	11.34 (22.2%)	0.47	11.24 (22.0%)	0.55	10.96 (21.5%)
[T1, ..., T4]	16/7	14/7	0.61	10.08 (19.8%)	0.66	10.17 (19.9%)	0.60	11.03 (21.6%)	0.61	10.01 (19.6%)	0.72	8.99 (17.6%)
	21/7	23/7	0.65	9.59 (18.8%)	0.68	9.63 (18.9%)	0.63	10.42 (20.4%)	0.69	9.51 (18.6%)	0.78	8.35 (16.4%)
[T1, ..., T5]	29/7	30/7	0.65	9.65 (18.9%)	0.69	9.52 (18.6%)	0.63	9.57 (18.7%)	0.70	9.40 (18.4%)	0.77	8.40 (16.5%)

TABLE VI
5F-CONV1D-BiLSTM-ATTENTION MODEL SUMMARY WITH SPECIFICATIONS

Layer type	Specification	Parameter
Conv1D	Filters: 64, Kernel: 1	448
Dense	Units: 64	4160
Conv1D	Filters: 32, Kernel: 1	2080
Dense	Units: 32	1056
Conv1D	Filters: 16, Kernel: 1	528
Dense	Units: 16	272
Residual connection	Units: 16	112
Bi-GRU	Units: 32	9600
Bi-LSTM	Units: 32	24832
Context vector multiply	Softmax activation	65
GlobalMaxPooling1D	-	0
Dense	Units: 1	65
Trainable parameters		43218
Nontrainable parameters		0
Optimizer parameters		86438
Total parameter		129656

could lead to better learning and further improvement in the model performance.

VI. CONCLUSION

Our findings highlighted the capability of UAS-based multispectral imaging over wide range of precision agriculture application, specifically in advancing spatial and spectral quality of UAS data, autosegmentation at subplot level, accurate feature extraction, feature engineering, and time-series yield prediction. In addition, the study also analyzed the impact of N-rate on tuber yield of the different potato cultivars. Our findings indicate that the RR cultivar consistently produced higher yields overall. In contrast, the RN cultivar consistently exhibits the lowest yields. The RB cultivar average tuber yield was found to be relatively stable, initially increasing with higher N-rates but showing a slight decrease as N-rates increases in both the years, although exceptions exist at specific N-rates.

This study employed a linear unmixing(LUM) model followed by nonbinary Otsu-based adaptive semantic autosegmentation to enhance and segment endmember spectral information present in raw orthomosaic images. This approach enabled efficient and accurate soil masking by isolating pixel coordinates using Otsu algorithm defined threshold. The comprehensive spatial and spectral feature extraction approach from the soil masked multispectral images provided detailed insight into vegetation dynamics, offering improved accuracy and robust characterization across different vegetation growth stages. The dynamic

changes in spatial and spectral features over time provided significance of the potato crop growth in capturing physiological, biochemical, and biophysical changes, which affects end season tuber-yield across different cultivars. Overall, by elucidating the relationships between spectral and spatial features, N-rates across cultivar responses, our analysis provided valuable insights on deep learning models for yield prediction by integrating time-series multispectral data, offering a robust framework to support data-driven decision-making in precision agriculture.

The partial least squares regression (PLSR) algorithm was employed to perform feature engineering/dimensionality reduction on the extracted features from soil-masked multispectral images. The PLSR analysis highlighted a shift in the significance of vegetation features from emergence (T1) growth stages to the maturity (T5) stage, reflecting dynamic crop development and physiological processes influencing yield formation. The six features with the highest VIP scores, namely, SR, CHLGR, MARI, SAVI2, OSAVI, and VF along with N-rate, were selected as robust inputs to train the proposed model. Notably, the integration of robust multisource datasets, incorporating spectral, spatial, and nutrient features such as SR, MARI, CHLGR, VF, and N-rate, demonstrated superior performance by outperforming other deep learning models, achieving a high R² value of 0.775 and a low percentage RMSE of 16.4%.

In our future work, we plan to integrate advanced remote sensing techniques, such as drone-based hyperspectral imaging and LiDAR, along with satellite data for comprehensive crop monitoring. We will explore cutting-edge deep learning architectures, including autoencoders, transformers, generative adversarial networks, diffusion model's, and incorporate reinforcement learning for real-time agronomic optimization. In addition, we intend to enhance the current model using long term temporal analysis to assess the impact of climate on yield, while cross-validation and hyperparameter tuning will be used to refine model performance. Our ultimate goal is to develop scalable, user-friendly interfaces for field use, ensuring applicability across different crops and field sizes.

ACKNOWLEDGMENT

The authors would like to thank Daniel O. Wall for creating shape file used in this study. They also would like to thank the Editor and anonymous reviewers for their valuable comments and suggestions, which have significantly contributed to improving this article.

REFERENCES

- [1] P. Jiang et al., "Prediction dynamics in cotton aphid using unmanned aerial vehicle multispectral images and vegetation indices," *IEEE Access*, vol. 11, pp. 5908–5918, 2023.
- [2] J. Deng et al., "Quantitative estimation of wheat stripe rust disease index using unmanned aerial vehicle hyperspectral imagery and innovative vegetation indices," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–11, 2023, Art. no. 4406111.
- [3] Y. Wu, Z. Mao, L. Guo, C. Li, and L. Deng, "Forest volume estimation method based on allometric growth model and multi-source remote sensing data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 8900–8912, 2023.
- [4] H. Liu et al., "A glimpse of ocean color remote sensing from moon-based Earth observations," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–11, 2022, Art. no. 4206811.
- [5] S.-E. Qian, "Hyperspectral satellites, evolution, and development history," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 7032–7056, 2021.
- [6] C. P. Donahue et al., "Bridging the gap between airborne and spaceborne imaging spectroscopy for mountain glacier surface property retrievals," *Remote Sens. Environ.*, vol. 299, 2023, Art. no. 113849.
- [7] M. Shimoni, R. Haelterman, and C. Perneel, "Hyperpectral imaging for military and security applications: Combining myriad processing and sensing techniques," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 101–117, Jun. 2019.
- [8] R. O. Green et al., *AVIRIS-Classic: L2 Calibrated Reflectance, Facility Instrument Collection, VI*, vol. 1, ORNL DAAC, Oak Ridge, TN, USA, 2023. [Online]. Available: <https://doi.org/10.3334/ORNLDAA/2154>
- [9] L. A. Zimmerman and J. P. Kerekes, "Comparison of methane detection using shortwave and longwave infrared hyperspectral sensors under varying environmental conditions," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 2517–2531, 2023.
- [10] K. Cawse-Nicholson et al., "NASA's surface biology and geology designated observable: A perspective on surface imaging algorithms," *Remote Sens. Environ.*, vol. 257, 2021, Art. no. 112349.
- [11] A. M. Wijata et al., "Taking artificial intelligence into space through objective selection of hyperspectral Earth observation applications: To bring the "brain" close to the "eyes" of satellite missions," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 2, pp. 10–39, Jun. 2023.
- [12] P. Buschkamp et al., "CHIME's Hyperspectral Imager (HSI): Status of instrument design and performance at PDR," in *Proc. Int. Conf. Space Opt.*, 2023, vol. 12777, pp. 1379–1398.
- [13] A. Torres-Rua et al., "Advances, challenges and opportunities in the application of uncrewed aerial systems in the characterization of water and energy fluxes in agriculture and urban environments by the aggear UAV research program," in *Proc. ASA, CSSA, SSSA Int. Annu. Meeting*, 2023.
- [14] M. Belgiu, M. Marshall, M. Boschetti, M. Pepe, A. Stein, and A. Nelson, "PRISMA and Sentinel-2 spectral response to the nutrient composition of grains," *Remote Sens. Environ.*, vol. 292, 2023, Art. no. 113567.
- [15] M. Wocher, K. Berger, J. Verrelst, and T. Hank, "Retrieval of carbon content and biomass from hyperspectral imagery over cultivated areas," *ISPRS J. Photogrammetry Remote Sens.*, vol. 193, pp. 104–114, 2022.
- [16] M. Maimaitijiang, V. Sagan, P. Sidiike, S. Hartling, F. Esposito, and F. B. Fritschi, "Soybean yield prediction from UAV using multimodal data fusion and deep learning," *Remote Sens. Environ.*, vol. 237, 2020, Art. no. 111599.
- [17] S. Bhadra, V. Sagan, S. Sarkar, M. Braud, T. C. Mockler, and A. L. Eveland, "PROSAIL-Net: A transfer learning-based dual stream neural network to estimate leaf chlorophyll and leaf angle of crops from UAV hyperspectral images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 210, pp. 1–24, 2024.
- [18] S. Torabian et al., "Nitrogen altered petiole nutrient concentration, yield, and quality of castle russet and echo russet potatoes in an irrigated arid region," *Amer. J. Potato Res.*, vol. 100, no. 6, pp. 464–478, 2023.
- [19] USDA, *Statistical Bulletins: Potatoes*, 2023 Summary. [Online]. Available: <https://downloads.usda.library.cornell.edu/usda-esmis/files/fx719m44h/6q184c88z/v118t7605/pots0924.pdf>
- [20] A. A. Zaeen, L. K. Sharma, A. Jasim, S. Bali, A. Buzzia, and A. Alyokhin, "Yield and quality of three potato cultivars under series of nitrogen rates," *Agrosyst. Geosci. Environ.*, vol. 3, no. 1, 2020, Art. no. e20062.
- [21] N. Lu et al., "Estimation of nitrogen nutrition status in winter wheat from unmanned aerial vehicle based multi-angular multispectral imagery," *Front. Plant Sci.*, vol. 10, 2019, Art. no. 1601.
- [22] M. A. Hossen, P. K. Diwakar, and S. Ragi, "Total nitrogen estimation in agricultural soils via aerial multispectral imaging and LIBS," *Sci. Rep.*, vol. 11, no. 1, 2021, Art. no. 12693.
- [23] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [24] Y.-X. Wang and Y.-J. Zhang, "Nonnegative matrix factorization: A comprehensive review," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 6, pp. 1336–1353, Jun. 2013.
- [25] B. Verma, R. Prasad, P. K. Srivastava, P. Singh, A. Badola, and J. Sharma, "Evaluation of simulated AVIRIS-NG imagery using a spectral reconstruction method for the retrieval of leaf chlorophyll content," *Remote Sens.*, vol. 14, no. 15, 2022, Art. no. 3560.
- [26] J. Cheng, C. Deng, Y. Su, Z. An, and Q. Wang, "Methods and datasets on semantic segmentation for unmanned aerial vehicle remote sensing images: A review," *ISPRS J. Photogrammetry Remote Sens.*, vol. 211, pp. 1–34, 2024.
- [27] M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, "A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction," *IEEE Access*, vol. 9, pp. 63406–63439, 2021.
- [28] J. Skobalski, V. Sagan, H. Alifu, O. Al Akkad, F. A. Lopes, and F. Grignola, "Bridging the gap between crop breeding and GeoAI: Soybean yield prediction from multispectral UAV images with transfer learning," *ISPRS J. Photogrammetry Remote Sens.*, vol. 210, pp. 260–281, 2024.
- [29] S. A. Yadav, R. Prasad, A. K. Vishwakarma, J. Sharma, B. Verma, and P.K. Srivastava, "Optimization of dual-polarized bistatic specular scatterometer for studying microwave scattering response and vegetation growth parameters retrieval of paddy crop using a machine learning algorithm," *Comput. Electron. Agriculture*, vol. 175, 2020, Art. no. 105592.
- [30] Y. Ma, Z. Zhang, Y. Kang, and M. Özdogan, "Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach," *Remote Sens. Environ.*, vol. 259, 2021, Art. no. 112408.
- [31] N. Ali et al., "Field-scale precision: Predicting grain yield of diverse wheat breeding lines using high-throughput UAV multispectral imaging," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 11419–11433, 2024.
- [32] B. Verma, R. Prasad, P.K. Srivastava, S. A. Yadav, P. Singh, and R. Singh, "Investigation of optimal vegetation indices for retrieval of leaf chlorophyll and leaf area index using enhanced learning algorithms," *Comput. Electron. Agriculture*, vol. 192, 2022, Art. no. 106581.
- [33] J. Zhao, Y. Zhong, X. Hu, L. Wei, and L. Zhang, "A robust spectral-spatial approach to identifying heterogeneous crops using remote sensing imagery with high spectral and spatial resolutions," *Remote Sens. Environ.*, vol. 239, 2020, Art. no. 111605.
- [34] Y.-C. Lin and A. Habib, "Quality control and crop characterization framework for multi-temporal UAV LiDAR data over mechanized agricultural fields," *Remote Sens. Environ.*, vol. 256, 2021, Art. no. 112299.
- [35] J. Sun, Z. Lai, L. Di, Z. Sun, J. Tao, and Y. Shen, "Multilevel deep learning network for county-level corn yield estimation in the us corn belt," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5048–5060, 2020.
- [36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [37] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.
- [38] D. She and M. Jia, "A BiGRU method for remaining useful life prediction of machinery," *Measurement*, vol. 167, 2021, Art. no. 108277.
- [39] S. M. M. Nejad, D. Abbasi-Moghadam, A. Sharifi, N. Farmonov, K. Amankulova, and M. László, "Multispectral crop yield prediction using 3D-convolutional neural networks and attention convolutional LSTM approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 254–266, 2023.
- [40] A. Vaswani et al., "Attention is all you need," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, 2017, vol. 30, pp. 6000–6010.
- [41] J.-W. Ma, C.-H. Nguyen, K. Lee, and J. Heo, "Regional-scale rice-yield estimation using stacked auto-encoder with climatic and MODIS data: A case study of South Korea," *Int. J. Remote Sens.*, vol. 40, no. 1, pp. 51–71, 2019.
- [42] T. Gervais, A. Creelman, X.-Q. Li, B. Bizimungu, D. De Koeyer, and K. Dahal, "Potato response to drought stress: Physiological and growth basis," *Front. Plant Sci.*, vol. 12, 2021, Art. no. 698060.

- [43] J. M. Blauer et al., "Rainier Russet: A dual use russet potato with long tuber dormancy, excellent process quality, and high early harvest packaging efficiency," *Amer J. Potato Res.*, vol. 101, no. 1, pp. 17–33, 2024.
- [44] R. Novy et al., "Alturas: A multi-purpose, russet potato cultivar with high yield and tuber specific gravity," *Amer J. Potato Res.*, vol. 80, pp. 295–301, 2003.
- [45] J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [46] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
- [47] G. Priyanka, P. Rajalakshmi, and J. Kholova, "Two dimensional histogram based on relative entropy thresholding for crop segmentation using UAV images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 3518–3521.
- [48] Z. Luo, W. Yang, Y. Yuan, R. Gou, and X. Li, "Semantic segmentation of agricultural images: A survey," *Inf. Process. Agriculture*, vol. 11, no. 2, pp. 172–186, 2023.
- [49] H. Yang et al., "Integrating multidimensional feature indices and phenological windows for mapping cropping patterns in complex agricultural landscape regions," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, pp. 7448–7459, 2024.
- [50] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sens. Environ.*, vol. 8, no. 2, pp. 127–150, 1979.
- [51] K. Yawata, T. Yamamoto, N. Hashimoto, R. Ishida, and H. Yoshikawa, "Mixed model estimation of rice yield based on NDVI and GNDVI using a satellite image," in *Proc. SPIE Conf. Remote Sens. Agriculture, Ecosyst., Hydrol.*, 2019, vol. 11149, pp. 380–385.
- [52] C. F. Jordan, "Derivation of leaf-area index from quality of light on the forest floor," *Ecology*, vol. 50, no. 4, pp. 663–666, 1969.
- [53] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, no. 1–2, pp. 195–213, 2002.
- [54] A. A. Gitelson, O. B. Chivkunova, and M. N. Merzlyak, "Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves," *Amer J. Botany*, vol. 96, no. 10, pp. 1861–1868, 2009.
- [55] A. A. Gitelson, A. Viña, V. Ciganda, D. C. Rundquist, and T. J. Arkebauer, "Remote estimation of canopy chlorophyll content in crops," *Geophys. Res. Lett.*, vol. 32, no. 8, 2005.
- [56] D. Major, F. Baret, and G. Guyot, "A ratio vegetation index adjusted for soil brightness," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 727–740, 1990.
- [57] J. Qi, A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian, "A modified soil adjusted vegetation index," *Remote Sens. Environ.*, vol. 48, no. 2, pp. 119–126, 1994.
- [58] G. Rondeaux, M. Steven, and F. Baret, "Optimization of soil-adjusted vegetation indices," *Remote Sens. Environ.*, vol. 55, no. 2, pp. 95–107, 1996.
- [59] F. Baret and G. Guyot, "Potentials and limits of vegetation indices for LAI and APAR assessment," *Remote Sens. Environ.*, vol. 35, no. 2–3, pp. 161–173, 1991.
- [60] J. M. Chen, "Evaluation of vegetation indices and a modified simple ratio for boreal applications," *Can. J. Remote Sens.*, vol. 22, no. 3, pp. 229–242, 1996.
- [61] C. S. Daughtry, C. Walther, M. Kim, E. B. De Colstoun, and J. McMurtrey III, "Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance," *Remote Sens. Environ.*, vol. 74, no. 2, pp. 229–239, 2000.
- [62] A. J. Richardson and J. H. Everitt, "Using spectral vegetation indices to estimate rangeland productivity," *Geocarto Int.*, vol. 7, no. 1, pp. 63–69, 1992.
- [63] A. A. Gitelson, "Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation," *J. Plant Physiol.*, vol. 161, no. 2, pp. 165–173, 2004.
- [64] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: A basic tool of chemometrics," *Chemometrics Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, 2001.
- [65] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [66] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 2, pp. 281–305, 2012.
- [67] M. Khakbazan, R. M. Mohr, D. J. Tomasiewicz, A. Nelson, and B. Bizimungu, "Economic response of potato to nitrogen rate, timing of nitrogen application, cultivar, and irrigation," *Can. J. Plant Sci.*, vol. 104, no. 6, pp. 652–669, 2024.
- [68] B. A. King, D. D. Tarkalson, D. L. Bjorneberg, and J. P. Taberna, "Planting system effect on yield response of Russet Norkotah to irrigation and nitrogen under high intensity sprinkler irrigation," *Amer J. Potato Res.*, vol. 88, pp. 121–134, 2011.
- [69] K. Djaman, S. Irmak, K. Koudahe, and S. Allen, "Irrigation management in potato (*Solanum Tuberosum L.*) production: A review," *Sustainability*, vol. 13, no. 3, 2021, Art. no. 1504.
- [70] W. Li, B. Xiong, S. Wang, X. Deng, L. Yin, and H. Li, "Regulation effects of water and nitrogen on the source-sink relationship in potato during the tuber bulking stage," *PLoS One*, vol. 11, no. 1, 2016, Art. no. e0146877.
- [71] N. Liu et al., "Growth stages classification of potato crop based on analysis of spectral response and variables optimization," *Sensors*, vol. 20, no. 14, 2020, Art. no. 3995.
- [72] J. G. Clevers, L. Kooistra, and M. M. Van den Brande, "Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop," *Remote Sens.*, vol. 9, no. 5, 2017, Art. no. 405.
- [73] R. Binte Mostafiz, R. Noguchi, and T. Ahamed, "Agricultural land suitability assessment using satellite Remote Sensing-derived soil-vegetation indices," *Land*, vol. 10, no. 2, 2021, Art. no. 223.



Suraj A. Yadav received the M.S. degree from the University of Mumbai, Mumbai, India, in 2014 and the Ph.D. degree from the Indian Institute of Technology (BHU), Varanasi, India, in 2022, both in physics.

He was a Postdoctoral Research Associate with the Electrical and Computer Engineering Department and Geosystem Research Institute, Mississippi State University, Mississippi State, MS, USA, from 2022 to 2023. He is currently a Postdoctoral Research Associate with the Department of Agricultural and Biological Engineering, Mississippi State University. His research focus includes advancing remote sensing technologies through computational electromagnetic simulations, computer vision, and deep learning models to enhance understanding of environmental factors based on data collected from ground/UAS/Satellite-based microwave/multispectral/hyperspectral sensors. These efforts contribute to the accurate modeling of wave propagation and scattering, which is vital for Earth system monitoring.



Xin Zhang received the B.S. of Agronomy degree in facility agriculture science and engineering from Gansu Agriculture University, Lanzhou, China, in 2013, the M.S. of Agriculture degree in horticulture engineering from Northwest A&F University, Xianyang, China, in 2016, and the Ph.D. of Biological and Agricultural Engineering degree in agricultural automation engineering from Washington State University, Pullman, WA, USA, in 2020.

She was a Postdoctoral Research Associate of Remote Sensing with the University of California, Davis, Davis, CA, USA, in 2020–2021. She is currently an Assistant Professor with the Department of Agricultural and Biological Engineering, Mississippi State University, Mississippi State, MS, USA. Her research interests include agricultural robotics, computer vision, AI in agriculture, unmanned aerial system (UAS), remote sensing, and smart agriculture.

Dr. Zhang was a Managing Editor of Computers and Electronics in Agriculture in 2022–2023. She was a committee member in the University Data Science Committee with Mississippi State University in 2022–2024. She was panelists/reviewers for multiple grant programs, such as USDA NIFA, NASA, and Western Regional Aquaculture Center (WRAC). She is a member of Club of Bologna and American Society of Agricultural and Biological Engineers (ASABE). She has been a certified UAS remote pilot under the FAA since 2017.



Nuwan K. Wijewardane received the B.Sc. degree in agricultural technology and management and the M.Sc. degree in agricultural and biosystem engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2010 and 2013, respectively, and the M.S. degree in agricultural and biological systems engineering and the Ph.D. degree in biological engineering from the University of Nebraska-Lincoln, Lincoln, NE, USA, in 2016 and 2019, respectively.

He is currently an Assistant Professor with the Department of Agricultural and Biological Engineering (ABE), Mississippi State University, Mississippi State, MS, USA. His current research interests include soil and plant sensing, spectroscopy, spectroscopy-based sensor design and fabrication, phenotyping, precision agriculture, and AI.



Max Feldman is currently a Research Geneticist with the U.S. Department of Agriculture (USDA), Agricultural Research Service (ARS), Temperate Tree Fruit and Vegetable Research Unit and is currently an Adjunct Professor with the Department of Horticulture, Washington State University, Pullman, WA, USA. His research is centered around application of high-throughput phenotyping, quantitative genetics, and marker assisted selection to potato germplasm enhancement breeding.

Dr. Feldman is a Senior Editor for the *American Journal of Potato Research*.



Ruijun Qin received the M.Sc. degree in soil science from the Chinese Academy of Agricultural Sciences, Beijing, China, and the Ph.D. degree in agronomy from the Swiss Federal Institute of Technology (ETH Zurich), Zürich, Switzerland.

He is currently an Extension Agronomist with Oregon State University's Hermiston Agricultural Research and Extension Center, Hermiston, OR, USA, and an Associate Professor with the Department of Crop and Soil Science, Osun State University, Osogbo, Nigeria. He was with the San Joaquin Valley Agricultural Research Center, USDA-ARS, Washington, DC, USA, and the University of California, Davis, Davis, CA, USA. He plays a key role in organizing growers meetings and field days in Northeastern Oregon and Southeastern Washington. His extension and research efforts focus on nutrient and water management, crop production, cropping systems, cover crops, organic amendments, environmental quality, and soil health.



Yanbo Huang is currently a Research Agricultural Engineer with the U.S. Department of Agriculture (USDA)-Agricultural Research Service (ARS), Genetics and Sustainable Agriculture Unit, Mississippi State, Mississippi, Mississippi State, MS, USA. Before joining USDA ARS, he was a Research Scientist with Texas A&M University at College Station, College Station, TX, USA, and the Chinese Academy of Mechanics and Electronics Industry, Beijing, China. He originated agriculture cybernetics (Agricultural Cybernetics, Huang and Zhang, Springer, 2021) a

new science which is the basis to study control and communication issues in agricultural production processes. His research interests include precision agricultural process modeling, optimization and control, biological and remote sensing image processing and analysis, and development of smart agricultural information technology.



Sathishkumar Samiappan received the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2014, the M.Tech. degree in computer science and engineering from Amrita Vishwa Vidyapeetham, Ettimadai, India, in 2006, and the B.Engg. degree in electronics and communication engineering from Bharathiar University, Coimbatore, India, in 2003.

He is currently an Associate Research Professor with the Department of Biosystems Engineering and Soil Sciences, University of Tennessee, Knoxville, TN, USA. His research interests include remote sensing, geospatial image analysis, machine learning, and their applications in agriculture, forestry, and environmental science.



Wyatt Young received the B.Sc. degree in agriculture with concentrations in precision agriculture and integrated crop management from Mississippi State University, Mississippi State, MS, USA, in 2020.

He is currently a Physical Science Technician serving with USDA ARS Genetics and Sustainable Agricultural Unit, Mississippi State, MS, USA. He specializes in planning UAS data collection. His research interests include precision agriculture application for spot spraying using UAS-based multispectral and hyperspectral imaging.



Francisco G. Tapia received the B.Sc. degree in agricultural and food systems and the Ph.D. degree in horticulture from Washington State University (WSU), Pullman, WA, USA, in 2021, majored in agricultural technology and management and minored in horticulture.

He is currently a Research Horticulturist with the U.S. Department of Agriculture, Agricultural Research Service (USDA-ARS) and an Adjunct Professor with the Department of Horticulture, WSU. His research is focused on understanding how hops, a key crop in the brewing industry, respond to abiotic stresses, particularly water stress. As water availability becomes a critical issue in agriculture, his work is vital for advancing crop health and yield amidst changing climate conditions.