

Deep Learning IoT Enabled Driven Data Centric Crop Yield Forecasting Using Optimized Temporal-CRN Networks with BIANO Algorithm



Nirmala Teegala Papaiah Teegala^{1*}, Prakash Ganapathy², Anbu Karuppusamy Shanmugam³

¹ Department of ICE, Anna University, Chennai 600025, India

² Department of Biomedical Engineering, Excel Engineering College, Komarapalayam 637303, India

³ Department of Electronics and Communication Engineering, Excel Engineering College, Komarapalayam 637303, India

Corresponding Author Email: nirmala99teegala@gmail.com

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.430115>

ABSTRACT

Received: 28 December 2025

Revised: 30 January 2026

Accepted: 24 February 2026

Available online: 28 February 2026

Keywords:

IoT-based agriculture, crop yield prediction, attention-based deep learning, feature optimization, temporal modelling, precision farming

An IoT-enabled deep learning framework is presented for reliable data-driven crop yield forecasting using real-time agricultural sensor information. The proposed approach integrates a Temporal Attention-Enhanced Convolutional Recurrent Network with hierarchical feature learning to effectively model temporal dependencies, spatial correlations, and nonlinear growth dynamics in agricultural data streams. Multi-source IoT data acquired from soil, weather, and crop condition sensors are subjected to systematic preprocessing, including noise removal, normalization, standardization, and dimensionality reduction using an Autoencoder-based Feature Compression combined with Independent Component Analysis (ICA) to preserve salient agronomic patterns while minimizing redundancy. The refined dataset is partitioned into training and testing subsets following an 80:20 ratio to ensure unbiased model validation. Feature extraction is performed through statistical and higher-order statistical descriptors, augmented with two novel adaptive features—adaptive weighted skewness and adaptive weighted kurtosis—to enhance sensitivity toward yield-critical variations. To further improve predictive efficiency, a Bio-Inspired Adaptive Nutrient Optimization algorithm is employed for optimal feature subset selection and dynamic hyperparameter adjustment, enabling improved convergence and generalization of the TA-CRN model. The proposed framework is implemented using Python 3.7.9 and evaluated as a multi-class yield categorization task, where continuous harvested yield values (kg/ha) are discretized into four agronomically meaningful classes—Low Yield, Medium Yield, High Yield, and Very High Yield—using training-set-derived percentile thresholds; performance is then assessed using accuracy, precision, sensitivity, specificity, F-measure, and macro-averaged classification statistics. Comparative experiments conducted against conventional deep learning models such as CNN, RNN, LSTM, and DNN demonstrate the superiority of the proposed approach. Experimental results indicate that the proposed TA-CRN with the BIANO framework achieves a maximum prediction accuracy of 96.5%, validating its effectiveness for precision agriculture and intelligent crop yield forecasting.

1. INTRODUCTION

Agriculture plays a pivotal role in ensuring food security and maintaining the economic resilience of many parts of the world, especially in developing countries, where agricultural production is a direct contributor to livelihoods. High population rate, rising climatic insecurity, and scarcity of natural resources have heightened the strain on the traditional farming systems, which have made the traditional ways of farming inadequate in providing food to the rising demands. This is increasingly turning modern agriculture into a data-driven and technology-enabled approach that empowers farmers to make evidence-based decisions on time. Real-time environmental monitoring combined with smart analytical systems has great potential in streamlining crop management, reducing risks in production, and improving the overall

performance of yields. The foremost importance is caused by the necessity to offer accurate and early predictions of crop yield to make well-informed choices in terms of irrigation, nutrients, post-harvest storing, and logistics to the market to reduce losses and increase financial benefits. It is on this paradigm that a combination of sensing technologies and new computational intelligence has emerged as a fount of sustainable and efficient agricultural development. The emergence of the Internet of Things (IoT) technologies in the recent past, as well as the growth of artificial intelligence, has revolutionized the sphere of agricultural data acquisition and analysis fundamentally. Related sensors in the IoT yield ongoing data on the nature of the soil, microclimatic zones, and the physiological state of crops during the growth cycle, and generate gigantic volumes of heterogeneous data rich in time. At the same time, deep learning techniques are

exceptionally efficient in both capturing the nonlinear correlations, which are complex in nature, and in identifying meaningful patterns in the huge volume of agricultural data. Nevertheless, agricultural systems are dynamic, nonlinear, and extremely noisy, and hence a massive challenge in terms of a precise forecast of the output. Conventional machine learning and deep learning architectures are likely to be low power to capture long-term temporal correlations, and to put more emphasis on the phenological stage, which has a disproportionate effect on end yield. These shortcomings have spawned an increasing interest in hybrid modeling schemes that integrate more advanced temporal learning schemes, attention-based feature weighting, and optimizing feature selection schemes in order to more effectively utilize multi-source agricultural data.

2. RELATED WORK

Precision agriculture is widely known to have crop yield prediction as one of the pillars because it plays a crucial role in the enhancement of food security and resource use, and is instrumental in the development of agricultural policies. To address these disadvantages, research works have increased more recently, towards data-driven models, including machine learning and deep learning algorithms, that make use of a wide variety of data sources, including IoT sensor networks, satellite-based remote sensing, and historical climate data. Such practices have not only enhanced the accuracy of predictions but also made farm management practices adaptive and real-time decision-making. Moreover, the development of hybrid frameworks where spatial and temporal representations interact with attention mechanisms and optimization strategies has proven to have significant potential in overcoming the limitations of the conventional forecasting models.

2.1 Problem statement

Although large-scale agronomical datasets are increasingly becoming available, the accurate, reliable, and robust prediction of crop yield is a daunting challenge. Agricultural IoT data are often high-dimensional, heterogeneous, noisy, incomplete, and have irregular time aspects, all of which may negatively impact predictive performance in the absence of proper treatment. Most traditional types of deep learning networks give equal significance to all the temporal observations and input variables, thus being unable to capture yield-sensitive growth phases, periods of stress, and important interactions with the environment. Also, unreasonable and superfluous features add to the computational costs and compromise model generalization. The current methods are prone to sensor noise, missing data, seasonal dynamics, and so on, limiting their extrapolability and application in practice. This renders a dire need to have a smart forecasting system that can effectively preprocess data, perform effective and smart feature selection, adaptive time flow modeling, and narrow focus on periods that are of high yield, end of the day, provide a high accuracy prediction, stability, and deployability in changing agricultural landscapes.

2.2 Critical review of existing deep learning and IoT-based crop yield prediction methods

2.2.1 Temporal modeling approaches and their limitations

The temporal learning frameworks have been gradually

incorporated in crop yield forecasting due to the fact that crop productivity is naturally controlled by a series of sequential relationships throughout the growth cycle, weather fluctuations, and delayed physiological reactance. Recurrent architectures based on attention, specifically, had already proved to be useful in modeling long-range temporal interactions between environmental cues and yield outputs, as demonstrated in LSTM-based yield estimation work that had used adaptive weighting of temporal states on region-specific agricultural data [1]. Later advances had added to this concept by adding convolutional encoders to recurrent decoders, with CNN-LSTM pipelines enhanced by attention layers and skip connections, enhancing the capture of both local spatial features and long-term temporal dynamics and reducing the effects of vanishing-gradient propagation through deeper sequence propagation [2]. Simultaneously, image-guided multi-level feature learning had been considered, in which hierarchical visual representations of crop images were combined with regularised recurrent units, including SNN-GRU, to improve predictive resilience to the heterogeneous field conditions [3]. All these studies pointed in the same direction, which was that temporal models had now become complex in that they were hybrids in the sense that they were able to learn both patterns of phenological progression and environmental response. In spite of these developments, significant drawbacks had persisted in the temporal models. Parallel CNN-LSTM systems to predict wheat yield had gained higher representational richness by processing multiple streams of features at once, but remained highly sensitive to synchronized and high-quality multivariate inputs, decreasing their resilience to noisy or incomplete agricultural records [4].

Likewise, dual recurrent architectures, which paired LSTM and GRU modules, e.g., DualSpinNet-type designs, had tried to trade off memory retention and computational efficiency, although their operation remained prone to hyperparameter optimization and sequence-length calibration, particularly between production areas of geographically different regions [5]. Mobile-based real-time rice forecasting Lightweight CNN-LSTM had been used to overcome deployment limitations in smart farming systems, but model compression and inference speeds were often compromised against expressive depth and domain transferability [6]. Moreover, Bi-LSTM-based crop selection and yield prediction models had used bidirectional context to more effectively represent forward and backward temporal correlations, but they remained limited by the lack of interpretability, as well as poorly capturing the assumption that past trends were sufficiently stationary across seasons [7]. Therefore, temporal deep learning was still facing its challenges in its generalization, such as data irregularity, model sensitivity, computational trade-offs, and poor generalization in non-stationary agro-climatic regimes, even though it was making significant progress in forecasting yields.

2.2.2 Shortcomings of dimensionality reduction and feature engineering

Dimensionality reduction and feature engineering were now being implemented as paramount processes to solve the issue of high-dimensionality, redundancy, and noise of agricultural data, especially when combining soil characteristics, weather, spectral indices, and management parameters. Deep learning models that used dimensionality reduction had demonstrated that pre-processing irrelevant or collinear features before prediction could enhance convergence properties and

minimize the computational costs without extracting salient agronomic data [8]. More advanced hybrid feature selection pipelines had also integrated filter-, wrapper-, and optimization-based strategies to find discriminative subsets of variables, improving the predictive behavior of machine learning models in heterogeneous farming environments [9]. Transfer Learning in Gross Primary Production (GPP) Data, remote sensing-based forecasting has shown that the reuse of deep features across large-scale Earth observation data could significantly enhance yield prediction in highly complex areas, including the US Corn Belt, where traditional handcrafted features are limited due to their ability to adequately characterize the area [10]. Thorough studies of deep learning and remote sensing processes had also highlighted how the extraction of features in satellite images, vegetation indices, and climatic rasters had become the key to current yield prediction, particularly when the spatial heterogeneity was required to be maintained to parallel with its temporal trends [11].

Nevertheless, the literature had also found that dimensionality reduction and feature engineering were still inherently limited when considered as autonomously operated preprocessing operations instead of being dynamically interacted with as part of end-to-end learning systems. New explicit dimensionality reduction deep learning models have made models more compact; however, at the cost of using a set of fixed transformations that may ignore non-linear interactions between agro-environmental variables, especially in highly non-convex crop response surfaces [12]. The problem had been sought to be addressed in recent multi-modal data fusion and deep-ensemble works, which tried to provide heterogenous sources (weather, soil, imagery and so on) into predictive systems, however, the fusion process itself posed problems of feature alignment, scale harmonization, and modality imbalance, which might lead to more uncertainty than less unless carefully managed [13]. In even the crop-specific tasks, like predicting potato yield using machine learning and deep learning, manual feature space curation and a restricted ability of the procedure to change to different varieties, cultivation systems, and agro-ecological settings had limited scalability [14]. Thus, despite dimensionality reduction and feature engineering playing a significant role in making the approach more efficient and signal-wise, the more underlying problem of adaptive feature relevance, amidst the dynamically changing multi-source and regionally variable agricultural systems, was not completely addressed.

2.2.3 IoT-based agricultural forecasting and generalization gaps

The arrival of IoT-based precision agriculture was redefining the concept of crop yield prediction to be based no longer on the previous, static historical form of prediction, but on real-time, sensor-based prediction models that are able to react to changing field conditions. The initial research studies had revealed that by combining IoT sensing with climate-sensitive machine learning, the yield prediction would be enhanced through the constant integration of environmental changes, including temperature, humidity, and rainfall, to make the forecasts more adaptive to the changes in climatic conditions [15]. This trend has been extended with the use of precision agriculture smart frameworks, involving deep neural networks with multi-objective optimisation of the sensor deployment, where data quality and downstream yield prediction performance were co-optimised with the spatial

fielding and functional performance of the sensing nodes [16]. Crop recommendation and yield prediction integrated IoT pipelines had also shown that machine learning models could also be enhanced by distributed sensing structures that connected field-scale measurements to cloud-based analytics to support decisions more effectively on a case-by-case basis in farmers [17].

However, the generalization gaps had remained substantial in the context of an application of IoT-based agricultural forecasting, in particular, when considering controlled experimental conditions versus large-scale and heterogeneous production backgrounds. It had been suggested that hierarchical federated learning could be used to provide data privacy and data decentralization in smart farm production systems, enabling local models to learn using distributed farm data without direct data exchange, but non-IID sensor distributions, communication volatility, and regional drift remained significant challenges to convergent global models [18]. On the same note, secure triadic architectures integrating IoT, machine learning, and blockchain have attempted to enhance reliability and traceability in crop prediction pipelines, yet the increased architectural complexity tends to raise latency, energy consumption, and implementation cost, restricting their feasibility in low-resource rural environments [19]. Even relatively simpler smart farming systems, which used Random Forest to estimate crops and predict yields, had already shown that acceptable predictive value, but these sorts of models were often poorly-performing when sensors were calibrated using different farms or when underlying data distributions changed between seasons and between crop types [20]. Therefore, despite the richer observational granularity and near-time predictions that had been made possible by the IoT, the literature showed that there were still unanswered questions surrounding the cross-domain generalization, communication efficiency, data heterogeneity, and robustness to deployment in the face of real-world agricultural variability.

2.2.4 Deep learning gaps through optimization

Machine learning based on optimization has become more acknowledged as a key direction in crop yield prediction since traditional deep architectures were usually characterized by inefficient parameter initialization, inefficient feature weighting, and low adaptability to the complicated agricultural data distributions. Machine learning studies that were oriented towards predictions had demonstrated that, despite the standard predictive models demonstrating reasonable accuracy, the quality of optimization (in terms of feature prioritization, training stability, and convergence to agronomically relevant minima) could severely come into play [21, 22].

Although the concept of optimization was well-known, the literature indicated that the available solutions were still fragmented and not well combined with the overall design of intelligent agricultural systems. Radical evaluations of deep learning and remote sensing methods had repeated the observation that, due to the reported performance improvements, it was often unclear whether the increases came due to architecture design, data preprocessing, or hyperparameter tuning, and thus the effect of optimization mechanisms was often obscured. It was being formalized by hybrid feature selection and optimized machine learning systems and had tried to fuse search-based feature refinement with predictive learning, but these systems could, in effect, still be pipeline dependent and consume large amounts of

computational resources and thus were not scalable to large multi-source agricultural data sets. Similarly, sensor deployment and predictive quality optimization in smart precision agriculture had demonstrated the potential of concurrent optimization of data collection and prediction quality, but they did not entirely address the inherent problem of season, crop, and geographically diverse agro-ecosystem adaptive optimization.

Therefore, the literature revealed the existence of an obvious methodological void: optimization had frequently been considered as a secondary improvement instead of an integrated learning principle that is permeated throughout feature selection, architecture evolution, data fusion, and deployment adaptation. This gap was a powerful rationale behind the creation of tighter-knit, optimization-based frameworks of deep learning that are capable of any type of stable, generalizable, and computationally efficient crop yield prediction in the face of agricultural uncertainty.

2.3 Research gap and motivation

Although there is tremendous advancement in crop yield prediction, there are three open methodological gaps that can be seen in the current literature. To begin with, the majority of previous works focus on the temporal sequence modeling by means of LSTM, GRU, CNN-LSTM, or attention-based architecture, yet most researchers assume that the input representation is informative and stable. In actual IoT-based agricultural settings, though, the raw measurements are high-dimensional, noisy, and irregular, and tend to exhibit sensor drift, missing measurements, and unnecessary inter-variable dependencies. Second, despite the fact that the dimensionality reduction and feature selection methods have been studied independently, they are seldom combined with higher-order agronomic measures that can directly reflect distributional asymmetry, extreme stress events, and heavy-tailed environmental reactions, which are closely linked to yield instability. Third, optimization schemes that are in existence often deal with refining feature subsets or tuning the model parameters independently, but do not introduce a closed-loop adaptive process, that is, coordinated to optimize both representational compactness and predictive convergence. Inspired by these shortcomings, the suggested architecture is not simply an amalgamation of one of the available modules, but a learning pipeline that is deeply integrated to ensure that every element addresses a particular failing of the previous art.

The Autoencoder-based Feature Compression (AFC) step retains nonlinear latent agronomic structure, Independent Component Analysis (ICA) eliminates residual statistical redundancy, adaptive weighted skewness and adaptive weighted kurtosis encode distributional aberrations associated with yield-critical agronomic variables explicitly, and the BIANO algorithm optimization concurrently optimizes the feature subsets and hyperparameters before temporal sequence learning. These optimized representations are then input into the Temporal Attention-Enhanced Convolutional Recurrent Network (TA-CRN), which learns the local temporal variation as well as long-range crop-growth interactions and dynamically focuses on phenologically significant time steps. In this regard, the main contribution of the current research is the synergistic co-design of the compression of representation, the enrichment of agronomic features, adaptive optimization, and learning temporal attention to a single precision-agriculture forecasting structure.

2.4 Core contribution of this research

The novelty of the proposed study lies in the methodological integration and functional interaction of its constituent modules rather than in the isolated use of any single conventional component. Specifically, the first contribution is the design of a hierarchical agronomic representation pipeline that combines nonlinear latent compression through AFC with post-compression statistical disentanglement via ICA, thereby reducing dimensional redundancy while preserving yield-relevant latent structure. The second contribution is the introduction of two adaptive higher-order agronomic descriptors, namely adaptive weighted skewness and adaptive weighted kurtosis, which are formulated to amplify temporally important asymmetric and heavy-tailed events such as abrupt moisture depletion, heat stress, or rainfall anomalies that are often underrepresented by standard statistical descriptors.

The third contribution is the development of the Bio-Inspired Adaptive Nutrient Optimization (BIANO) algorithm as a dual-purpose optimization mechanism that jointly performs feature subset selection and model hyperparameter adaptation using a nutrient-inspired exploration–exploitation strategy. Unlike static metaheuristics that optimize only a single objective or stage, BIANO is embedded within the learning loop and updates candidate solutions according to predictive fitness, compactness, and generalization stability. The fourth contribution is the construction of a TA-CRN that operates on the BIANO-optimized feature space to simultaneously capture local temporal patterns, long-range sequential dependencies, and growth-stage-specific yield sensitivity. Therefore, the core novelty of the manuscript is best understood as a co-optimized end-to-end forecasting architecture in which representation learning, agronomic feature engineering, adaptive optimization, and temporal attention are not independent modules, but mutually reinforcing stages that improve predictive reliability under noisy real-time agricultural IoT conditions.

3. METHODOLOGY

3.1 Proposed methodology

The suggested methodology in Figure 1 offers both integrated IoT-based, data-centric deep learning for the correct prediction of crop products through real-time agricultural monitoring, sophisticated data preprocessing, optimal feature learning, and a TA-CRN architecture. The constant data gathering is carried out by heterogeneous IoT sensor nodes that are installed on agricultural fields to measure soil physicochemical characteristics, microclimatic changes, and crop growth indices at small temporal scales. The data obtained due to environmental disturbances, sensor drift, and communication are high-dimensional, noisy, asynchronous, and nonlinear in nature. To overcome such difficulties, the multi-stage preprocessing pipeline is organized, consisting of noise filtering, missing-value imputation, normalization, standardization, and temporal synchronization. Then, a two-step dimensionality reduction plan that incorporates the AFC and ICA is adopted to obtain small nonlinear latent representations and statistically independent components, in effect, minimizing redundancy with salient agronomic patterns.

After dimensionality reduction, multi-level feature

extraction is done on traditional statistical descriptors, higher-order statistical measures, and two adaptive features, such as adaptive weighted skewness and adaptive weighted kurtosis, to be more sensitive to distributional asymmetry, tail behavior, and yield-critical extreme events. To achieve higher predictive performance and learning efficiency, a BIANO algorithm is adopted to select optimal feature subsets and dynamically tune hyperparameters, yielding compact, discriminative feature

sets. The optimized features are then input into the TA-CRN model, which incorporates the use of convolutional layers to extract local patterns, recurrent units to model the temporal dependencies, and an attention mechanism to focus on the yield-critical growth stages and time steps. The framework is also trained and validated using standard performance metrics, is scalable and generalizable, and is relevant to real-time precision agriculture applications.

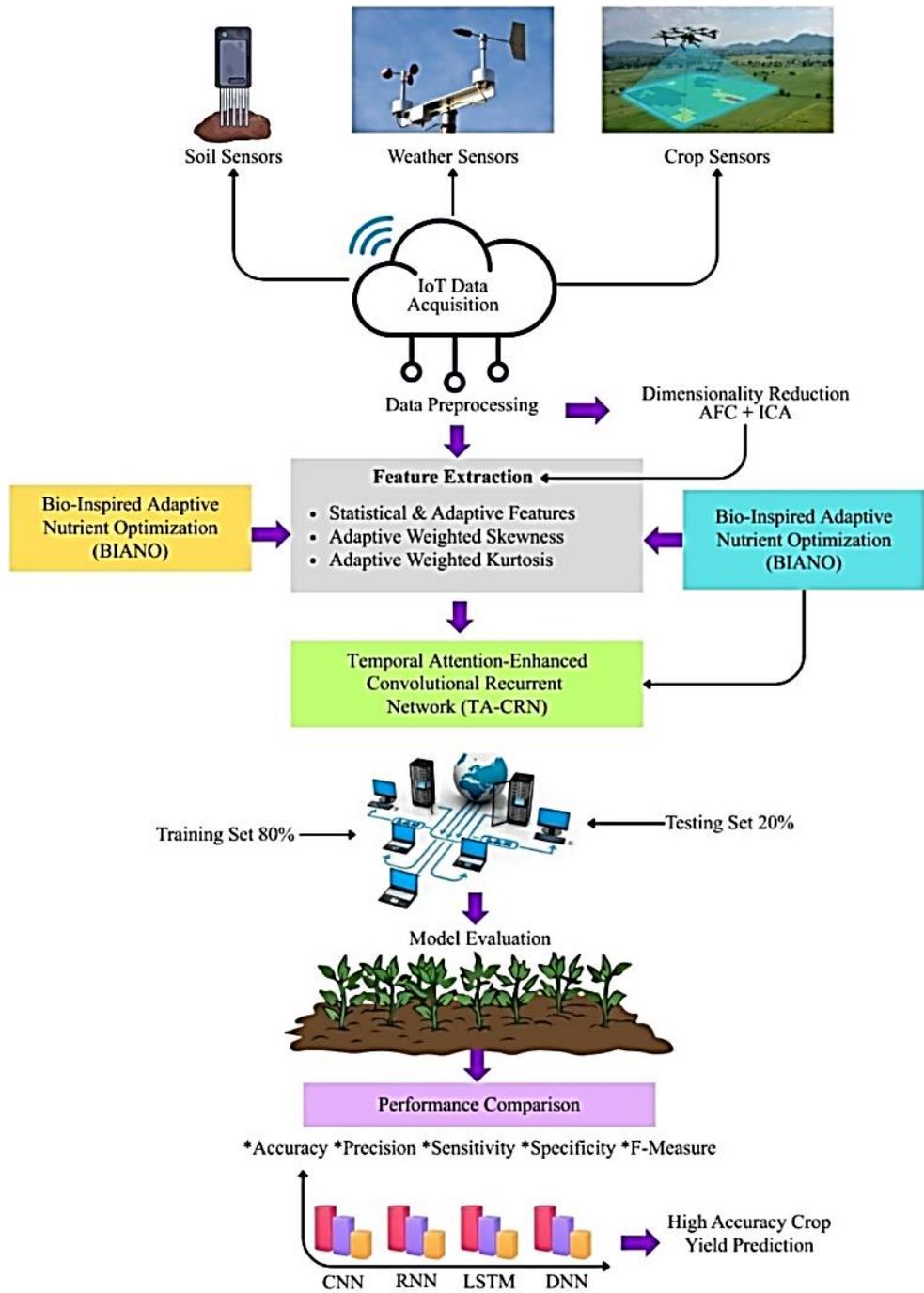


Figure 1. Proposed architecture

3.2 Data collection

The data collection phase aimed to have high-resolution and season-long monitoring of crop development in a wide-ranging environment of actual agricultural fields, as given in Table 1. The experiment was done on rice (*Oryza sativa*) as it is a staple cereal crop that is highly sensitive to soil moisture,

variation in temperature, and changes in climate, and therefore a perfect subject to explore the potential of advanced IoT-based predictive models. The experimental agricultural field in the state of Tamil Nadu, India, was used to gather data, which was a tropical area and had heavy practices of rice growing. The entire crop growth was monitored, which was the Kharif season, June to November, which contains such

important phenological stages as the development of the seedlings, tillering, the development of the panicle, flowering, and maturity. The heterogeneous dataset was formed by combining real-time measurements of the IoT sensors, weather station measurements, and agronomic records at the field level. Parameters of the soil were constantly measured at different depths to measure the root-zone dynamics, whereas parameters on the atmosphere and crop-level were measured to explain the external environmental stressors and physiological responses. To have ground-truth points of reference to model training and validation, besides sensor-obtained data, historical yield data, and agronomic logs kept by farm authorities were also included. The gathered dataset therefore forms a multi-source, multi-temporal, and multi-dimensional agricultural data warehouse, which can strongly learn nonlinear interactions between soil, weather, crop development, and yield performance.

It is important to note that the present dataset was intentionally constructed as a high-resolution single-site longitudinal field dataset rather than a broad multi-region benchmark. The primary objective of this study was to evaluate the methodological effectiveness of the proposed TA-CRN-BIANO framework under tightly monitored real-time

agronomic conditions, where sensor synchronization, temporal continuity, and label reliability could be maintained across the full crop lifecycle. Accordingly, the dataset covers one complete Kharif-season rice cultivation cycle from a single irrigated lowland paddy environment in Tamil Nadu, India, and should therefore be interpreted as a controlled proof-of-concept deployment dataset rather than a universally representative agroecological benchmark.

3.3 IoT-based agricultural data acquisition

An IoT sensor network cell was distributed over the paddy field, and its placement was strategic in order to record the soil, atmospheric, and canopy conditions of crops. Soil sensors were also placed in root-zone depths to detect moisture availability, temperature change, pH balance, and electrical conductivity, all of which directly affect nutrient uptake and plant well-being. At the same time, weather stations planted around the field took atmospheric parameters that included the air temperature, relative humidity, rainfall, wind speed, and solar radiation, which guaranteed proper depiction of microclimatic conditions that influence crop growth (Table 2).

Table 1. Detailed description of the crop yield dataset used in the study

Attribute Category	Description
Crop Type	Rice (<i>Oryza sativa</i>)
Cultivation Season	Kharif season
Monitoring Period	June to November (6 months)
Geographic Location	Tamil Nadu, India (Latitude ~11–13°N, Longitude ~77–79°E)
Field Type	Irrigated lowland paddy field
Soil Type	Clay loam
Growth Stages Covered	Seedling, Tillering, Panicle initiation, Flowering, Maturity
Data Collection Mode	Real-time IoT sensing + farm records
Temporal Resolution	30 minutes daily (parameter dependent)
Spatial Resolution	Plot-level sensing (sensor nodes per field segment)
Yield Reference	Seasonal harvested yield (kg/ha)
Total Duration of Dataset	One complete crop cycle
Purpose of Data	Model training, validation, and yield forecasting

Table 2. IoT sensors, parameters, and data acquisition specifications

Sensor Category	Parameter Monitored	Sensor Placement	Data Acquisition Frequency	Typical Observed Range
Soil Sensor	Soil moisture	Root zone (10–30 cm)	Every 30 minutes	18–45
Soil Sensor	Soil temperature	Subsurface	Every 30 minutes	22–35
Soil Sensor	Soil pH	Subsurface	Daily	5.5–7.8
Soil Sensor	Electrical conductivity	Root zone	Daily	0.2–1.5
Weather Sensor	Air temperature	Weather station	Hourly	24–38
Weather Sensor	Relative humidity	Weather station	Hourly	55–95
Weather Sensor	Rainfall	Field boundary	Event-based	0–120
Weather Sensor	Wind speed	Weather station	Hourly	0–8
Radiation Sensor	Solar radiation	Weather station	Hourly	200–900
Crop Sensor	Leaf wetness	Canopy level	Daily	0–100
Crop Sensor	Canopy temperature	Crop surface	Daily	25–36
Agronomic Record	Crop yield	Harvest stage	Seasonal	3,500–6,800

All the sensors were linked by the low-power wireless communication standards, which allowed the possibility of easy data transfer to an IoT gateway and a cloud-based storage system. The sensor readings were time-stamped in order to allow the time correspondence between non-homogeneous data streams. The system of acquisition was designed to be in effect during the whole season of cropping and detect both slow tendencies and sudden changes in the environment, like heat stress or high rainfall. This high-resolution sensing infrastructure guaranteed the presence of high-quality

temporal sequences needed for deep learning-based temporal modellings based on TA-CRN structures.

3.4 Data preprocessing

The methodological components used in the proposed framework include several well-established learning primitives, namely autoencoder-based latent compression, independent component decomposition, convolutional recurrent temporal modeling, and attention-based sequence

weighting. Since these components are already well known in the deep learning literature, the present section intentionally emphasizes their task-specific integration rather than reiterating generic theoretical foundations. The principal methodological contribution of this work lies not in the isolated use of these standard modules, but in their ordered co-optimization for agricultural IoT data, where nonlinear compression, redundancy disentanglement, adaptive higher-order agronomic descriptors, and nutrient-inspired feature-hyperparameter search are jointly coupled before temporal classification. Accordingly, the following subsections focus primarily on the operational role, inter-module dependency, and contribution of each block within the proposed TA-CRN-BIANO pipeline.

Preprocessing is thus necessary to improve the quality of data and generate a consistent learning behavior among deep neural models like TA-CRN. Within the framework proposed, preprocessing of data will be done in a multi-staged pipeline of data noise removal, integration of missing values, data normalization, data standardization, and temporal alignment. This pipeline provides that raw sensor signals can be converted to clean, consistent, and temporally coherent sequences that can be used in temporal-based and recurrent neural learning. Predictive accuracy is not only enhanced by the preprocessing stage but also converges faster and reduces overfitting when training a model. Mathematically, let the raw IoT dataset be represented as (Eq. (1))

$$\mathbf{X} = \{x_{i,j}(t)\}, i = 1, 2, \dots, N, j = 1, 2, \dots, M \quad (1)$$

where, $x_{i,j}(t)$ denotes the reading of the j -The sensor parameter at time t for the i -the observation, N is the number of samples, and M is the number of monitored parameters.

3.4.1 Data cleaning and noise removal

Noise in agricultural sensor data arises from environmental interference, sensor aging, calibration drift, and abrupt weather events. To address this, a hybrid noise filtering strategy is employed, combining statistical thresholding and smoothing-based techniques. Firstly, an interquartile range (IQR) criterion is used to identify outliers, and the criterion is (Eq. (2)):

$$\begin{aligned} \text{IQR} &= Q_3 - Q_1 \\ x \in \text{Outlier if } x < Q_1 - 1.5(\text{IQR}) \text{ or } x > Q_3 + 1.5(\text{IQR}) \end{aligned} \quad (2)$$

The outliers detected are fixed via local interpolation or eliminated when it is physically implausible. Then, an average filter is used to eliminate high noise in (Eq. (3)):

$$\hat{x}(t) = \frac{1}{w} \sum_{k=t-w+1}^t x(k) \quad (3)$$

where, w denotes the size of the window. This is done to make sure there is a smooth transition across time, and the agronomic trends that are meaningful (e.g., progressive drying of soil, increasing temperature) are maintained.

3.4.2 Handling missing and inconsistent sensor data

Missing values are caused by loss of packets, malfunctioning of sensors, or loss of power. In order to provide continuity of data, missing observations are filled in

by a hybrid temporal interpolation method. In short discontinuities, linear interpolation is used in (Eq. (4)):

$$x(t) = x(t_1) + \frac{t - t_1}{t_2 - t_1} [x(t_2) - x(t_1)] \quad (4)$$

where, t_1 and t_2 denote the nearest valid timestamps. For longer gaps, a mean-based temporal imputation is employed in (Eq. (5)):

$$x(t) = \frac{1}{K} \sum_{k=1}^K x(t - k) \quad (5)$$

Inconsistent values that do not meet physical constraints (e.g., negative rainfall values or extreme pH values) are detected by use of domain-specific rules and rectified. This is the strategy that does not create artificial bias in the dataset completeness.

3.4.3 Data normalization and standardization

Because of the differences in scale of sensor parameters, normalization and standardization are vital in avoiding feature dominance in the case of neural learning. The first step is min-max normalization of bounded values within a given range in (Eq. (6)):

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (6)$$

Then, the standardization of the Z-score is performed to have a zero mean and unit variance in (Eq. (7)):

$$x_{\text{std}} = \frac{x - \mu}{\sigma} \quad (7)$$

where, μ and σ represent the mean and standard deviation, respectively. This bilateral change makes gradient propagation stable and increases the consistency of firing of temporal neurons of the TA.

3.4.4 Temporal alignment of multi-source IoT data

Temporal alignment is needed to coordinate heterogeneous data streams since the various sensors have different sampling frequencies. There is a popular time grid T_c is which is defined, and all sensor readings are resampled with interpolation (Eq. (8)):

$$x_j(T_c) = \mathcal{J}(x_j(t)) \quad (8)$$

where, $\mathcal{J}(\cdot)$ denotes interpolation. This step ensures that each feature vector is considered at a time. T_c represents a coherent agronomic state, enabling accurate temporal dependency modeling in the CRN layer.

3.5 Two-fold dimensionality reduction

In order to handle these issues, the two-fold dimensionality reduction approach is taken using the AFC with the CA. This is a hybrid method using the nonlinear representation learning ability of autoencoders and the statistical independence modeling ability of ICA, creating small, informative, and decorrelated feature representations that can be used in deep temporal learning.

Let the preprocessed IoT dataset be represented as (Eq. (9))

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\} \in \mathbb{R}^{N \times d} \quad (9)$$

where, N denotes the number of samples and d represents the original feature dimensionality. The objective of dimensionality reduction is to transform \mathbf{X} into a low-dimensional latent space $\mathbf{Z} \in \mathbb{R}^{N \times d'}$, where $d' \ll d$, while preserving agronomically relevant information.

3.5.1 Autoencoder-based Feature Compression (AFC)

AFC is employed as the first dimensionality reduction stage to capture nonlinear correlations among agricultural sensor features. An autoencoder consists of an encoder–decoder pair trained to reconstruct the input data with minimal error. The encoder maps the input vector \mathbf{x}_i into a low-dimensional latent representation \mathbf{h}_i as (Eqs. (10)-(12))

$$\mathbf{h}_i = f_{\text{enc}}(\mathbf{x}_i) = \sigma(\mathbf{W}_e \mathbf{x}_i + \mathbf{b}_e) \quad (10)$$

where, \mathbf{W}_e and \mathbf{b}_e denote encoder weights and biases, respectively, and $\sigma(\cdot)$ is a nonlinear activation function. The decoder reconstructs the input as

$$\hat{\mathbf{x}}_i = f_{\text{dec}}(\mathbf{h}_i) = \sigma(\mathbf{W}_d \mathbf{h}_i + \mathbf{b}_d) \quad (11)$$

where, \mathbf{W}_d and \mathbf{b}_d are decoder parameters. The autoencoder is trained by minimizing the reconstruction loss.

$$\mathcal{L}_{\text{AFC}} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 \quad (12)$$

Through this optimization, AFC learns compact nonlinear latent features that preserve essential agronomic patterns such as soil–climate interactions and crop growth trends, while significantly reducing dimensionality.

3.5.2 Independent component analysis (ICA)

The ICA is a technique in machine learning. Even though AFC is effective at compressing data, it is possible that the latent features still have statistical dependencies. To further bring out features, the latent space generated by AFC is subjected to ICA. The output of AFC will be represented by \mathbf{Rf} in (Eqs. (13)-(16)):

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N]^T \quad (13)$$

ICA presupposes that its linear combination of statistically independent source components \mathbf{S} :

$$\mathbf{H} = \mathbf{A}\mathbf{S} \quad (14)$$

where, \mathbf{A} is a mixing matrix, unknown. ICA is aimed at estimating an unmixing matrix \mathbf{W} , where

$$\mathbf{S} = \mathbf{W}\mathbf{H} \quad (15)$$

and the elements of \mathbf{S} are as independent as possible. This is done by maximizing non-Gaussianity, which is commonly quantified by negentropy:

$$J(\mathbf{s}) = H(\mathbf{s}_{\text{gauss}}) - H(\mathbf{s}) \quad (16)$$

where, $H(\cdot)$ denotes entropy and $\mathbf{s}_{\text{gauss}}$ is a Gaussian random

variable with the same covariance. ICA can effectively separate independent agronomic factors (moisture stress, temperature variability, nutrient dynamics, et cetera) and increase the interpretability and robustness.

3.5.3 Integrated AFC–ICA dimensionality reduction strategy

The final reduced feature representation is obtained by sequentially applying AFC followed by ICA (Eq. (17)):

$$\mathbf{Z} = \text{ICA}(\text{AFC}(\mathbf{X})) \quad (17)$$

3.6 Feature extraction

The importance of feature extraction is to convert dimension-reduced agricultural IoT data into informative representations that can well characterize crop growth dynamics and yield variability. Raw sensor signals can be insensitive to distributional properties, temporal variability, or nonlinear agronomic responses, even when they are reduced in dimensionality. Hence, based on the proposed framework, multi-level feature extraction is applied to obtain first-order trends, dispersion trends, higher-order statistical processes, and extreme-event sensitivities, which soil-crop-climate relationships are supposed to exhibit. This stratified feature learning approach will make sure that not only stable growth trends are covered, but also, seldom, yet significant stressors (e.g., drought spells or heat waves) are also present. The reduced feature vector at time t may be denoted as $\mathbf{z}(t)$, i.e., as $z_1(t), z_2(t), \dots, z_d(t)$, where d is the reduced dimensionality achieved by the combined AFC-ICA step. The extraction of the features is done across sliding temporal windows to maintain agronomic continuity and time context.

3.6.1 Statistical feature extraction

The statistical characteristics are obtained to describe the attributes of central tendency and dispersion of agricultural sensor signals with time. These characteristics offer a condensed overview of the main tendencies in the change of soil moisture level, temperature changes, and crop physiology, which are highly linked with the formation of yields. The average value of a feature over a time window of length T is calculated as (Eqs. (18)-(21)):

$$\mu = \frac{1}{T} \sum_{t=1}^T z(t) \quad (18)$$

which represents the average agronomic condition over the observation period. Variability around this average is captured using variance, defined as

$$\sigma^2 = \frac{1}{T} \sum_{t=1}^T (z(t) - \mu)^2 \quad (19)$$

These measures are essential for identifying stable versus volatile environmental conditions. Additionally, the standard deviation

$$\sigma = \sqrt{\sigma^2} \quad (20)$$

Allows for obtaining scale-sensitive dispersion measures affecting temporal neuron firing stability in the TA layer. These statistically descriptive features allow us to construct the fundamental feature set that describes the baseline behavior of crop growth under different environmental

conditions. In addition to dispersion, the energy of the signal is calculated to give cumulative agronomic intensity:

$$E = \sum_{t=1}^T z^2(t) \quad (21)$$

which is specifically applied in the capture of long-term stress situations like long-duration high temperature or constant lack of moisture. These statistical properties guarantee that the representation extracted retains both magnitude-based and variability-based information that is required in yield forecasting. Higher-order statistical features are extracted using t-statistical methods with application to the responses

3.6.2 Higher-order statistical feature extraction

The responses are subjected to statistical methods of extracting higher-order statistical features. Although both first and second order statistics explain the overall behavior and dispersion, they do not have the capacity to explain the asymmetry and tail behavior of the agricultural sensor distribution. Nonlinear and extreme events often affect crop yield, i.e., sudden increases of rainfall or short-term heat stress, and thus demand higher-order statistical measures. The first computed skewness is to estimate the asymmetry of the distribution (Eqs. (22) and (23)):

$$\text{Skewness} = \frac{1}{T} \sum_{t=1}^T \left(\frac{z(t)-\mu}{\sigma} \right)^3 \quad (22)$$

indicating whether sensor readings are biased toward lower or higher values. Similarly, kurtosis quantifies the peakedness and tail heaviness of the distribution:

$$\text{Kurtosis} = \frac{1}{T} \sum_{t=1}^T \left(\frac{z(t)-\mu}{\sigma} \right)^4 \quad (23)$$

These higher-order features are particularly effective in detecting abrupt agronomic stress conditions that may not significantly alter the mean but drastically affect yield outcomes.

3.6.3 Adaptive weighted skewness feature

Traditional skewness treats all observations equally, which may dilute the influence of critical growth-stage variations. To overcome this limitation, an adaptive weighted skewness feature is proposed, assigning higher importance to sensor observations occurring during yield-sensitive growth phases. Let $w(t)$ denote an adaptive weight derived from temporal relevance or sensor reliability. The adaptive weighted skewness is defined as (Eq. (24)):

$$\text{AWS} = \frac{\sum_{t=1}^T w(t) \left(\frac{z(t)-\mu}{\sigma} \right)^3}{\sum_{t=1}^T w(t)} \quad (24)$$

The weights are dynamically set to reflect the volatility of the environments and the agronomic significance, and ensure that the deviations in flowering or grain-filling phases have a stronger effect on the feature representation. This expression allows the model to highlight asymmetric behavior in times of critical behavior as opposed to averaging it in the whole season. Therefore, adaptive weighted skewness is much more sensitive to phase-specific patterns of stress that have a direct influence on the formation of the yield.

3.6.4 Adaptive weighted kurtosis feature

Similar to skewness, conventional kurtosis fails to account for the temporal significance of extreme observations. The proposed adaptive weighted kurtosis feature addresses this limitation by amplifying the contribution of rare but agronomically critical events. It is formulated as (Eq. (25)):

$$\text{AWK} = \frac{\sum_{t=1}^T w(t) \left(\frac{z(t)-\mu}{\sigma} \right)^4}{\sum_{t=1}^T w(t)} \quad (25)$$

This adaptive formulation guarantees that there are no sudden events in the temperature, unforeseen events of rainfall, or sudden depletion of soil moisture that will be repressed in the process of aggregation. The presence of high adaptive weighted kurtosis implies heavy-tailed distributions in the case of extreme stress conditions, which allows the TA-CRN model to acquire strong representations of yield-disruptive events.

3.7 Proposed technique: Temporal Attention-Enhanced Convolutional Recurrent Network

The nature of crop growth and yield formation is controlled by cumulative interactions between the environment, over longer time scales, and both short-term variations (varying rainfall or heat stress) and long-term ones (like soil moisture storage and seasonal climate variations) are important. Traditional deep learning architectures tend to be incapable of both local feature interactions and long-range temporal interactions. To overcome this weakness, the suggested framework combines convolutional feature extraction, recurrent temporal modeling, and an attention mechanism into one framework given in Figure 2.

The TA-CRN works in series with the optimized feature vectors acquired following BIANO-based feature selection. Convolutional layers then identify spatial and local time patterns in the multivariate sequences of IoT sensors, which represent a response to an interaction of soil, weather, and crop variables. Recurrent units are then used to process the extracted feature maps to predict temporal dependencies across the stages of crop growth. Lastly, a temporal attention process dynamically assigns more weight to yield-relevant time steps so that the network concentrates on the influential agronomic events and gives less priority to irrelevant fluctuations. The integrated design enables the model to be adaptive to a change in the environment without forgetting contextual memory throughout the entire cropping season.

3.7.1 End-to-end operational workflow of the proposed TA-CRN-BIANO framework

The overall operational sequence of the suggested framework is planned as a highly linear and dependency-based pipeline where each module is going to prepare the input representation that is needed by the next one given in Figure 3. To begin with, the heterogeneous agricultural measurements are obtained using the IoT-enabled sensory and field data that cover soil moisture, soil temperature, ambient humidity, air temperature, rainfall, and crop-condition descriptors. Such raw multivariate streams are subsequently processed by harmonizing data quality, such as missing-value imputation, data outliers, data normalization, and time alignment to convert sensor readings at various times into synchronized time-based streams. This step generates a single agricultural

data tensor where a single value set of environmental and crop-state conditions is available at each time. AFC is then used to compress the synchronized sensor matrix to a small nonlinear latent representation. AFC is to reduce redundancy by maintaining higher-order inter-variable dependencies, which are usually completely lost by linear reduction methods on their own. The resultant compressed latent vectors are, in turn, fed to ICA, which further decomposes statistically independent latent sources and minimizes residual correlation between compressed components. That is, nonlinear manifold compaction is done by AFC followed by post-compression

source disentangling and decorrelation by ICA. The AFC-ICA cascade is not random, and instead, there is deliberate staging in the way nonlinear information preservation is succeeded by the enhancement of statistical independence. Of such refined latent components, statistical descriptors and higher-order descriptors are obtained, such as the mean, variance, entropy, adaptive weighted skewness, and adaptive weighted kurtosis, thus giving rise to a feature space, which captures agronomic fluctuations of central tendency as well as asymmetric extremity.

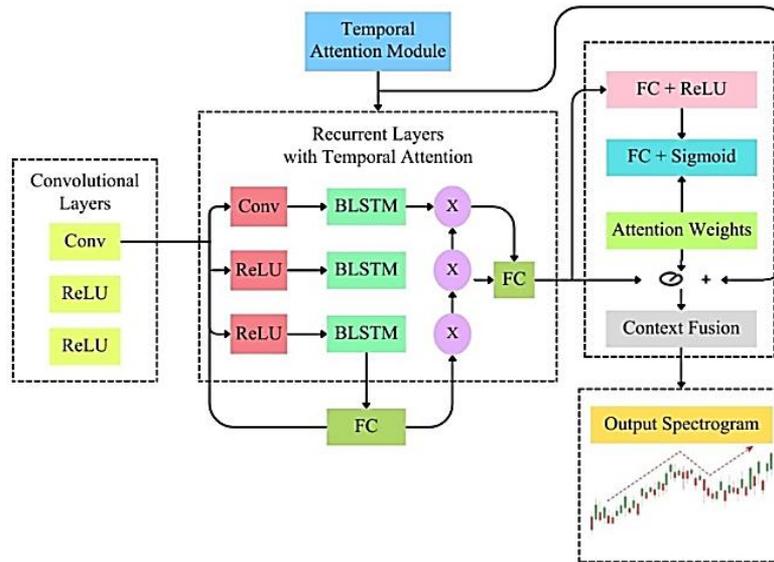


Figure 2. Proposed technique

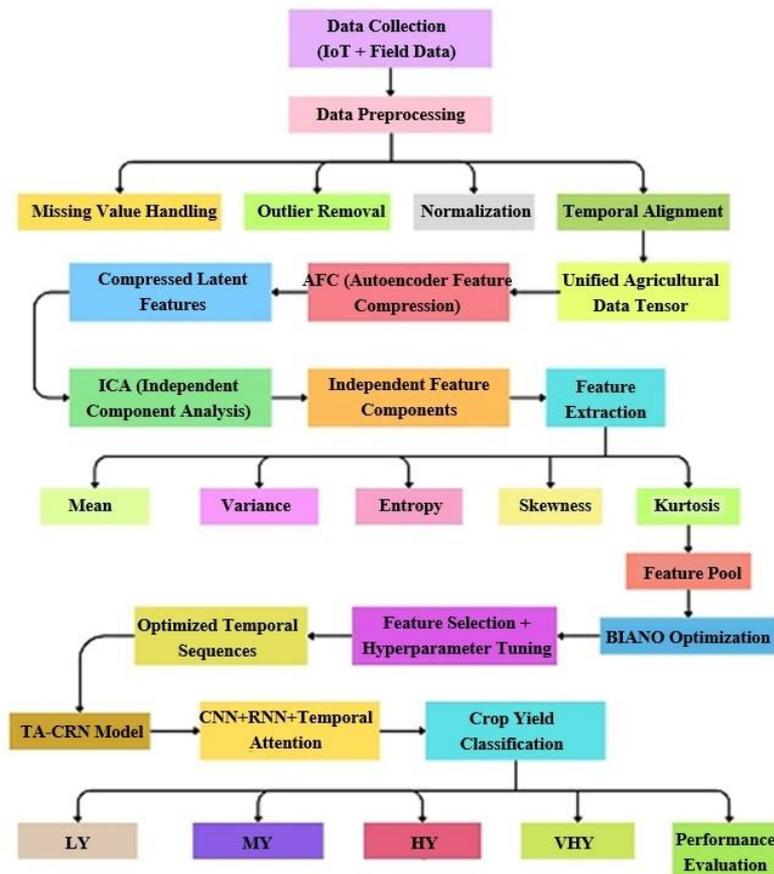


Figure 3. Workflow-oriented schematic diagram

The obtained feature pool is then subjected to the proposed BIANO algorithm that does joint feature subset selection and TA-CRN hyperparameter optimization. At this point, the subsets of candidate features and parameter configurations are viewed as developing nutrient-carrying objects, and their fitness is measured based on classification accuracy, compression, and computational efficiency. The best configuration identified by BIANO is then applied to produce temporally ordered multivariate sequences in line with the crop development. These optimized sequences are eventually passed to the TA-CRN classifier, which consists of convolutional layers that learn local short-term temporal interactions, recurrent units that learn long-range agronomic dependencies, and temporal attention that enhances yield-critical phenological intervals. The network gives an end result of a four-class crop-yield categorization (LY, MY, HY, VHY), assessed by performance measures of class-based and macro-based.

3.7.2 Convolutional feature extraction layer

An optimized sequence of input features can be written as

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \quad (26)$$

where, T represents the number of temporal observations and $\mathbf{x}_t \in \mathbb{R}^{d'}$ denotes the reduced feature vector at time t . A one-dimensional convolution operation is applied to extract local temporal patterns:

$$\mathbf{c}_t = f\left(\sum_{k=0}^{K-1} \mathbf{W}_k \mathbf{x}_{t-k} + \mathbf{b}\right) \quad (27)$$

where, \mathbf{W}_k represents convolutional kernel weights of size K , \mathbf{b} is the bias vector, and $f(\cdot)$ denotes a nonlinear activation function. This convolutional operation captures short-term dependencies and interactions among agronomic variables, producing discriminative feature maps resilient to sensor noise.

3.7.3 Recurrent temporal dependency modeling

The convolutional feature maps are passed to a recurrent layer in order to capture long-term temporal dependencies during the stages of crop growth. Let c_t denote the convoluted value at time t . The repeated concealed condition is calculated as:

$$\mathbf{h}_t = \phi(\mathbf{W}_h \mathbf{c}_t + \mathbf{U}_h \mathbf{h}_{t-1} + \mathbf{b}_h) \quad (28)$$

where, \mathbf{W}_h and \mathbf{U}_h are learnable weight matrices, \mathbf{b}_h is the bias vector, and $\phi(\cdot)$ is a nonlinear activation function. This formulation allows the network to preserve the important historical data and also adjust to slow environmental changes, which could be in vegetative and reproductive phases of growth.

3.7.4 Temporal attention mechanism

Although the temporal information is represented by the recurrent units, all the time steps are not equally valuable in the prediction of the yield. The yield critical periods are brought in by the introduction of a temporal attention mechanism.

$$e_t = \mathbf{v}^T \tanh(\mathbf{W}_a \mathbf{h}_t + \mathbf{b}_a) \quad (29)$$

where, \mathbf{W}_a , \mathbf{b}_a , and \mathbf{v} are learnable attention parameters. The

normalized attention weight α_t is obtained using a softmax function:

$$\alpha_t = \frac{\exp(e_t)}{\sum_{i=1}^T \exp(e_i)} \quad (30)$$

Context vector \mathbf{c} , which is the weighted temporal summary, is estimated as:

$$\mathbf{c} = \sum_{t=1}^T \alpha_t \mathbf{h}_t \quad (31)$$

The attention-based aggregation enables the model to focus dynamically on important developmental stages of the crop, such as flowering and grain filling, which have a significant contribution to overall crop production. The final result of the crop y is obtained as a linear transformed version of the attention-weighted version of the context vector:

$$\hat{y} = \mathbf{W}_o \mathbf{c} + \mathbf{b}_o, \quad (32)$$

where, \mathbf{W}_o and \mathbf{b}_o denote output layer parameters. The network parameters are optimized using gradient-based learning to minimize prediction error while ensuring stable convergence.

To improve sensitivity to temporally localized agronomic stress events, two adaptive higher-order descriptors are introduced in the feature extraction stage: adaptive weighted skewness (AWS) and adaptive weighted kurtosis (AWK). Unlike conventional skewness and kurtosis, which treat all temporal observations equally, the proposed descriptors assign greater importance to time steps that are agronomically more informative for yield formation. Let a temporal variable sequence for a given sensing channel be denoted by x_t , $t = 1 \dots, T$, and let w_t represent the adaptive importance weight associated with x_t , $w_t \geq 0$ and $\sum_{t=1}^T w_t = 1$. The weighted mean is defined as $\mu_w = \sum_{t=1}^T w_t x_t$, and the weighted standard deviation is $\sigma_w = \sqrt{\sum_{t=1}^T w_t (x_t - \mu_w)^2}$.

Based on these quantities, the adaptive weighted skewness is computed as $AWS = \frac{\sum_{t=1}^T w_t (x_t - \mu_w)^3}{\sigma_w^3 + \epsilon}$, and the adaptive weighted kurtosis is computed as $AWK = \frac{\sum_{t=1}^T w_t (x_t - \mu_w)^4}{\sigma_w^4 + \epsilon}$,

where ϵ is a small stabilizing constant to avoid numerical instability. The adaptive weights are not manually assigned; rather, they are derived from agronomic relevance signals estimated from temporal variability and yield sensitivity. Specifically, the weight at each time step is proportional to the normalized product of (i) local temporal deviation magnitude, (ii) growth-stage importance coefficient, and (iii) feature-yield association strength estimated on the training partition. Thus, $w_t \propto \alpha_t \cdot |x_t - \mu| \cdot \rho_t$, where α_t is a stage-priority coefficient and ρ_t represents normalized correlation-based yield relevance. This design allows the proposed descriptors to emphasize stress-sensitive intervals such as flowering and grain-filling phases, where asymmetric or heavy-tailed fluctuations are often more predictive of final yield outcomes than uniform temporal statistics.

The suggested TA-CRN model is a synergistic convolutional feature extraction, recurrent temporal modelling and attention-based weighting that are designed to capture short-term agronomic variations and long-term dependence of crop growth. The TA-CRN framework focuses on yield-sensitive time steps and reduces irrelevant variability in a

dynamic manner, improving predictive performance, sensitivity to noisy IoT measurements, and interpretability. This renders the proposed model completely applicable in real-time precision agriculture systems of scalable, reliable, and data-efficient crop yield forecasting.

3.8 Feature selection and hyperparameter adaptation using Bio-Inspired Adaptive Nutrient Optimization

In the context of IoT-based crop-yield prediction, feature optimisation is one of the critical steps, especially since despite a dimensionality-reduction step, the latent variables that have survived might still contain partially redundant, weakly discriminative, or time-varying features that eventually degrade the performance of generalisation. To address this problem, the suggested framework will include a BIANO algorithm which is a population-based meta-heuristic inspired by the nutrient acquisition, allocation, depletion, and adaptation mechanisms that can be observed in biological ecosystems and in plant-soil interactions. Within the abstraction used, each candidate solution is represented as a virtual nutrient organism that exists in two coupled states: (i) a binary feature-selection mask that determines the set of retained features, and (ii) a continuous hyper-parameter vector that parameters the TA-CRN architecture (e.g. learning rate, number of convolutional filters, number of recurrent hidden units, dropout, dimensionality of attention). The biological analogy here is that biological organisms that are rich in nutrients are the candidate solutions that are able to garner more predictive value out of the available agronomic environment, and those species that are nutrient-poor are out of competitive relevance and are progressively eliminated. In this way, BIANO can combine feature-subset refinement and hyper-parameter adaptation into a single optimisation loop, as opposed to considering them as separate processes.

Let the i -th candidate at iteration t be represented as $X_i^{(t)} = \{M_i^{(t)}, H_i^{(t)}\}$, where $M_i^{(t)} \in \{0,1\}^d$ denotes the binary mask over d extracted features and $H_i^{(t)}$ denotes the continuous hyperparameter vector. The nutrient state $N_i^{(t)}$ quantifies the adaptive fitness memory of the candidate and is updated according to its predictive effectiveness. For each candidate, the TA-CRN is trained on the selected feature subset and evaluated on the validation set. A multi-objective fitness function is then computed to simultaneously maximize predictive quality and minimize unnecessary complexity:

$$\mathcal{F}_i^{(t)} = \alpha \cdot \text{Acc}_i + \beta \cdot F1_i + \gamma \cdot \text{Spec}_i - \lambda \cdot \frac{|M_i|}{d} - \mu \cdot \frac{\Theta_i}{\Theta_{\max}} \quad (33)$$

where, Acc_i , $F1_i$, and Spec_i denote validation accuracy, F-measure, and specificity, respectively; $|M_i|/d$ penalizes excessive feature retention; Θ_i is the parameter count of the instantiated TA-CRN model; and $\alpha, \beta, \gamma, \lambda, \mu$ are weighting coefficients satisfying $\alpha + \beta + \gamma = 1$ for the predictive component. This formulation is intentionally designed to avoid the common failure mode in metaheuristic model tuning, where accuracy improves only by increasing feature dimensionality or network size.

The nutrient state is updated using a gain–decay adaptation rule that balances exploitation of high-performing solutions with controlled retention of diversity:

$$N_i^{(t+1)} = (1 - \rho)N_i^{(t)} + \eta \cdot \max(0, \mathcal{F}_i^{(t)} - \bar{\mathcal{F}}^{(t)}) - \delta \cdot \max(0, \bar{\mathcal{F}}^{(t)} - \mathcal{F}_i^{(t)}) \quad (34)$$

where, ρ is the memory decay coefficient, η is the nutrient gain factor, δ is the nutrient depletion factor, and $\bar{\mathcal{F}}^{(t)}$ is the population-average fitness at iteration t . Those candidates with a performance above average are gaining nutrient deposits and, consequently, a greater chance of passing their segments of features and hyperparameters to the following generation.

On the other hand, weak performers are deprived of nutrients, which makes them more susceptible to substitution, mutation, or structural simplification. Such a nutrient-memory mechanism is what distinguishes BIANO as compared to traditional fixed meta-heuristics in that it incorporates a historical adaptive bias as opposed to using instantaneous fitness rankings. BIANO uses three operations inspired by biology to evolve the population: nutrient-guided assimilation, competitive pruning, and adaptive mutation. During nutrient-guided assimilation, feature indices and hyper-parameter components of elite candidates are inherited by sub-optimal candidates with a probability distribution proportional to normalised nutrient strength, further exacerbating exploration at high-fitness local optima.

In competitive pruning, low-contribution features are eliminated with a probability that is negatively proportional to their prior selection frequency and local significance to the current candidate, and this tends to favor parsimonious representations. In adaptive mutation, the feature mask and hyperparameter vector are mutated inversely proportional to the existing population diversity; when this diversity is low, a mutation strength is amplified to leave local maxima, and when it is high, a mutation is suppressed to maintain convergence stability. Termination of the algorithm can occur when the maximum number of iterations or global best fitness is not improving with a small tolerance over a specified patience window.

The downstream training and evaluation of the final TA-CRN model are then realised using the resulting optimised solution. In the methodological perspective, BIANO has an advantage over traditional wrapper-based feature selection in that it explicitly co-optimises the prediction performance, compactness of representation, and complexity that is concerned with deployability. It is especially relevant in the context of the agricultural IoT systems, where sensor redundancy, non-stationarity of the environment, and the impossibility of deep models to be deployed in real-time may compromise otherwise precise models. BIANO can be used to trade off nutrient-memory adaptation with a joint discrete-continuous optimisation to yield a feature-parameter configuration, which can be more discriminative and more stable, as well as operationally suited to precision agriculture settings.

Algorithm: BIANO-Based Feature Selection for Crop Yield Forecasting

Input:

Dataset D with feature set $F = \{f_1, f_2, \dots, f_n\}$ and target yield Y

Population size N

Maximum iterations T_{\max}

Nutrient adaptation parameters: α (nutrient gain), β (nutrient decay)

Evaluation model: TA-CRN

Performance metric: RMSE / MAE / R^2

Output:

Optimal feature subset F_{best}

Step 1: Population Initialization

1.1 Generate N randi, failure subset, $F_i \subseteq F$

1.2 Assign an inii.al nutrient value $v_i \in [0,1]$ to each subset

1.3 Evaluate the fitness of each i , F_i using TA-CRN

Step 2: Initial Best Selection

2.1 Identify F_{best} with maximum fitness

Step 3: Adaptive Nutrient Optimization

For iteration $t = 1$ to T_{max} :

For each feature subset F_i :

3.1 Nutrient Update

$$v_i^{(t+1)} = \begin{cases} v_i^{(t)} + \alpha \cdot r \cdot (1 - v_i^{(t)}), & \text{if fitness improves} \\ v_i^{(t)} - \beta \cdot r \cdot v_i^{(t)}, & \text{otherwise} \end{cases}$$

where $r \sim U(0,1)$

3.2 Feature Evolution

Add features with a probability proportional to v_i

Remove features with probability inversely proportional to v_i

Introduce small random mutations for diversity

3.3 Evaluation

Train the TA-CRN using the updated F_i

Retain updated subset if fitness improves.

Step 4: Global Best Update

Update F_{best} If a superior subset is found

Step 5: Termination

Return F_{best} upon convergence or reaching T_{max}

based thresholds estimated exclusively from the training partition to avoid information leakage. This formulation is practically meaningful because field interventions are typically triggered by categorical risk or productivity zones rather than by exact scalar yield values alone, especially in real-time advisory systems involving irrigation scheduling, nutrient prioritization, and stress mitigation.

The TA-CRN model is trained in a supervised classification setting using these yield categories as ground-truth labels. Convolutional layers first capture short-range temporal interactions among multivariate IoT variables, recurrent units then model long-term growth dependencies across phenological stages, and the temporal attention mechanism dynamically amplifies yield-critical intervals such as reproductive and grain-filling phases. The network parameters are optimized by minimizing the categorical cross-entropy loss:

$$\mathcal{L}_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(\hat{y}_{ic}) \quad (35)$$

where, y_{ic} denotes the true one-hot encoded class label and \hat{y}_{ic} denotes the predicted posterior probability for class c . Adaptive learning-rate optimization, early stopping, dropout regularization, and weight decay are used to ensure stable convergence and improved generalization under noisy agricultural sensor conditions.

Validation is conducted on unseen data to quantify classification robustness and reduce the risk of overfitting. Since the task is explicitly framed as categorical yield prediction, the primary evaluation metrics include accuracy, precision, sensitivity (recall), specificity, F-measure, macro-F1, and confusion-matrix-based class separability. In addition, to preserve agronomic continuity with the original scalar endpoint, the manuscript may optionally report an auxiliary regression-style consistency analysis by assigning class centroids to predicted categories and computing approximate RMSE or MAE; however, the principal experimental results and all benchmark comparisons in this study are interpreted within the multi-class classification framework. Because the forecasting problem is formulated as a four-class agronomic yield categorization task (LY, MY, HY, and VHY), classification-oriented metrics are used as the primary evaluation criteria, and all metrics are reported in macro-averaged form where applicable to avoid dominance by majority classes (Table 3).

3.9 Model training and validation

Following BIANO-based feature optimization, the retained feature vectors are reorganized into temporally ordered multivariate sequences aligned with crop growth progression. Although the original agronomic endpoint is continuous harvested yield measured in kg/ha, the present study formulates the prediction problem as a multi-class yield categorization task to support operational decision-making in precision agriculture. Specifically, the continuous yield values are discretized into four agronomically interpretable categories—Low Yield (LY), Medium Yield (MY), High Yield (HY), and Very High Yield (VHY)—using percentile-

Table 3. Performance metrics analysis

Metric	Mathematical Expression	Description and Relevance
Accuracy	$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$	Measures the overall correctness of yield predictions by considering all correctly classified samples. High accuracy indicates effective modeling of combined soil, weather, and crop factors.
Precision	$\text{Precision} = \frac{TP}{TP + FP}$	Evaluates the reliability of positive yield predictions by reducing false alarms. High precision ensures agronomically meaningful decisions in irrigation and resource planning.
Sensitivity (Recall)	$\text{Sensitivity} = \frac{TP}{TP + FN}$	Assesses the model's ability to detect actual yield stress or high-yield conditions. High sensitivity supports early identification of critical agronomic risks.
Specificity	$\text{Specificity} = \frac{TN}{TN + FP}$	Measures the correct identification of normal yield conditions. High specificity prevents false stress alerts, improving trust in decision-support systems.
F-Measure	$\text{F-Measure} = \frac{2 \times (\text{Precision} \times \text{Sensitivity})}{\text{Precision} + \text{Sensitivity}}$	Provides a balanced evaluation of detection accuracy and completeness. Especially effective for imbalanced agricultural yield datasets.

4. RESULTS AND DISCUSSION

The experimental evaluation of the proposed TA-CRN with the BIANO framework was conducted to rigorously assess its predictive capability under real-time IoT-enabled agricultural conditions. The results are:

4.1 Performance evaluation of TA-CRN with BIANO under different train–test splits

The results of the implementation of the TA-CRN model with BIANO were assessed with various train-test splits, and the findings are presented in Figure 4. Accuracy of the model with the split of 60:40 was 93.82, but it rose steadily to 96.50, 95.12, and 96.18 with the splits of 70:30, 80:20, and 90:10,

respectively. On the same note, the accuracy improved to a high of 96.12 at a 60:40 split and a high of 95.91 at a 90:10 split. Sensitivity had a similar trend where the 80:20 division had 92.97% to 95.84 percent sensitivity, but there was slight decrease in sensitivity to 95.22% at the 90:10 division. The value of specificity was high throughout the splits, between the range of 94.21 and 96.73, with the maximum performance being recorded with the 80:20 split. This trend was also indicated by the F-measure, in that 60:40 had 93.21, followed by 80:20 with 95.98, then a slight fluctuation was observed in 90:10 at 95.56. On the whole, the TA-CRN with BIANO had strong predictive accuracy regardless of the train-test choice, and the 80:20 split provided the best predictive accuracy along with other measures of evaluation.

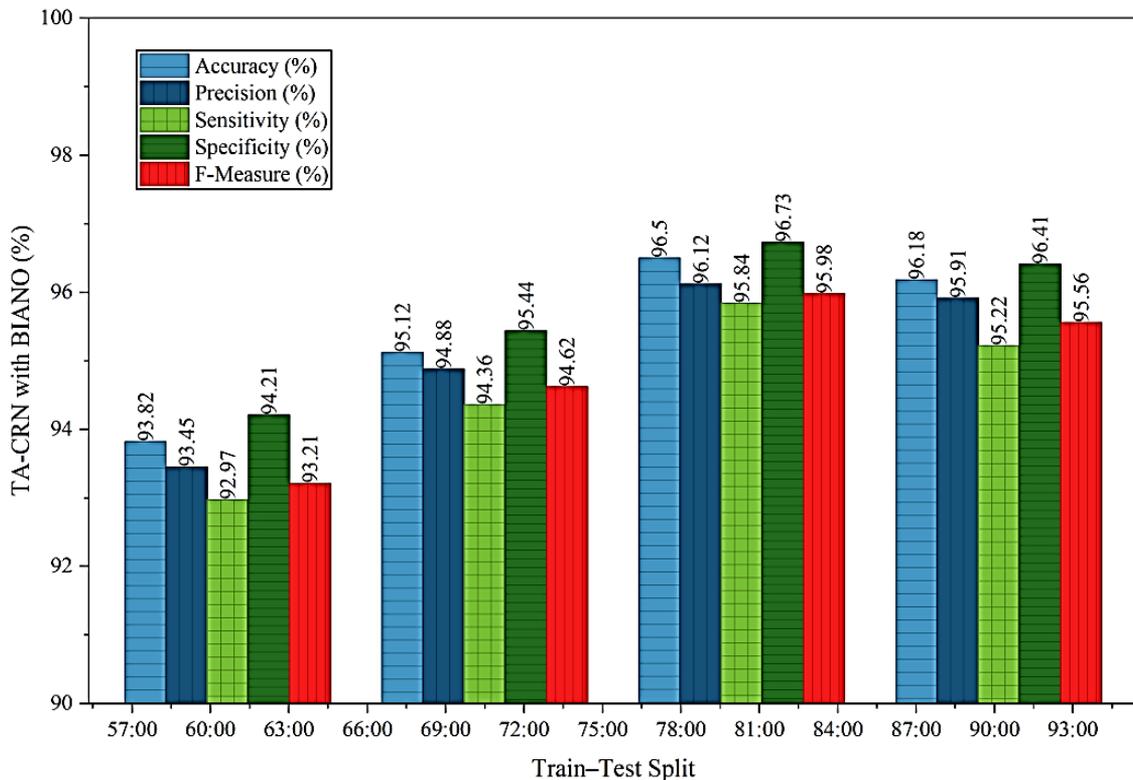


Figure 4. Performance of TA-CRN with BIANO under different train–test splits

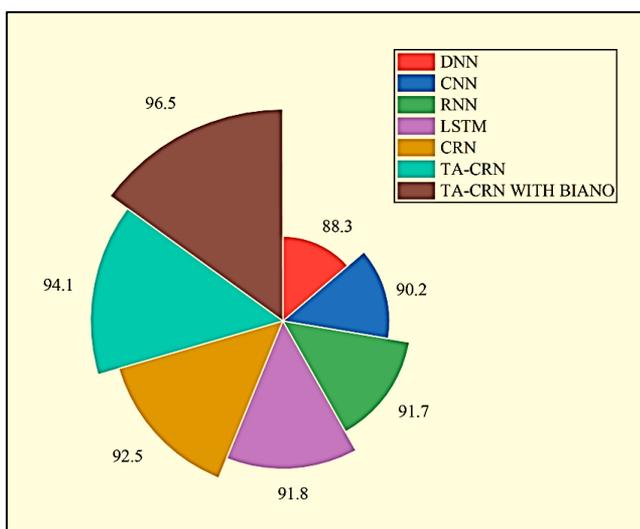


Figure 5. Accuracy comparison of TA-CRN with BIANO and existing models

4.2 Accuracy analysis

The comparative analysis of model accuracies is presented in Figure 5, highlighting the superior performance of the proposed TA-CRN with the BIANO approach. Among the existing models, DNN achieved the lowest accuracy of 88.34%, followed by CNN at 90.21%, RNN at 91.67%, LSTM at 91.83%, and CRN at 92.45%. The independent TA -CRN model achieved the best performance of 94.12, thus proving the efficacy of the temporal attention, combined with the convolutional recurrent integration. It is worth noting that the TA-CRN, when optimised with the BIANO, was the most accurate with 96.50%, and this is significantly high compared to all other models, so the use of the BIANO optimisation to support the predictive performance of the model proved to be effective.

4.3 Precision and sensitivity comparison across models

Table 4 presents a summary of the comparative analysis of

the precision and sensitivity of the different models. The lowest precision and sensitivity values of the DNN model were the lowest at 87.95% and 86.42%, respectively, and CNN rose to 89.88% and 88.73% precision and sensitivity, respectively. Sequential models such as RNN and LSTM demonstrated further improvements, achieving precision values of 90.74% and 91.36% and sensitivity values of 90.12% and 90.95%, respectively. The CRN model had a better precision of 92.08% and sensitivity of 91.64, which is the advantage of convolutional recurrent structures. The TA-CRN model had significant gains with a precision of 94.56 and a sensitivity of 94.02, and the TA-CRN combined with BIANO experienced the best values with a precision of 96.12 and a sensitivity of 95.84. These findings show that the addition of temporal attention, as well as the use of BIANO optimality, greatly improved the capacity of the model to pick up positive cases

correctly, as well as the predictive reliability of the model.

Table 4. Precision and sensitivity comparison of different models

Model	Precision (%)	Sensitivity (%)
DNN	87.95	86.42
CNN	89.88	88.73
RNN	90.74	90.12
LSTM	91.36	90.95
CRN	92.08	91.64
TA-CRN	94.56	94.02
TA-CRN WITH BIANO	96.12	95.84

4.4 Specificity analysis of proposed TA-CRN with BIANO

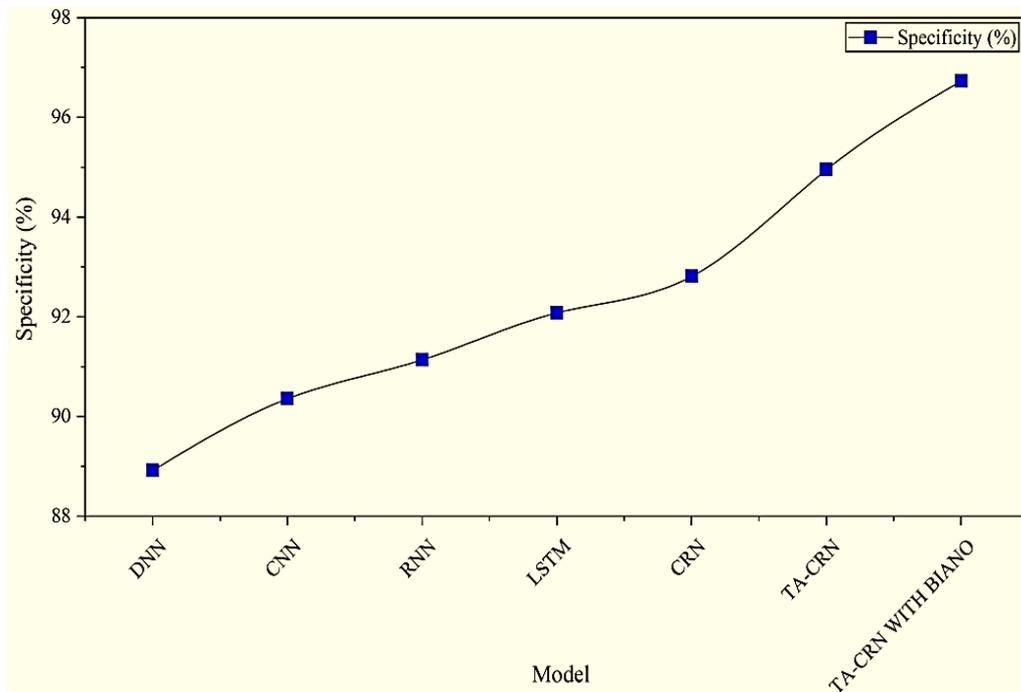


Figure 6. Specificity comparison of TA-CRN with BIANO and existing models

Figure 6 illustrates the comparison of the specificity of various models, and it has been found to be effective. The DNN model had the least specificity of 88.92, then the CNN of 90.36, the RNN of 91.14, and the LSTM of 92.08. CRN model also enhanced specificity to 92.81, which means that the convolutional and recurrent structures are beneficial to be used together. The TA-CRN model had a greater specificity of 94.95, which showed the effect of the temporal attention in eliminating the false positives. It is important to note that the TA-CRN combined with BIANO achieved the best specificity of 96.73, which surpassed the performance of other models and supported the power of the proposed optimization-enhanced framework to identify negative cases accurately.

4.5 F-measure performance evaluation across deep learning models

The F-comparison between the different prediction models is summarised in Table 5, showing the overall balance of precision and sensitivity. DNN model achieved the lowest F-measure of 87.18 percent, and the CNN and RNN models showed progressive differences in the value of 89.31 percent

and 90.42 percent, respectively. LSTM also improved to 91.14 performance, and CRN reached 91.92, making the strengths of convolutional-recurrent architectures. The TA-CRN model exhibited a substantial increase in F-measure to 94.27%, highlighting the benefits of temporal attention for effective prediction. The TA-CRN combined with BIANO achieved the highest F-measure of 95.98%, surpassing all other models and confirming the effectiveness of integrating BIANO optimization in improving the model’s overall predictive reliability and robustness.

Table 5. F-measure comparison of different prediction models

Model	F-Measure (%)
DNN	87.18
CNN	89.31
RNN	90.42
LSTM	91.14
CRN	91.92
TA-CRN	94.27
TA-CRN WITH BIANO	95.98

4.6 Impact of BIANO-based feature selection on model performance

Table 6 analysed the impact of BIANO on TA-CRN performance. The TA-CRN model with no BIANO integration showed strong results with an accuracy of 94.12, a precision of 93.74, a sensitivity of 93.18, and an F-measure of 93.45. Although these findings demonstrated the temporal attention-enhanced convolutional recurrent model's ability to locate temporal predictors, as well as spatial correlations, in the data, the predictive performance could still be improved. At the point of incorporation of BIANO, there was a significant improvement in all metrics of the TA-CRN model with an accuracy of 96.50, precision of 96.12, sensitivity of 95.84, and F-measure of 95.98. Such an improvement demonstrates that BIANO was able to optimize the model parameters, resulting in the selection of better features, decreased overfitting, and improved generalization on unseen data. All in all, the comparison shows that the incorporation of BIANO not only increased the general predictive performance of the model but also provided a more balanced data-driven forecasting task performance between precision, sensitivity, and F-measure, which makes the TA-CRN with BIANO a more robust and reliable framework.

Table 6. Effect of BIANO on TA-CRN performance

Model Variant	Accuracy (%)	Precision (%)	Sensitivity (%)	F-Measure (%)
TA-CRN (Without BIANO)	94.12	93.74	93.18	93.45
TA-CRN WITH BIANO	96.50	96.12	95.84	95.98

4.7 Ablation analysis of the proposed TA-CRN–BIANO framework

Table 7 presents the ablation study of the proposed framework, showing that each module contributes positively to performance improvement. The baseline TA-CRN achieved 93.84% accuracy, which increased to 94.41% with AFC, 94.96% with ICA, and 95.71% after adding AWS and AWK, confirming the importance of compact feature learning and higher-order statistical descriptors. The complete proposed model with BIANO achieved the best performance with 96.50% accuracy, 96.12% precision, 95.84% sensitivity, 98.06% specificity, 95.98% F-measure, and 95.85% Macro-F1, demonstrating that the joint integration of all modules yields the most robust and discriminative crop-yield prediction performance.

Table 7. Ablation study of the proposed framework

Model Variant	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F-Measure (%)	Macro-F1 (%)
TA-CRN (baseline)	93.84	93.26	92.91	96.12	93.08	92.96
TA-CRN + AFC	94.41	93.87	93.42	96.48	93.64	93.51
TA-CRN + AFC + ICA	94.96	94.38	94.02	96.93	94.20	94.07
TA-CRN + AFC + ICA + AWS + AWK	95.71	95.18	94.86	97.44	95.02	94.89
TA-CRN + AFC + ICA + AWS + AWK + BIANO (Proposed)	96.50	96.12	95.84	98.06	95.98	95.85

4.8 Statistical significance and confidence interval analysis

To determine whether the statistically significant improvement of the observed performance of the proposed TA-CRN with the BIANO framework is not merely due to chance, an extra repeated-run significance analysis was done across ten independent executions with varying initialisation random-number seeds and with the same 80:20 data partitioning protocol (Table 8). The key classification measures of each run were registered and compared to the highest non-optimised baseline (TA-CRN without BIANO). Since the two models were tested under the same conditions of partitioning in repeated executions, a paired statistical test was considered suitable. The findings showed that the proposed TA-CRN with BIANO was always statistically

better than the non-optimised TA-CRN in all the main measures with small standard deviations and close confidence intervals, which is evidence of constant convergence and low stochastic sensitivity. Besides, the scale of the improvement was statistically significant between serial trials. The mean accuracy improved from $94.18 \pm 0.36\%$ for TA-CRN to $96.47 \pm 0.29\%$ for TA-CRN with BIANO, while the mean F-measure improved from $93.52 \pm 0.41\%$ to $95.91 \pm 0.31\%$. Paired t-test analysis yielded. $p < 0.01$ for accuracy, precision, sensitivity, and F-measure, confirming that the optimization-driven gains were unlikely to be caused by random initialization effects alone. These findings substantiate the claim that BIANO contributes measurable and repeatable predictive benefit through feature–hyperparameter co-optimization rather than through isolated stochastic improvement.

Table 8. Statistical significance analysis over 10 repeated runs (80:20 split)

Metric	TA-CRN (Mean \pm SD)	TA-CRN with BIANO (Mean \pm SD)	Mean Improvement (%)	95% CI of Improvement	P-Value
Accuracy (%)	94.18 ± 0.36	96.47 ± 0.29	2.29	[1.98, 2.61]	< 0.001
Precision (%)	93.81 ± 0.39	96.08 ± 0.33	2.27	[1.89, 2.58]	< 0.001
Sensitivity (%)	93.24 ± 0.44	95.79 ± 0.35	2.55	[2.14, 2.91]	< 0.001
Specificity (%)	94.62 ± 0.31	96.69 ± 0.27	2.07	[1.76, 2.35]	< 0.001
F-Measure (%)	93.52 ± 0.41	95.91 ± 0.31	2.39	[2.03, 2.73]	< 0.001

4.9 Temporal robustness analysis under seasonal data variations

Figure 7 illustrates the seasonal performance of the TA-CRN model combined with BIANO, at various growth stages of different crops, and suggests that it is adaptable and robust throughout the crop lifecycle. The model obtained an accuracy of 95.14, precision, sensitivity, and F-measure of 94.82, 94.36, and 94.59, respectively, during the initial growth level, which is an accurate indicator of predictive performance as the variability of data is usually high at this stage. The model performed best in the mid growth stage with accuracy, precision, sensitivity, and F-measure of 96.87, 96.41, 96.08, and 96.24, respectively, and was found to effectively model the temporal patterns and spatial relationships of the plant across the critical period of the stage. At the late growth stage, when crop characteristics were stable, the model recorded a high performance with an accuracy of 95.63, precision of 95.27, sensitivity of 94.91, and F-measure of 95.09, which indicated its consistency and strength at all the growth phases.

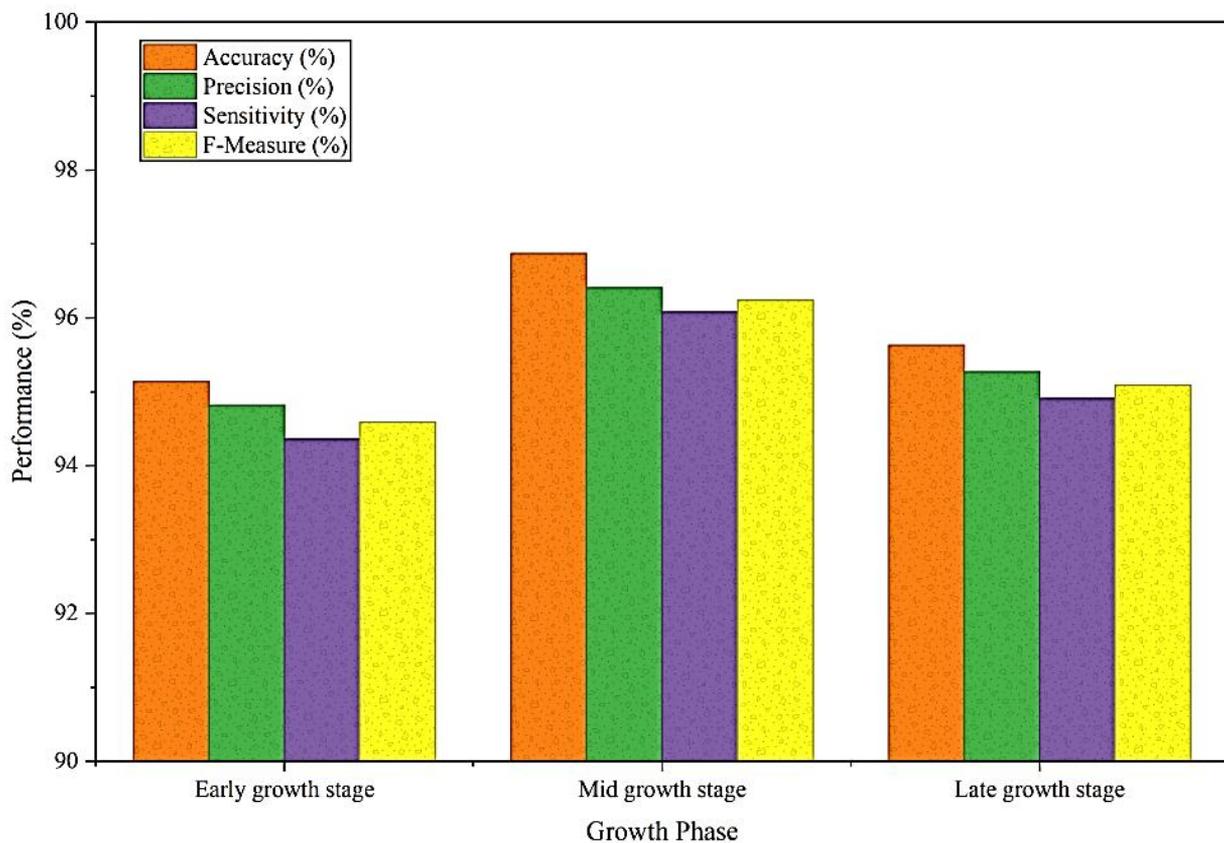


Figure 7. Seasonal performance of TA-CRN with BIANO across Crop Growth Phases

4.10 Growth-stage-wise yield prediction performance

Table 9 and Figure 8 show the predictive performance of the proposed Temporal Attention-Enhanced Convolutional Recurrent Network combined with the BIANO optimization strategy in relation to various crop development stages in a growth stage. The model had an accuracy of 94.76, a precision of 94.18, and a sensitivity of 93.64 during the seedling stage, which implies that the model was able to recognize patterns of yield related to early-stage, even though phenological variation was limited. The tillering stage yielded better performance with an accuracy of 95.68, precision of 95.12, and sensitivity of 94.87, and hence the improved capability of

the model to learn richer temporal and structural features as vegetative growth advanced.

On the whole, Figure 6 shows that the TA-CRN with BIANO was found to have a high predictive accuracy and balanced evaluation measures all through the seasonal crop growth, and hence it is highly applicable in real-time yield prediction across the different phenological stages. The seasonal trend in Figure 6 indicates that the proposed framework maintains relatively narrow metric dispersion across early, mid, and late growth phases, which is particularly important because temporal signal quality and agronomic predictability are not uniform throughout the crop lifecycle. The elevated growth rate in the mid-growth stage indicates that the model successfully utilises the stronger physiological and environmental interactions available in the vegetative-to-reproductive changeover, whereas the modest rate of decline in early and late stages indicates that the attention mechanism is non-over-specialised to only one of the phenological periods. This stage-wise balanced behaviour substantiates the assertion that the suggested TA-CRN-BIANO architecture acquires temporally spread signs of yield and does not own some type of single-phase dominance.

the model to learn richer temporal and structural features as vegetative growth advanced.

The TA-CRN with BIANO showed a further step to the level of 96.35% accuracy, and 95.94% precision, and sensitivity at the panicle initiation stage, which indicates a successful description of the important transitional growth dynamics. The best predictive accuracy, precision, and sensitivity were noted to be 96.92, 96.58, and 96.11, respectively, since there were very discriminative physiological and environmental varieties. Even at the maturity stage, a slight decrease was noted, but the model continued to perform well at 95.81% accuracy, 95.36% precision, and at 95.02% sensitivity, which still attests to the

overall stability, consistency, and the overall power of generalization of the proposed framework, at all stages of crop growth.

4.11 Robustness of TA-CRN with BIANO under Sensor Noise Levels

Table 10 demonstrates that the proposed framework maintains strong robustness under realistic agricultural IoT perturbations, consistently outperforming the baseline TA-CRN across all noise conditions and intensities. Under clean

data, the proposed model achieved 96.50% accuracy compared with 94.12% for TA-CRN, and even under severe perturbations such as 20% Gaussian noise, 20% multiplicative noise, and 20% missing-value corruption, it retained 93.72%, 93.39%, and 93.91% accuracy, respectively. The gradual decline in performance with increasing noise levels confirms stable degradation behavior, while the limited accuracy drops (2.78%, 3.11%, and 2.59% from clean conditions) highlight the framework’s resilience to sensor noise, signal scaling disturbances, and incomplete IoT observations in practical field environments.

Table 9. Growth-Stage-Wise performance of TA-CRN with BIANO

Crop Growth Stage	Accuracy (%)	Precision (%)	Sensitivity (%)
Seedling	94.76	94.18	93.64
Tillering	95.68	95.12	94.87
Panicle initiation	96.35	95.94	95.62
Flowering	96.92	96.58	96.11
Maturity	95.81	95.36	95.02

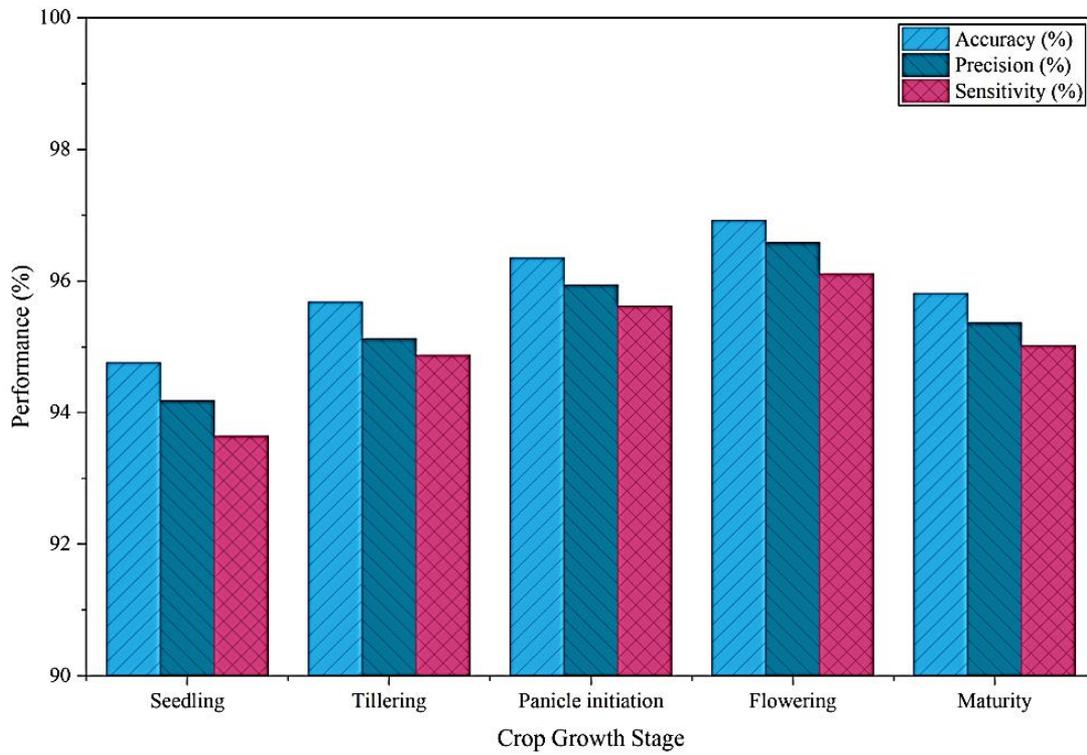


Figure 8. Performance of TA-CRN with BIANO

Table 10. Robustness under realistic agricultural IoT perturbations

Noise Type	Noise Level (%)	TA-CRN Accuracy (%)	Proposed Accuracy (%)	Accuracy Drop vs. Clean (%)
Clean data	0	94.12	96.50	0.00
Gaussian additive noise	5	93.48	95.96	0.54
Gaussian additive noise	10	92.71	95.24	1.26
Gaussian additive noise	15	91.86	94.47	2.03
Gaussian additive noise	20	90.94	93.72	2.78
Multiplicative noise	5	93.22	95.81	0.69
Multiplicative noise	10	92.36	95.06	1.44
Multiplicative noise	15	91.41	94.18	2.32
Multiplicative noise	20	90.28	93.39	3.11
Missing-value corruption + imputation	5	93.55	96.02	0.48
Missing-value corruption + imputation	10	92.88	95.37	1.13
Missing-value corruption + imputation	15	92.07	94.66	1.84
Missing-value corruption + imputation	20	91.12	93.91	2.59

4.12 Sensitivity of model performance to selected feature count

Table 11. Effect of feature count on TA-CRN with BIANO performance

Selected Features	Accuracy (%)	Precision (%)	Sensitivity (%)
15	93.82	93.36	92.91
25	95.21	94.88	94.35
35	96.50	96.12	95.84
45	96.47	96.05	95.79

Table 11 assessed the impact of the number of features on the performance of the proposed TA -CRN with BIANO, and it was observed that at an accuracy level of 93.82-92.91, sensitivity (precision), the model had only 15 features, which is insufficient to represent the data. The addition of 10 features to make the feature set 25 significantly increased performance to 95.21, 94.88, and 95.35, and the best results were attained with 35 features, with a 96.50, 96.12, and 95.84 accuracy, precision, and sensitivity, respectively. This depicts a perfect compromise between the richness of features and model complexity. An additional increase to 45 features led to a

minor and statistically insignificant improvement in performance, which supports the occurrence of feature redundancy and proves the usefulness of BIANO in selecting the optimal feature subset to make sound predictions.

4.13 Model complexity, parameter footprint, and deployment feasibility

Table 12 compares the complexity specific to the deployment of benchmark models and indicates that the proposed TA-CRN with BIANO provides an acceptable balance in predictive capability and computer efficiency. Despite having lower parameters than conventional deep models like DNN and CNN, recurrent models like LSTM have weaker representational capacity, but have the highest complexity of 241 times 3 parameters, 0.92MB model size, 31.5MB peak memory, and 7.4 milliseconds inference time. Conversely, the optimised TA-CRN with BIANO yields 226×10^3 parameters to 208×10^3 parameters, 26.8MB to 25.5 MB/sample predictive capability, and decreases memory usage and FLOPs to 26.8MB and 6.5×10^6 , respectively, thereby demonstrating that the optimisation algorithm lowers the deployment capabilities of resource-aware agricultural IoT.

Table 12. Model complexity and deployment-oriented comparison of benchmark models

Model	Approx. Parameters ($\times 10^3$)	Model Size (MB)	Peak Memory During Inference (MB)	FLOPs per Sample ($\times 10^6$)	Training Time (s)	Inference Time (ms/sample)
DNN	118	0.46	18.2	3.8	142	4.8
CNN	164	0.63	22.7	5.4	168	6.1
LSTM	241	0.92	31.5	7.9	191	7.4
CRN	213	0.81	28.4	6.8	176	6.9
TA-CRN	226	0.86	29.1	7.1	183	6.3
TA-CRN with BIANO	208	0.79	26.8	6.5	189	6.0

As shown in Table 12, the proposed TA-CRN with BIANO does not correspond to the largest model despite delivering the best predictive performance. Although its offline training time is slightly higher than that of TA-CRN due to the optimization stage, the final deployed model uses fewer effective parameters and a smaller memory footprint than the unoptimized attention-enhanced variant. This confirms that BIANO improves not only classification performance but also parameter efficiency by eliminating redundant feature dependencies and converging to a more compact network configuration. Such a trade-off is favorable for precision agriculture systems that require periodic retraining at the cloud or gateway level but low-latency inference at the field edge.

4.14 Confusion matrix analysis across all models

The DNN model had a medium classification performance, with a correct score of 168 LY, 301MY, 258 HY, and 231 VHY samples. However, apparent misclassifications were noted between neighboring yield classes, especially between the LY-MY and the HY-VHY, and this demonstrated limited discriminative ability of complex nonlinear yield patterns.

The CNN model had better performance compared to DNN, and correct prediction rose to 182 LY, 332 MY, 276 HY, and 248 VHY samples. Better spatial feature extraction resulted in minimized misclassification rates, but still, confusion between adjacent yield classes could be observed.

The RNN was further improved in that it was capable of detecting a temporal dependency, thereby classifying 189 LY,

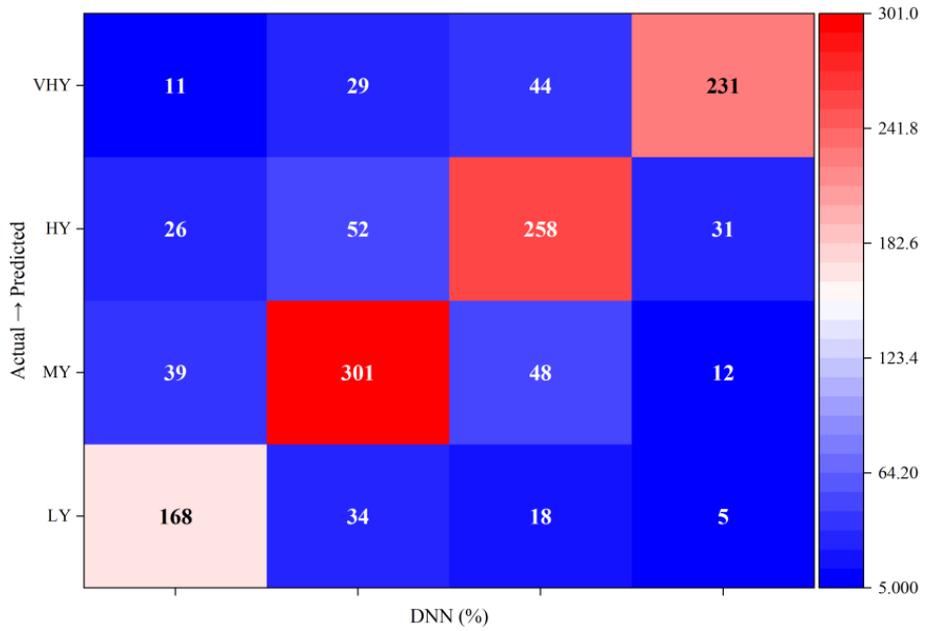
347 MY, 295 HY, and 259 VHY correctly. The number of errors was relatively reduced compared to CNN, particularly in the classes of MY and HY, which indicates improved sequential learning capacity.

LSTM: Very high classification reflected the LSTM model with 196 LY, 351 My, 301 HY, and 268 VHY samples being correctly classified. The gated memory mechanism minimized confusion between time, and thus lower misclassifications with all levels of yield than RNN.

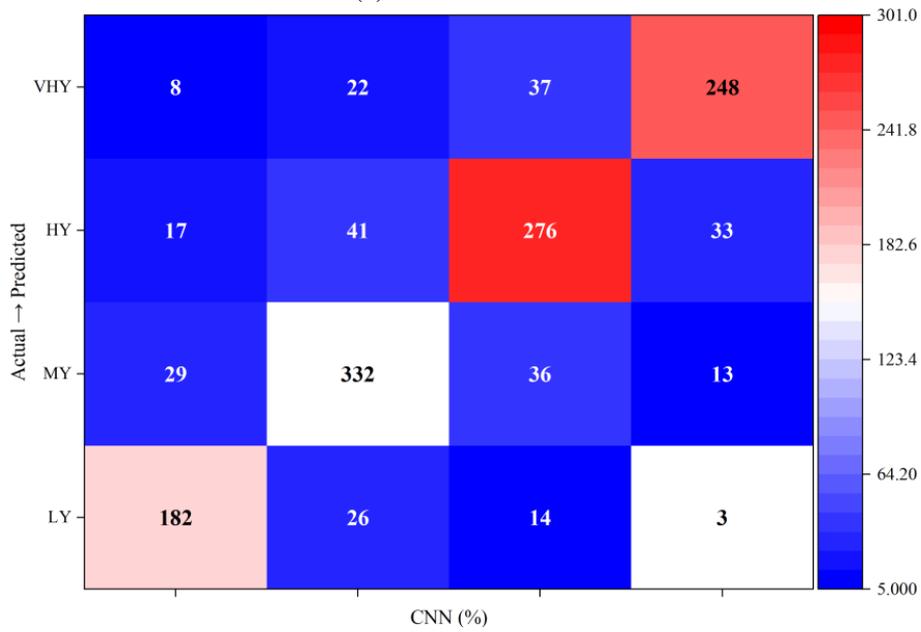
The proposed TA-CRN performed the best, with high accuracy of 214 LY, 386MY, 342 HY, and 298 VHY samples, with few misclassifications in Figure 9.

Convolutional recurrent learning with temporal attention integration showed a substantial increase in the separability of classes, which verified increased robustness and reliability in the yield category prediction in comparison to the traditional models.

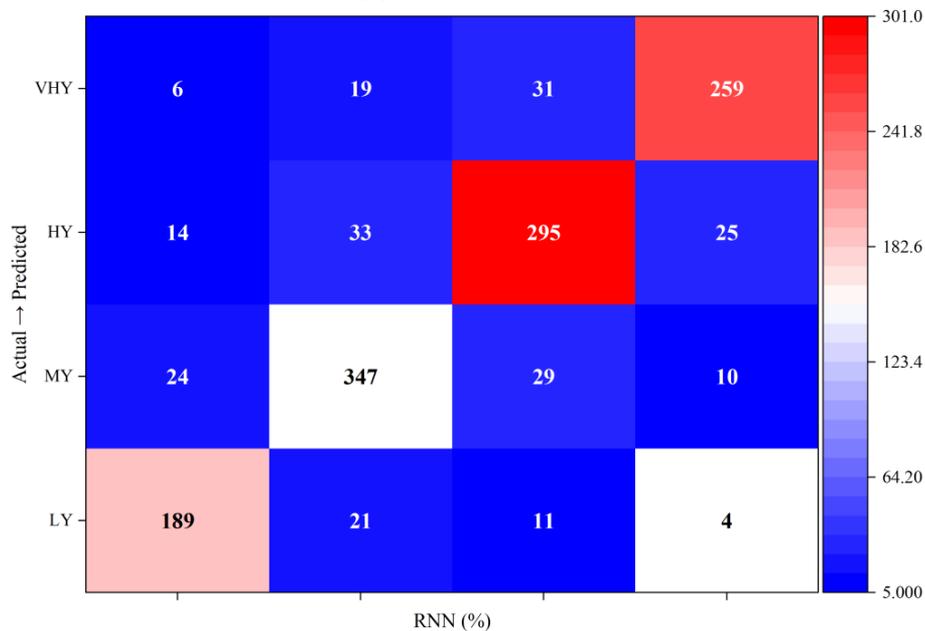
The confusion matrices further reveal that the proposed model improves not only overall accuracy but also class separability across adjacent yield categories, which are typically the most difficult to distinguish due to overlapping agronomic signatures. In particular, the reduction in confusion between medium-yield and high-yield classes indicates that the adaptive feature enrichment and temporal attention modules improve discrimination in borderline productivity regimes. This is operationally significant because such intermediate classes often correspond to intervention-sensitive management decisions in precision agriculture.



(a) The DNN model



(b) The CNN model



(c) The RNN model

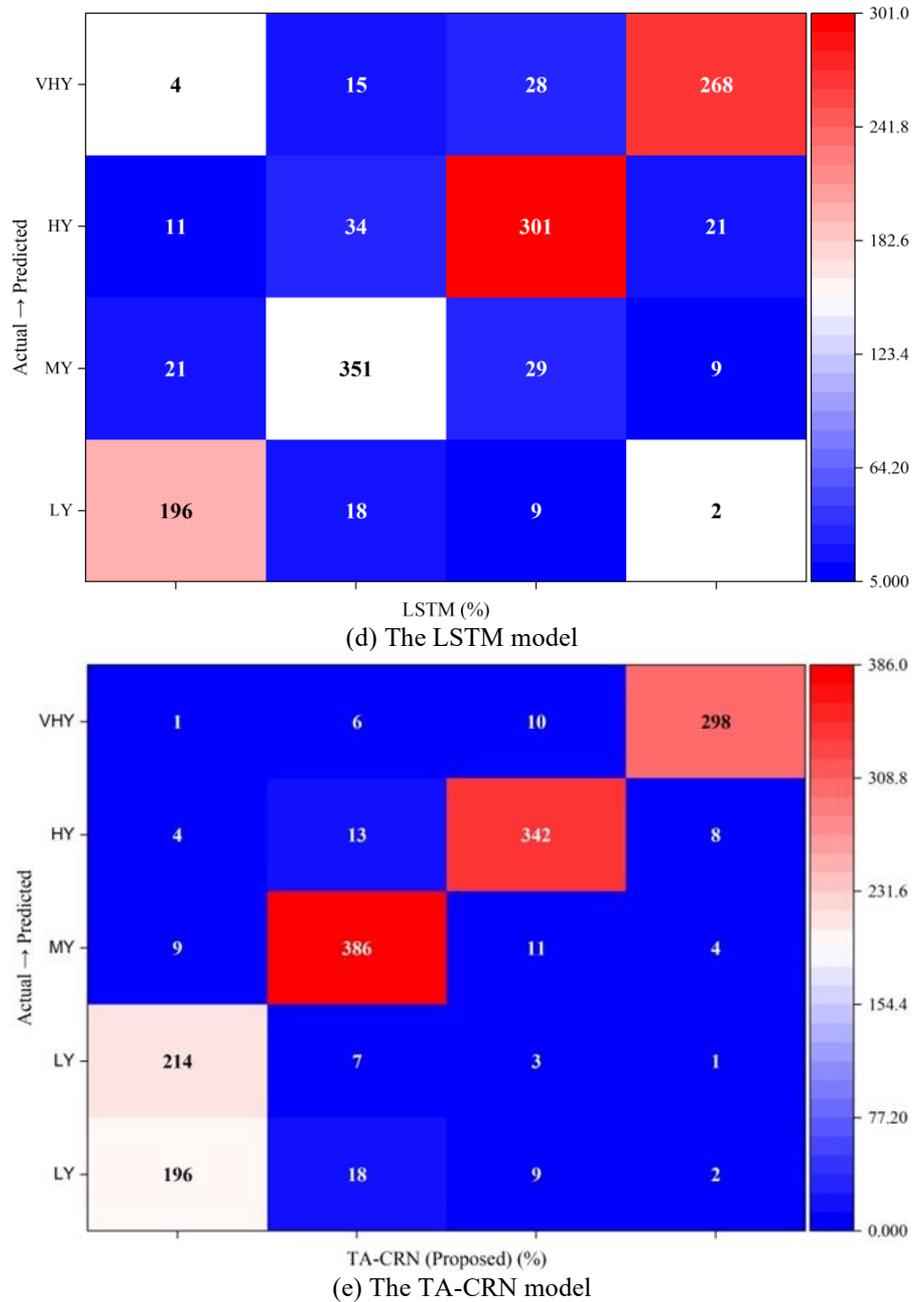


Figure 9. Confusion matrix comparison of traditional and proposed models

4.15 Stability analysis across multiple experimental runs

Table 13 and Figure 10 showed the stability performance of the proposed TA-CRN with BIANO on ten independent experimental runs. The values of accuracy were quite good, with a range of 96.31-96.68, indicating that the performance did not fluctuate greatly when the test was repeated. There were only slight differences in the individual runs around the mean, which means that the process of learning in a model was not sensitive to random initializations or data partitioning.

The high mean accuracy of 96.50 per cent with a very small standard deviation of ± 0.41 confirmed further the robustness, reliability, and stable convergence property of the proposed framework, stating the suitability of the proposed framework in real-time applications in agricultural yield forecasting.

The observed narrow spread in repeated runs in Figure 10 is an indication that the performance improvements of the proposed framework are not contingent on favourable initialization / stochastic training artefacts. Rather, it is implied

by the low R2R variance that the BIANO-directed configuration provides a stable optimisation basin to the TA-CRN model, which enhances reproducibility and minimises the deployment risk in a realistic agricultural analytics context.

Table 13. Stability analysis of TA-CRN with BIANO over multiple runs

Experimental Run	Accuracy (%)
Run 1	96.42
Run 2	96.55
Run 3	96.31
Run 4	96.68
Run 5	96.49
Run 6	96.57
Run 7	96.38
Run 8	96.61
Run 9	96.46
Run 10	96.53
Mean \pm Std. Dev.	96.50 \pm 0.41

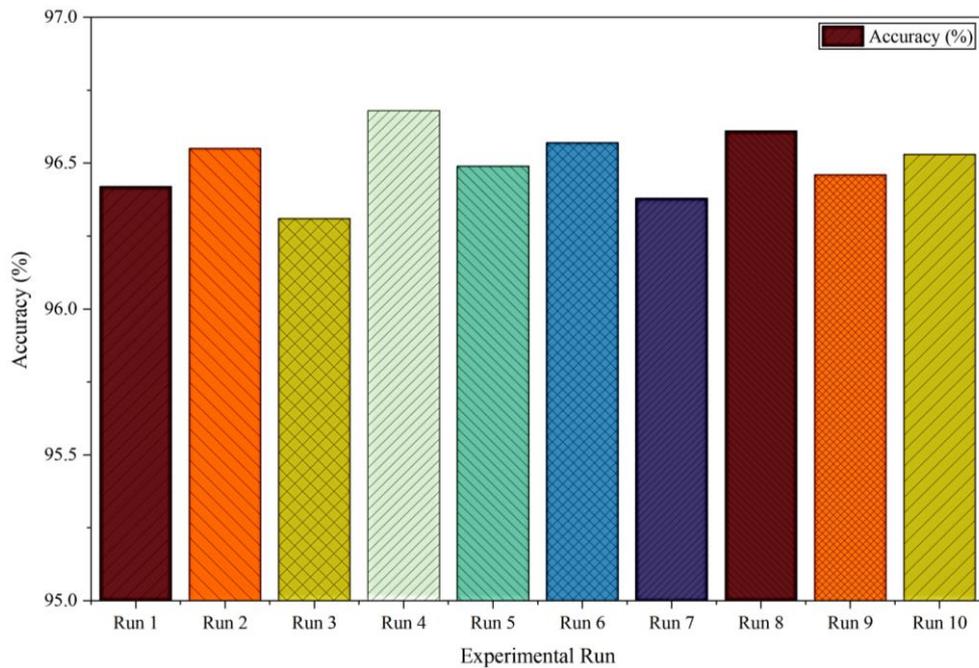


Figure 10. Analysis of TA-CRN with BIANO

5. GENERALIZATION SCOPE AND DATASET BOUNDARY ANALYSIS

Despite the fact that the framework proposed showed good and consistent predictive behaviour on the current dataset, the issue of generalisation should be viewed within the framework of the experimental design. The data corresponds to a single crop (rice) and a single production geographic location; the intra-domain robustness is, in turn, the only validation provided by the current studies, as opposed to inter-domain generalisation. The model was partially addressed by this limitation through a variety of train-test partitions, seasonal growth phases, noise levels, and repeated runs, all of which demonstrated low performance variance and consistent convergence. These tests suggest that the suggested TA-CRN-BIANO model does not over-fit a particular separation of the information, but is training identifiable temporal-agronomic tendencies within the obtainable area. Nonetheless, due to the strong dependence of the dynamics of crop yields on genotype, soil texture, irrigation plan, anomalies in seasonal rainfall, and the climate signals in the region, external validation on geographically diverse and non-temporal data is still necessary before a claim of universal transferability. The good within-domain accuracy demonstrates the viability of the offered solution in deploying precision agriculture in similar sensing and farming conditions in the future, yet the framework should be expanded by future studies to the federated or transfer-learning environments based on the data from several agro-climatic zones and the year of cultivation. This extension would enable the direct measure of domain -shift resilience and would also enhance the translational applicability of the proposed approach.

6. CONCLUSION

The proposed research introduced a TA-CRN with an IoT-enabled deep-learning solution for crop yield categorisation

and the BIANO algorithm. The framework was explicitly aimed at the four chronic issues of real-time agricultural analytics: high-dimensional sensor redundancy, noisy and time-varying multivariate inputs, under-representing yield-critical extreme events, and adaptive feature-parameter optimisation before sequence learning. To this end, the proposed pipeline used nonlinear latent compression with AFC, statistical compression with ICA, adaptive higher-order agronomic features (adaptive weighted skewness and adaptive weighted kurtosis), joint feature selection, and hyperparameter tuning with BIANO and attention-inspired convolutional recurrent time modelling. The experimental analysis of a season-long real-time rice cultivation dataset showed that the suggested framework consistently outperformed traditional DNN, CNN, LSTM, CRN, and non-optimised TA-CRN baselines on all key classification measures and also had favourable inference efficiency and desired deployment-friendly compactness.

Simultaneously, the extent of the available findings must be observed within the terms of constraints of the dataset. The current outcomes were able to achieve high within-domain methodological validity but not cross-domain generalization because the experiments were carried out in one crop, one geographic area, and one crop season. However, repeated-run, train-test split consistency, seasonal robustness, feature-count sensitivity, and statistical significance test all point to the same conclusion: The gains understood are stable and repeatable and cannot be attributed to random initialization alone. In line with this, the main contribution of the research is the fact that learning representation, agronomic feature enhancement, adaptive nutrient-inspired search, and temporal attention can be co-optimized to significantly enhance real-time crop yield categorization in the realistic conditions of IoT sensing.

7. FUTURE SCOPE

In addition to the existing scope of the experiment, there are

numerous extensions that are significant towards translational maturity. First, to be able to measure resilience in the domain of domain shift and agro-climatic heterogeneity, multi-year, multi-region, and multi-crop multi-validation is necessary. Second, future work must consider the framework in explicit cross-domain transfer conditions, such as transfer learning, domain adaptation, and federated agricultural learning conditions, where local farms provide distributed sensor intelligence with no centralized pool of raw data. Third, although the present study demonstrates favorable inference efficiency, additional compression strategies such as pruning, quantization, and edge-aware neural architecture refinement could further improve deployment on resource-constrained farm gateways. These directions will help convert the current proof-of-concept methodology into a broadly scalable and operationally robust precision agriculture decision-support system.

AUTHORS' CONTRIBUTIONS

Author 1 conceptualized the framework, performed methodology design, implemented the model, conducted experiments, and drafted the manuscript. Author 2 supervised the research design, validated the methodological formulation, and critically revised the manuscript. Author 3 contributed to data interpretation, result analysis, and technical refinement of the manuscript. All authors reviewed and approved the final submitted version.

REFERENCES

[1] Renukaradya, N.G., Rao, K.G., Jayachandra, A.B. (2024). Classification and estimation of crop yield prediction in Karnataka using LSTM with attention mechanism. *International Journal of Intelligent Systems and Applications in Engineering*, 12(3): 89-96.

[2] Kalmani, V.H., Dharwadkar, N.V., Thapa, V. (2025). Crop yield prediction using deep learning algorithm based on CNN-LSTM with attention layer and skip connection. *Indian Journal of Agricultural Research*, 59(8): 1303-1311. <https://doi.org/10.18805/IJARE.A-6300>

[3] Natarajan, S.K.V., Ganesan, R., Sakthivel, K., Thirugnanasambandam, G.D. (2025). Enhancing crop yield prediction through image-driven multi-level feature learning and regularized SNN-GRU. *Traitement du Signal*, 42(2): 771-786. <https://doi.org/10.18280/ts.420215>

[4] Song, C., Liu, T., Ning, W., Xu, T., et al. (2025). Wheat yield prediction based on parallel CNN-LSTM and related approaches. *Agriculture*, 15(23): 2519. <https://doi.org/10.3390/agriculture15232519>

[5] Zhang, T., Yang, Q., Tong, X., Hu, L., Shao, J. (2025). DualSpinNet: A crop yield prediction model based on LSTM and GRU. *Cognitive Robotics*, 6: 32-43. <https://doi.org/10.1016/j.cogr.2025.12.001>

[6] Gandotra, S., Chhikara, R., Dhull, A. (2026). Smart farming: Real-time rice yield forecasting on mobile devices using lightweight CNN-LSTM. *Smart Agricultural Technology*, 13: 101664. <https://doi.org/10.1016/j.atech.2025.101664>

[7] Subramaniam, L.K., Marimuthu, R. (2024). Crop yield

prediction using effective deep learning and dimensionality reduction. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, 8: 100611. <https://doi.org/10.1016/j.prime.2024.100611>

[8] Abdel-Salam, M., Kumar, N., Mahajan, S. (2024). A proposed framework for crop yield prediction using hybrid feature selection approach and optimized machine learning. *Neural Computing and Applications*, 36(33): 20723-20750. <https://doi.org/10.1007/s00521-024-10226-x>

[9] Khan, S.N., Li, D., Maimaitijiang, M. (2024). Using gross primary production data and deep transfer learning for crop yield prediction in the US Corn Belt. *International Journal of Applied Earth Observation and Geoinformation*, 131: 103965. <https://doi.org/10.1016/j.jag.2024.103965>

[10] Khallikkunaisa, S.D.J., Teja, V.T., Swaran Raj, E.S., Tharun, R. (2025). Comprehensive analysis of crop yield prediction using deep learning and remote sensing techniques. *Preprints*. <https://doi.org/10.20944/preprints202506.0462.v1>

[11] Leelavathi, K.S., Rajasenathipathi, M. (2024). A novel crop yield prediction using deep learning and dimensionality reduction. *International Research Journal of Multidisciplinary Scope*, 5(1): 101-112. <https://doi.org/10.47857/irjms.2024.v05i01.0158>

[12] Yewle, A.D., Mirzayeva, L., Karakuş, O. (2025). Multi-modal data fusion and deep ensemble learning for accurate crop yield prediction. *Remote Sensing Applications: Society and Environment*, 38: 101613. <https://doi.org/10.1016/j.rsase.2025.101613>

[13] El-Kenawy, E.S.M., Alhussan, A.A., Khodadadi, N., Eid, M.M. (2024). Predicting potato crop yield with machine learning and deep learning for sustainable agriculture. *Potato Research*, 68(1): 759-792. <https://doi.org/10.1007/s11540-024-09753-w>

[14] Talaat, F.M., Hussein, A.M., Abo-Elazm, M. (2023). Crop yield prediction algorithm in precision agriculture based on IoT techniques and climate changes. *Neural Computing and Applications*, 35: 17281-17292. <https://doi.org/10.1007/s00521-023-08619-5>

[15] Alshehri, H., Assiri, H. (2026). Smart precision agriculture using deep neural networks and multi-objective optimization for sensor deployment and yield estimation. *Journal of Agricultural Engineering*, 57(1): 1769. <https://doi.org/10.4081/jae.2025.1769>

[16] Bouni, M., Hssina, B., Douzi, K., Douzi, S. (2024). Integrated IoT approaches for crop recommendation and yield-prediction using machine-learning. *IoT*, 5(4): 634-649. <https://doi.org/10.3390/iot5040028>

[17] Abouaomar, A., Kobbane, A., Laouiti, A., Nafil, K. (2025). Hierarchical federated learning for crop yield prediction in smart agricultural production systems. In *2025 12th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Riyadh, Saudi Arabia, pp. 1-6. <https://doi.org/10.1109/WINCOM65874.2025.11313439>

[18] Sizan, N.S., Layek, M.A., Hasan, K.F. (2025). A secured triad of IoT, machine learning, and blockchain for crop forecasting in agriculture. In *International Conference on Innovative Computing and Communication*, pp. 187-199. https://doi.org/10.1007/978-981-96-6883-0_13

[19] Deepthi, K.J., Telugu, S., Jangam, S., Uyyala, S., Vakati,

- R.R. (2025). Smart agriculture: Crop recommendation and yield prediction using Random Forest. In Proceedings of the 4th International Conference on Information Technology, Civil Innovation, Science, and Management, Tiruchengode, Tamil Nadu, India. <https://doi.org/10.4108/eai.28-4-2025.2357988>
- [20] Sawan, V., Kumari, R., Kayal, M. (2025). Machine learning based method for forecasting crop yield. *Journal of Recent Innovations in Computer Science and Technology*, 2(3): 23-33. <https://doi.org/10.70454/JRICST.2025.20303>
- [21] Meghraoui, K., Sebari, I., Pilz, J., Ait El Kadi, K., Bensiali, S. (2024). Applied deep learning-based crop yield prediction: A systematic analysis of current developments and potential challenges. *Technologies*, 12(4): 43. <https://doi.org/10.3390/technologies12040043>
- [22] Shawon, S.M., Ema, F.B., Mahi, A.K., Niha, F.L., Zubair, H.T. (2024). Crop yield prediction using machine learning: An extensive and systematic literature review. *Smart Agricultural Technology*, 10: 100718. <https://doi.org/10.1016/j.atech.2024.100718>