Journal of Theoretical and Applied Information Technology
15th January 2026. Vol.104. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IUX -ADNET: MULTI-FUSION DATA ANALYSIS WITH ATTENTION APPROACH COMBINED WITH IUX DENSE MODEL FOR ACCURATE COTTON YIELD PREDICTION

PORANDLA SRINIVAS[1], DR. A SURESH*[2],

[1] Research Scholar, Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, 603203 Chennai, Tamil Nadu, India
[2] Associate Professor, Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chengalpattu, 603203 Chennai, Tamil Nadu, India
Email : srinivas.research@yahoo.com, prisu6esh@yahoo.com

ABSTRACT

Optimized approaches in agricultural productivity are known and critical factors to predict and identify the precise cotton yield prediction. This paper incorporates an innovative model architecture with the IUX framework using custom functionalities (I (identify), U (SDE), and X (explicit-DE)) implemented to integrate the multi-fusion approach on data. The proposed methodology describes the three models (i) IUX-Dense Neural Net, (ii) IUX-Attention model, and (iii) IUX-Fused Model, which are concatenated with both IUX-Dense and IUX-Attention models. These models are extensively trained with a multimodal-fused dataset (Mendeley dataset + weather-parameter-based synthetic dataset) that is imbued with different features. The experimental evaluations are implemented with 25 epochs, and its performance stability with early stopping criteria demonstrates the IUX-fused approach is far superior with 99.99% accuracy with both disease classification and yield predictions with the IUX-optimized solution with the least MSE (0.05), RMSE (0.24), and MAE (0.188) and the optimal EVS (explained variance score) (97.58). Similar comparisons with machine learning and other SOA architectures are made and tabulated with performance results on both classification and yield prediction analysis. Finally, the proposed framework with the IUX model approach, as IUX-ADNET demonstrates, has far superior capabilities and offers scalable and more reliable solutions towards cotton yield predictions and disease classification.

Keywords: *Stochastic Differential Equations (SDE), Differential Equations (DE), Explained Variance Score (EVS), Identify Feature with Stochastic Explicit differential Equations (IUX), Deep Learning (DL), Machine Learning (ML)*

1. INTRODUCTION

Accurate prediction of cotton yield is still a challenge to resolve, owing to the high dynamic, interdependent and noisy nature of agricultural data. This study was undertaken to address the critical gap in the existing prediction systems that fail to simultaneously model the variation in weather, soil nutrients, severity of diseases, and multimodal field observations in a unified and noise resilient way. To overcome these limitations, in this work, an original and innovative IUX-based framework that uses Deterministic-Stochastic modeling, deep learning, custom attention mechanisms, and multi-fusion dataset design is introduced. In contradistinction with traditional models which are based on single modal or linear assumptions, the proposed IUX-Dense, IUX-Attention, IUX-Fused, are so designed which will capture the non-linear interaction with the environment, can detect the hidden pattern, from the heterogeneous data and can adapt with missing or corrupted measurements from the field. This methodological advancement is the backbone of the novelty of this study that proposes not only an intelligent model but also a game-changer feature for classifying the disease in cotton and projecting yield.

Considering the yield factors in agriculture, cotton production needs yield prediction accuracy, as it can reduce the problems associated with managing the resources, monitoring the supply chain, and developing policy in the field of agriculture. Recent machine learning algorithms and profound learning models have provided real-time field observations and statistical models used to estimate yield during traditional times. UAV imagery of high resolution has been processed by multiscale convolutional neural networks (CNNs) for yield prediction and more detailed information about crop development and conditions [1]. However, machine learning models based on earth observation time series that are interpretable are very usable at the cost of somewhat reduced model prediction precision in favor of increased model interpretability [2]. But Random Forest (RF) models are still being used because they can handle mixed data such as soil and climate and generate coarse regional yield amounts [3]. At present, the research initiatives are employing ensemble and hybrid models to better generalize under a restrictive data set. According to research work in [4], the fusion of synthetic and field data, using ensemble methods, improves cotton yield prediction accuracies when operating on scarce real-world data sets. The efficiency of the RNNs in the crop data analysis is demonstrated using its LSTM variant to test for time dependency, which appeared first under wheat using cotton data, because the two have the same seasons, as mentioned in [5]. Outside, various base learners, including decision trees and SVMs, are combined across complex ensemble expert systems in a variety of agro-climatic conditions [6].

According to [7], based on spatial and temporal data, the combination of hybrid GNN and RNN models is remarkably efficient, particularly for regions where there exists detailed terrain information for which standard long short-term memory networks have been recommended in [7]. Satellite-based models and regression techniques are applied for several investigations. CNNs can be used to process multispectral satellite images and predict market prices and yield as well as provide a combined view for agricultural economic planning [8]. This decomposition-and-ensemble methodology, which was originally applied to human motion prediction, has shown utility in modeling time series of complex yield-influencing factors in agriculture [9]. As research carried out in [10] shows, multi-objective modeling is necessary in sustainable agriculture since optimization algorithms optimized on input parameters such as irrigation and fertilization are carried out. An important step forward is taken by the researchers utilizing a time series-based binary classification system or semi-supervised deep learning based on regression on top of prediction performance [11, 12, 13]. In particular, the topic of cotton yield prediction bears emphasis because it necessitates developing interdisciplinary methods that combine spatial trajectory modeling obtained by autonomous systems [14] and extensive algorithm comparisons using domain-dependent feature engineering [15].

A. 1.1 Problem Statement:

Cotton yield has been predicted as a combination factor and therefore is very crucial in the weather, soil, and stress factors prediction. The change in the climatic conditions, like the change in temperature and rainfall, and the extreme weather conditions, like the drought or flood, can cause some notable effects on the cotton growth. Atypically higher trends in temperatures could lead to heat stress, and unpredictable rainfall could lead to either drought or waterlogging, which is not favorable for the development of the crop. Moreover, shifts in weather extremities disrupt the usual growth cycles in crops, giving rise to difficulties related to the calculation of yield patterns. There are also soil-related issues (level of fertility and availability of nutrients, such as nitrogen, phosphorus, and potassium) that make predictions even more complex. Uneven health of soil in different areas will result in highly unreliable harvests, and in case nutrient status is not ideal, cotton will fail to develop and will have a direct impact on cotton productivity (as discussed in Figure 1). As has been highlighted, researchers M.S. Kuriakose and T. Singh [16] and R. Rashid et al. [17] have indicated that the significance of the incorporation of soil and weather data in prediction modeling has been a challenging task due to inconsistency and poor availability of data.

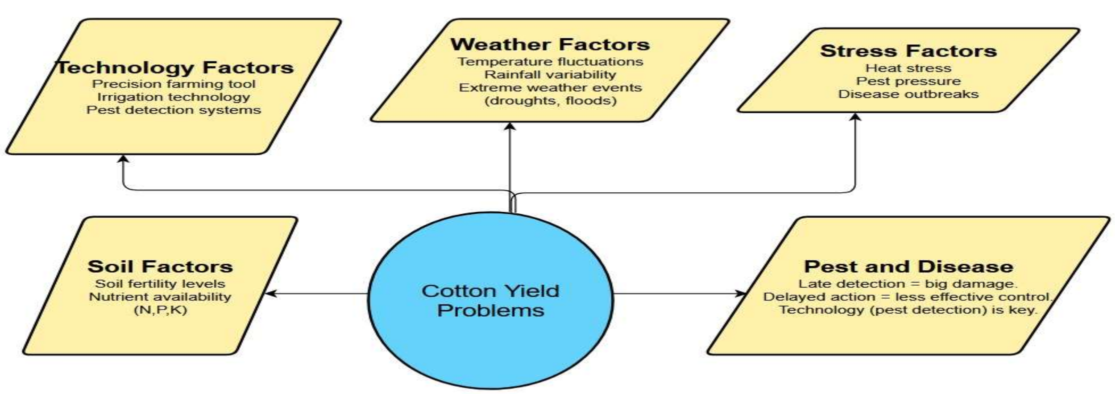


Figure 1: Representing The Overall Problems For Yield Prediction For Cotton Crops

Furthermore, other stressors, including pest infestations and disease outbreaks, increase yield prediction uncertainty. Such pests as the cotton bollworm (direct damage to the crop), and such diseases as Fusarium wilt (reduced plant health, yield potential). These stressors serve yet another unpredictable factor that needs to be incorporated into any model. M. S. Pest and disease data should be regarded as a necessity, as Isik et al. [2] and S. Wu et al. [12] assert but acquiring such data in real-time and across large agricultural areas is not an easy task. Although farmers can use technologies such as precision farming tools, pest detection systems, and remote sensing to address some of these issues, until then, their dependence on high-quality, region-specific data continues to constitute significant limitations. These weather, soil, and stress interactions form a complex outcome, and there is a need to unite these different parameters within a robust data-driven model to accurately forecast cotton yield.

B. Challenges:

The complexities of multi-variable integration, including weather, soil, and disease factors, make cotton yield prediction take a lot of effort to integrate because of the interdependence of the variables, which are at times unpredictable, on crop health and productivity. Variation in weather factors, including extreme temperatures and rainfall, has a strong impact on cotton growth and soil conditions that may include fertility and nutrient availability, making the prognosis even trickier. Consequently, the proposed solution would comprise a robust dataset of disease prediction and detection, which tackles the critical place of pest and disease outbreaks in the reduction of cotton yield. This data set may be combined with the weather and soil factors, as well as the heuristic features, which are based on real-time data, to predict better cotton yield according to various conditions involved in the model. The proposed ensemble model, which implies the combination of dense and attention models, is more accurate in prediction since it presupposes the usage of complex relationships of the variables. The dense model is extremely effective in its characteristic's discovery, and the attention mechanism ensures that emphasis is placed on the most essential effects and the priority is assigned to the most important weather, soil, and illness signs. A hybrid solution like this makes it easier to engage in additional data that makes more plausible and realistic forecasts about the farmers.

C. Importance of Cotton yield

Cotton yield prediction is a key challenge that has yet to be resolved because it relies on very different and co-dependent factors such as weather fluctuations, soil nutrient imbalances and unpredictable pest and disease outbreaks; all of which cause significant distortion to the accuracy of the current predictive models. These conditions introduce noise, missing values and inconsistent patterns in both field and remote sensing data which makes it difficult to generalized the typical machine learning and deep learning methods between regions or seasons. As a result, inaccurate yield estimates lead to considerable risks for farming communities to plan their input usage, irrigation and pest control; these are also high risks for supply-chain managers to plan with predictable production levels; and these are high risks for policymakers when designing agriculture-supporting programs and market measures. Therefore, the absence of a robust, multimodal, and noise-resilient prediction framework is a major issue for every stakeholder who relies on the accurate prediction of cotton yield for the purpose of operational, economic and strategic decision-making.

D. Research gap

Despite the existence of many sources of dependency in yield of cotton with weather, soil fertility, and biological stressors, current prediction systems lack the implementation of these multi-source variables in a unified, noise resilient, and real-time analytical model. Existing approaches have problems of inconsistent field data, incomplete pest and disease observations, and highly fluctuating environmental conditions so that their predictions are not reliable across regions and seasons. While previous research focuses on the role of integration of soil and weather factors, there are limited mechanisms to handle the missing values, non-linear interactions, as well as multimodal signals such as UAV image and disease severity characteristics. This shows a distinct research gap, the lack of an adequate, integrated and explainable framework for the joint modelling of environmental variability, the soil and the effects of pests and diseases under uncertainty. Tackling such a gap is important given the many farmers, supply-chain planners and policymakers who rely on accurate yield forecasts for purposes of resource deployment, crop protection choices, corn and commodity price determination and long-term agricultural planning. And without an integrated and adaptive prediction model, stakeholders continue to face production risks, economic instability and inefficiencies in farm management.

E. Proposed Solution

The proposed ensemble model, which implies the combination of dense and attention models, is more accurate in prediction since it presupposes the usage of complex relationships of the variables. The dense model is extremely effective in its characteristic's discovery, and the attention mechanism ensures that emphasis is placed on the most essential effects and the priority is assigned to the most important weather, soil and illness signs. A hybrid solution like this makes it easier to engage in additional data that makes more plausible and realistic forecasts about the farmers. Also, satellite images and UAV images provide the image of the crops filled with visual data of the status that can reveal the onset of any disease, infestation of insects, or stress that is not evident in the numerical data alone. These images are converted with the assistance of the deep learning algorithms which can identify the state of crops and interpret the growth patterns and anomalies.

The IUX-Dense approach that combines the dense neural network with attention mechanisms enables the model to place a weighted emphasis on different data sources, whether numerical or visual, depending on relevance to the current crop stage. The IUX offers cumulative Deterministic and Stochastic Model, becoming more deterministic based on each pattern of I-U-X functionality. The term I refers to Identify the patterns of Data features, U referring to Random permutated stochastic behavior with differential equations affecting the environment variability. While X refers to explicitly stochastic Differential Equations with neural net. To improvise more efficient and better perspective of the design the proposed model is incorporated with custom multi-stage architecture is implicated with IUX functionality as "Custom-Class". The approach on the custom class implores on different aspects of the design criteria chosen based on data factors of noise and its domain variant. These noises have distinct features affecting the environmental criteria (soil measurements missing, unrecorded climate affects, remote sensing environments) to be considered with multi-fused dataset. Finally, to incorporate the overall classification and Yield prediction, the proposed model introduces custom loss and approximation with SDE's are explicitly more sensitive to noise behaviors and pattern identification which are implemented based on custom formulations as mentioned in section 3.5. Similarly, the fusion based on Mendeley datasets provide high-quality curated environmental and agricultural data, while Google hackathon project datasets provide real-world observations from field trials. Through the integration of these datasets, the proposed model with IUX-Dense and IUX attention develops a powerful model that is not only able to forecast yield but also provides practical intelligence on how to enhance farm management practices, provide mitigation measures against risks of altered environmental conditions, and create the best use of resources. This facilitates the model effectiveness and increases the veracity and reliability of cotton yield predictions towards the theory of the development of the model and its architecture, as stated in the objectives.

F. Objectives:

- Phase 1: Cotton Leaf Disease Feature Extraction and Classification IUX-DNNs, IUX-Attention-Based Model, and Custom IUX-Dense Model are developed and compared to accurately analyzing multiple modalities (color, texture, and shape) of cotton leaf images for disease classification.
- Phase 2: To implement a Custom IUX-Attention-Enhanced Disease Impact Analysis to make critical disease features stand out as well as categorize disease severity levels (Critical, Moderate, Mild).
- Phase 3: Yield Forecasting Integrated Disease and Environmental Factors: Develop a IUX-Fused yield prediction model using classification outputs from disease and synthesized environmental (temperature, humidity, and rainfall) data to make severity-adjusted, high-accuracy cotton yield predictions with a low RMSE.

G. Overview

The major body of the paper is organized in four areas as follows: Section two provides an extensive literature review of previous studies on cotton yield prediction, and outlines what has so far been a problem concerning weather, soil, and disease factors. The gaps in the strategies that exist in the literature are also noted and will be dealt with by the proposed approach. Section three describes the proposed methodology, i.e., a thorough description of the ensemble methodology that involves the combination of Dense Neural Networks (DNNs) and Attention Mechanism to work with numerical and image-based data so that it is possible to predict the appropriate yield. The block diagram and algorithms in addition to the system configuration to process and predict data are also included in this section. The fourth section provides results and discussion that illuminates the effectiveness of the proposed model over the existing methods. The results show the model's enhanced accuracy, especially in combining analysis of weather and disease information with the soil conditions. Finally, in section five, the closure of the paper, by summarizing the major achievements of the research, concludes with further directions for improving the model, and its applicability to other crops.

2. BACKGROUND

A. Littérature Survey

H. Niu et al. [1] uses a convolutional neural network (CNN)-based model to predict the in-season cotton yield based on UAV-acquired RGB imagery. This work finally provided a scale-aware CNN so that the introduced CNN could process high-resolution data as a key for accurate yield predictions at various phenological stages. Although the model's results in precision agriculture were promising, the model's dependence on high-quality UAV data limits its adoption in all regions, as UAV data is not always available in all regions. For example, M. S. Isik et al. [2] introduced an interpretable model that used Earth observation series for cotton yield forecasting with a special focus on a transparent model through explainable machine learning models. Though not as accurate as deep CNN approaches, this model constructed multiple satellite timelines and was able to showcase the value of temporal data for gaining insight into yield trends. On the contrary, N. R. Prasad et al. [3] used a random forest (RF) model to forecast regional yields of cotton based on climatic and soil data. Though they were robust in dealing with heterogeneous datasets, they did not achieve the fine resolution capabilities available in models based on CNN. A. Utilizing real field data and synthetic datasets, data augmentation techniques were used in an ensemble learning framework due to Mitra et al. [4] to improve yield prediction. However, this approach introduced bias from synthetic inputs into the trained models (hence affecting tractable model calibration), which in turn improved generalization.

S. Sharma et al. [5], in predicting wheat yield with agricultural sequential data, showed the need to consider the temporal relationship in the sequence of crop data. Although they focused on wheat, the LSTM models were found suitable for predicting the cotton yield based on long-term weather and phenological data, and D. Tripathi and S. K. Biswas [6] proposed a precise ensemble expert system through a maximized accuracy of several base learner types (decision trees, SVM, etc.) for various cropping types. One advantage of their system is that it works well with yield factors and their nonlinear relationships, such as rainfall and temperature, and is therefore appropriate in any type of agricultural environment. J. Fan et al. [7] introduced a new geospatial and temporal crop yield forecasting using Graph Neural Networks and Recurrent Neural Networks. However, this approach could not capture spatial correlations and sequential dependencies, which is a key drawback. This approach outperformed the traditional LSTM models, but the resultant models in this approach would depend on high complexity and data requirements, which would hamper their use in low-resource settings. M. S. Gastli et al. [8] used deep learning together with satellite imagery to analyze simultaneously crop yield and crop price. Nevertheless, their dual-output CNN model may also face the overfitting problem since they lack sufficient labeled training data. Meanwhile, they also proposed a decomposition-and-ensemble framework for analyzing agricultural time series data by suggesting that their framework offers fairly good stability in the prediction performance in the fluctuating field conditions (which could also fix the issue of variable crop yield).

Consequently, J. Jiang et al. [10] have applied various multi-objective optimization algorithms that needed to be adapted for use in agriculture, i.e., optimizing irrigation, fertilizer levels, and pest control with opposing objectives including sustainability, yield, and cost. This multi-parameter approach to precision farming in cotton is important to modern precision farming practices. A. The authors of Arami et al. [11] introduced a model for binary time series forecasting as a classification problem and proved its effectiveness in distinguishing the crop stress stages. The authors showed with this technique high recall to detect critical classes for cotton to determine low-yield risk zones using field data, and S. Wu et al. [12] proposed the use of semi-supervised regression based on deep learning to achieve yield prediction using a few labeled data. In unevenly distributed field-trial data regions, this model was successful and could improve the prediction accuracy compared with the counterparts of fully supervised. J. Jiang et al. [13] incorporated multi-omics interaction networks to predict maize yield, suggesting the integration of genomic data into AI models, which could significantly enhance yield predictions by incorporating genetic factors alongside environmental and soil data. C. Anderson et al. [14] developed a pedestrian behaviour prediction framework in autonomous vehicles, which was later adapted for agricultural systems,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

including modelling crop growth and predicting harvest outcomes under shifting environmental influences.

In their work, R. Rashid et al. [15] compared the performance of several machine learning algorithms, like support vector machines, random forests, and artificial neural networks, to predict the crop yield, highlighting the requirement of localized feature engineering and domain-specific model training for the target crop, namely cotton. S. M. Ardhanaree Kuriakose and T. Singh [1] predict Indian crop yields using LSTM networks where seasonal data such as rainfall and temperature were used and found that it predicted much better than any feedforward neural networks. Nevertheless, the time series data was missing. R. Machine learning models such as RF and SVM were revalidated by Rashid et al. [17] for the prediction of palm oil yield, presenting a prediction of cotton yield from their potential application. This study is concerned with feature selection and optimization to enhance the models' precision. S. According to Addu et al. [18], environmental stressors such as pollution, weather anomalies, etc., ought to be considered in the prediction of the yield of crops. D. Tripathi and S. K. Biswas [19] emphasized their sustainability-oriented approach, where factors other than productivity are to be considered, which is very relevant for cotton, the water-intensive crop, and corroborated their yield forecasting work using ensemble expert systems, which continue to show superior performance utilizing the cross-validation strategy and the combination of synthetic and real-world datasets. J. Based on this, Fan et al. [20] further developed the GNN-RNN framework that can also be adapted to regional variability, applied to different regions having high predictive accuracy when using satellite-based vegetation indices and historical yield data, and offers a good application to cotton farming in various regions.

Similarly, in [21] the same model was used to revisit [21] that predicts satellite images and [21] crop prices forecasting. However, they failed in bad image quality regions and cloud-prone regions. S. Cotton prediction [22] is also like wheat prediction; providing weather data with the wheat prediction may be combined with the cotton prediction, and soybean prediction based on UAV, multispectral, and weather data [23] should be extended with adaptive cotton prediction to achieve additional accurate cotton prediction. D. Yoon et al. [24] found that such a decomposition spectrum framework for cotton farming, as well as the perspective of the magnitude of the effect of the agronomic practice on the outcome, and the operational decisions used in the framework for the decomposition–ensemble framework of yield variability. Optimization algorithms for aspects of irrigation and fertilization considered to be crucial parts of precision agriculture are proposed by J. Jiang et al. [25]. In addition to their direct application, these models would allow cotton farmers to get increased yield and environmental sustainability through more effective and informed resource management. Based on remote sensing data.

Zhang et al. [26] proposed another rice yield prediction model with a hybrid model of CNN and RNN, which can yield better prediction of the rice yield by integrating spatial and temporal patterns into the data and duplicating the prediction to cotton. Y. Liu et al. [27] presented a fuzzy neural network model to predict wheat uptake for dealing with uncertainty and fuzzy logic problems to enhance the uncertainty in the prediction of the place of unstable weather fluctuation. Because cotton could be grown in the direct effect of the weather and soil, that was the reason. T. N. They demonstrated that tree-based ensemble models in predicting corn yields based on data from Cox's data on weather and soil moisture data, such as the one from Nguyen et al. [28], can also be utilized as an application of cotton farming, with real-time weather and soil moisture data for maximizing resources used.

In the paper [29], H. Li et al. propose deep reinforcement learning (DRL) for predicting crop yield, such as maize, to minimize waste of resources, including irrigation and fertilizer, to increase productivity. Water and nutrient management are critical challenges in cotton production, and especially favourable results can be expected using this approach in cotton. V. Rice yield forecast based on environmental and genetic data by the combination of RF and deep learning model is developed by S. Kumar et al. [30], which has high accuracy and applies to cotton due to its high dependency on genotype x environment interactions. J. Artificial neural networks linked with remote sensing data were also used by Yang et al. [31] to predict the yield of sugarcane, a strategy that could be applied to cotton to improve yield accuracy in geographical regions. C. R. To optimize the use of nitrogen for maximum yield, support vector regression (SVR) applied on soybean yield prediction by Smith et al. [32] is a methodology that could be used for cotton.

In their work, F. M. Alston et al. [33] used machine learning techniques to merge multiple sources of remote sensing data for forecasting barley yield and demonstrated the success of fusing data from multiple sources. The approach is shown to have good potential for improving cotton yield forecasting, particularly in areas lacking or inconsistent with ground truth data. Similarly, G. Zhao et al. [34] introduced an LSTM-based model for predicting wheat yield that takes in meteorological and soil data as input. Two key drivers for crop productivity, soil characteristics, and climatic conditions are both demonstrated by this method to be excellent for cotton yield estimation, especially in this environment.

B. Research Gap

Many recent machine learning and deep learning models exist for crop yield prediction, and some of them can be adapted to better cotton yield prediction. Temporal memory is already useful for multi-step predictions, and multimodal deep learning models using UAV, multispectral, and weather data together are much more accurate for soybean and maize crops. These yield variability decomposition frameworks, fuzzy neural networks, and decision tree-based ensembles have also tried to deal with uncertainty and yield variability with the integration of real-time yield data.

Table 1. Representing The Overall Comparison Of The Algorithm And Approaches With Performance Metrics

Ref. No.	Algorithm / Approach	Accuracy (%)	Limitations
[1]	Scale-aware CNN using UAV imagery	93.2	Rely on high-quality UAV data; limited availability across all regions
[2]	Interpretable ML with EO time series	87.5	Less accurate than CNNs; depends on satellite imagery timelines
[3]	Random Forest (climate & soil data)	81.4	Cannot provide fine-resolution predictions; limited spatial granularity
[5]	LSTM for wheat yield (temporal modeling)	89.6	Requires complete time series; sensitive to missing data
[7]	Hybrid GNN-RNN (spatial + temporal learning)	95.8	High complexity; not suitable for low-resource or low-data environments
[12]	Semi-supervised regression with deep learning	91.7	Needs both labeled and unlabeled data; limited by label sparsity
[18]	ML integrated with environmental sustainability	84.3	Environmental stressors can dominate model outcomes; complex interactions

The technique itself has also evolved into a powerful technique of deep reinforcement learning that can optimize irrigation and fertilizer use required for sustainable management of cotton farming resources. Hybrid models, including CNNs, RNNs, and environmental and genetic data, trained on rice, wheat, and barley exhibit very strong performance, and the grouping of spatial and temporal features appears to be of positive value. It draws upon these insights to train an IUX-Dense and Attention mechanism, where the two mechanisms are jointly pre-trained on the multi-source data to extract and select the most essential features for the final prediction. The innovation and scalability of this approach make it a promising and innovative approach to modern cotton yield forecasting; the robustness and prediction accuracy usually improve in areas with spottily recorded or poor data quality.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

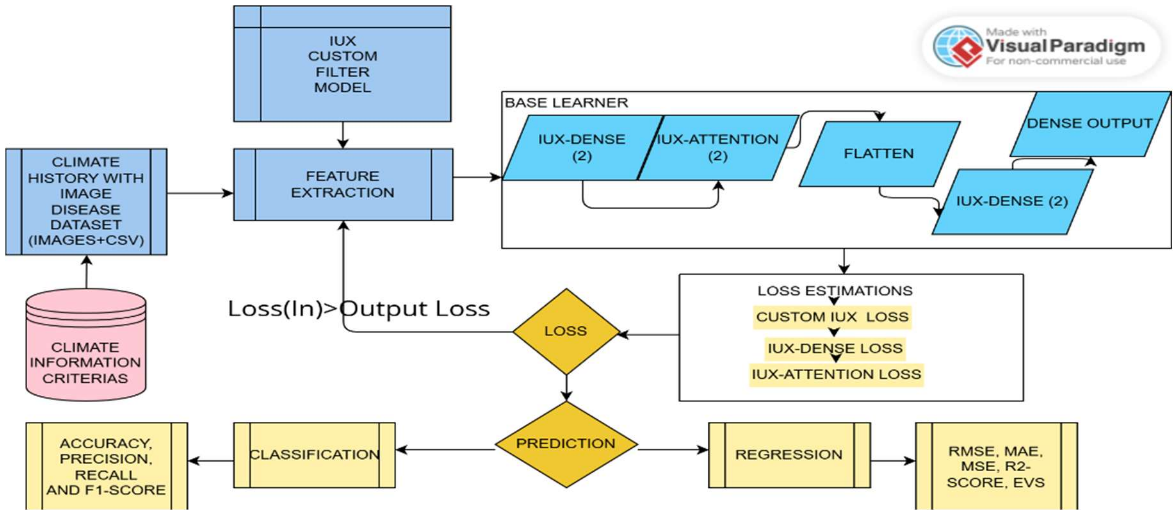


Figure 2: Representing The Overall Flow Architecture Model For The Proposed IUX-Fused (Desne+Attention) Mechanism

3. MATERIALS AND METHODS

The research adopts an experimental, multimodal, and model-driven design structured into three sequential phases: (i) cotton leaf disease classification, (ii) disease severity impact analysis, and (iii) cotton yield forecasting using multi-source fused data. The study integrates two datasets—Mendeley leaf disease images and a synthetic weather–soil–environment dataset—and preprocesses them through normalization, augmentation, and missing-value handling procedures. Three models are developed within the custom IUX computational framework: the IUX-Dense, IUX-Attention, and the IUX-Fused architecture, each incorporating deterministic and stochastic differential equation–based feature transformations. The models are trained for 25 epochs with early stopping, using custom IUX loss functions to optimize noise-sensitive pattern recognition and cross-modal learning. Performance evaluation is conducted using accuracy, RMSE, MSE, MAE, and EVS metrics, and comparative analyses against conventional ML and state-of-the-art DL models are executed to determine the improvement offered by the proposed methodology. Finally, the protocol validates the efficacy of the IUX-Fused model for real-world agricultural applications by assessing its robustness, generalization ability, and predictive stability across disease and yield tasks.

The proposed methodology is aimed at enhancing the extraction as well as the interpretation of functional features for diseases on cotton crops in the real-time available agricultural datasets. The approach works based on a phase design framework and utilizes deep-seated ensemble learning techniques for performing both the classification of disease and the prediction of yield. In the first phase, a hybrid deep learning model designed with dense and attention mechanisms, as well as ridge regression and gradient boosting, is combined such that a class of high precision in the disease severity levels is produced. The second phase in yield prediction consists of an IUX-based feature extraction (IUX filter FE) for feature differentials that are calculated with ranking amongst disease, soil, and weather condition parameters as shown in Figure 2. A complete design of this system leads to a robust, scalable, and data-driven solution to the crop health assessment and yield forecasting problem in a complex agricultural environment.

A. Algorithm Protocol

Algorithm 1: IUX-Based Cotton Yield Prediction Protocol

Input:
  D_img = Mendeley leaf disease image dataset
  D_env = Synthetic weather–soil–environment dataset
  Params = {epochs = 25, batch_size, learning_rate}

Output:
  Y_pred = Predicted cotton yield

  C_pred = Classified disease category

Procedure Begin
  // Phase 1: Data Preparation
  1. Load D_img and D_env
  2. Preprocess D_img:
      - Resize, normalize, augment

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Extract color, texture, shape features
3. Preprocess D_env:
   - Handle missing values
   - Normalize environmental variables
   - Encode categorical factors
4. Fuse datasets → D_fused

// Phase 2: IUX Model Construction
5. Define IUX components:
   I = Identify deterministic feature patterns
   U = Apply stochastic behavior via SDE-based noise modeling
   X = Compute explicit SDE-driven feature transformations
6. Build IUX-Dense model architecture
7. Build IUX-Attention model
8. Concatenate outputs → Construct IUX-Fused model
9. Implement Custom IUX Loss Function

// Phase 3: Model Training
10. Train each model (Dense, Attention, Fused) on D_fused
11. Apply early stopping for stability
12. Save best model weights

// Phase 4: Evaluation
13. Evaluate models using:
    Accuracy, MSE, RMSE, MAE, EVS
14. Compare against baseline ML/DL models
15. Select best-performing model → M_best

// Phase 5: Prediction
16. Use M_best to output:
    C_pred for disease classification
    Y_pred for yield prediction

End Process

Algorithm 1 presents the entire research protocol for the IUX-based cotton yield prediction system that starts with the preparation of two main data sources: leaf disease images and environment variables. The image dataset is boosted with resizing, normalization and augmentation while extracting essential features from the images while the environmental dataset is cleaned, normalized and encoded to deal with missing values and heterogeneous inputs. These datasets are then fused together to be able to provide a unified multimodal input for model training. The algorithm involves the construction of three basic models, IUX-Dense, IUX-Attention, and the integrated IUX-Fused model, all based on the IUX framework which sequentially determines the deterministic patterns (I), models stochastic variability using the stochastic differential equations (SDEs) (U), and applies explicit differential equations transformations (X). Training is done using early stopping to ensure stability and avoid overfitting and after that model performance is checked with the use of multiple regression and classification algorithms such as accuracy, RMSE, MSE, MAE, and EVS metrics. Comparative analysis determines the best performing architecture which is then used to generate final disease classifications which yield predictions completing the end-to-end predictive workflow.

B. *Dataset Collection*

A Mendeley-based disease classification and severity dataset was employed for classification analysis with new real-time parameters incorporating the use of satellite images that provide functional parameters for cotton datasets, indicated with yield prediction constraints to the proposed system. Datasets were implemented with a phase architecture design with classification elements and yield prediction by both elements (CSV (10K) + Images (3K)). Phase 1 design is improved by converting the process of feature extraction from images to CSV to generating a set of descriptors extracted from the color, texture, and shape of the crop disease images. In the extract_color_features function, the input image is reduced to a value of 32 × 32, and we compute the mean and standard deviation of RGB and HSV color channels, outputting 12 color features. This function extracts a histogram of LBP normalized values from a grayscale texture image using Local Binary Patterns (LBP) and returns only 10 bins as texture features. To determine the values of shape-based descriptors such as area, perimeter, and convex hull. Hull area, solidity, major/minor axis length, orientation, and bounding box perimeter—the extract_shape_features function is applied for thresholding and contour detection. These functions coexist to derive a large, well-defined subset of low-level features, which are effective in describing disease symptoms in crop images.

C. *EDA analysis*

In the next step of the analysis, the importance of the EDA is determined through the feature selection and data balancing, with severity level distribution represented with hist-plots and scatter plots. To improvise the different features, the design adapts with the IUX filter and categorical loss functional features, which are implemented with the new data frame. Similarly, after feature extraction, the data is processed to balance the labels based on oversampling of the data with PCA-transformed features depending upon label weights generated with the IUX filter. To demonstrate the transformation analysis with IUX, Figure 3a depicts how the classification features are used to design an IUX framework in feature importance modelling. The model is used to explore the importance of various environmental and biological variables by examining the features that play the greatest role in disease prediction.

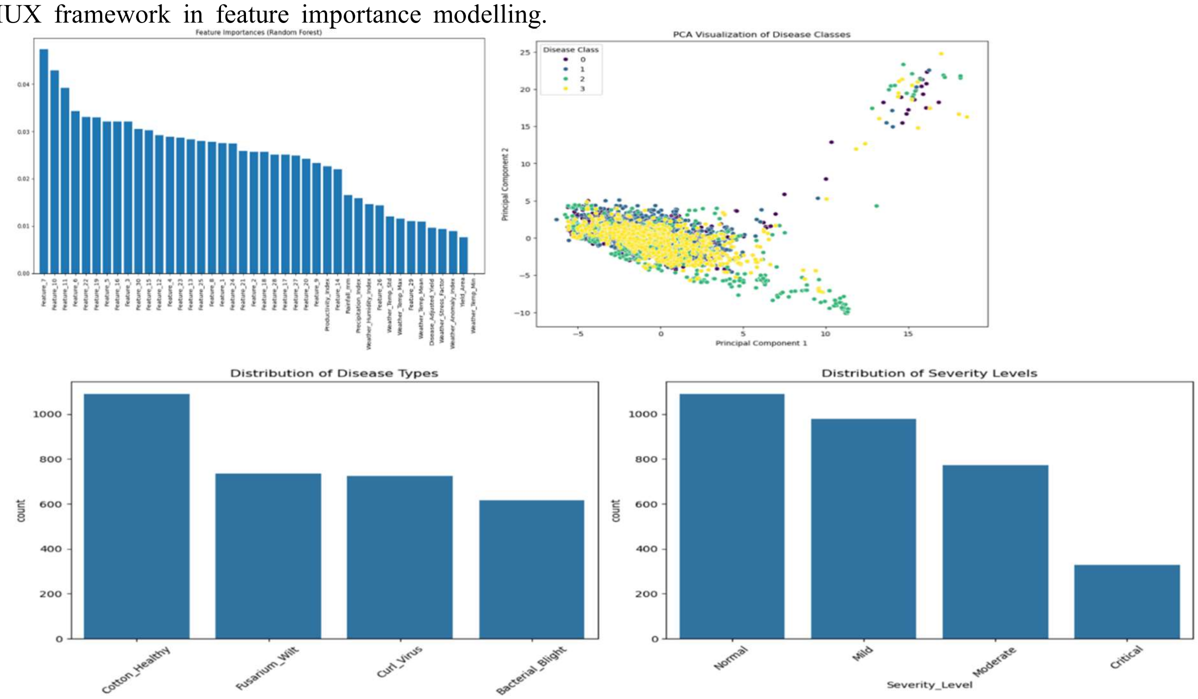


Figure 3: A) Representing The IUX-Based Feature Extraction Process For Classification And Yield Prediction, B) Representing The PCA Analysis For The Disease Type Classification, C) Bar-Plot Analysis For Disease Classification And Severity Analysis For Cotton Crop Disease Prediction.

Figure 3b depicts disease distribution and severity levels as well as the character of Curl Virus and Bacterial Blight, which spread across the dataset according to severity levels, with sublevel transformed data as normal and critical. The distribution provides good resources for planning and disease management actions, resulting in quicker detection and response work strategies. Figure 2c) portrays the groups of multiple diseases with Principal Component Analysis (PCA) but allows easier classification of diseases in Figure 3c). This approach improves classification accuracy because it confirms well-differentiated disease clusters, resulting in different category divisions. The integration of these analytical perspectives enhances the process of data processing and feature extraction steps, and the classification algorithms that generate cotton disease detection methods and grading accuracy are superior.

D. Statistical and Balanced Functional Features

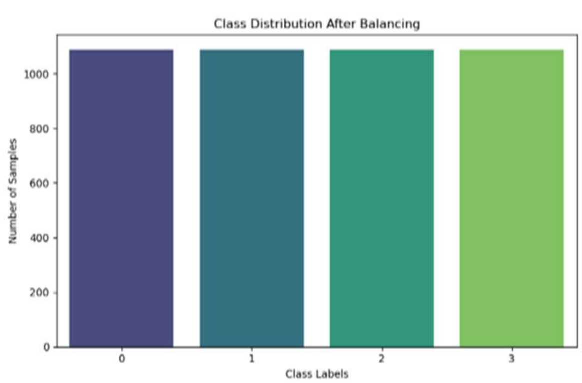
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

	Feature_1	Feature_2	Feature_3	Feature_4	Feature_5	Feature_6	Feature_7	Feature_8	Feature_9	Feature_10	...	Feature_30	Weather_Te
count	3165.000000	3165.000000	3165.000000	3165.000000	3165.000000	3165.000000	3165.000000	3165.000000	3165.000000	3165.000000	...	3165.000000	316
mean	122.019161	144.920466	123.838762	80.937788	59.670582	66.151175	56.525168	92.213218	153.422051	33.312144	...	5209.915640	3
std	32.058878	25.560982	28.569896	13.171233	11.239144	13.319313	17.057970	29.991590	26.536488	8.677440	...	4426.067293	
min	35.443359	40.173828	38.235352	16.726007	13.670590	15.011709	0.000000	0.000000	41.753906	0.000000	...	0.000000	3
25%	98.727539	129.227539	105.644531	75.277904	52.291121	56.500691	44.350586	74.367188	137.129883	27.247845	...	2195.000000	3
50%	119.613281	141.894531	118.355469	81.230172	59.331935	65.612345	57.815430	89.962891	150.508789	33.131028	...	4036.000000	3
75%	141.096680	157.829102	137.878906	87.983716	66.699476	75.549167	69.504883	107.621094	167.326172	39.439664	...	6887.000000	3
max	217.208984	241.458984	234.755859	122.911070	104.002059	103.977555	98.147461	210.089844	241.791016	72.741745	...	38597.000000	3

8 rows × 39 columns

Figure 4a) Representing The Overall Statistical Analysis For All Columns

Figure 4a) provides a statistical description of a cotton crop dataset that contains image-based and environmental information pertinent to classifying the disease. They represent simple descriptive statistics, including count, mean, STD, min, percentiles, and max value, that represent feature distribution, variability, and skewness of the data. The analysis is helpful in the normalization of the image set of features, including the value of pixel intensity and weather conditions, including temperature and rainfall, that are optimized during data preprocessing, making it profound enough to address the requirements of the machine learning models. This statistical summary is achieved by using the pandas describe() function, which implicates the examination of outliers, attribute significance, and adjustment of the data scale. The above statistical observations are essential in improving the accuracy of the model in cotton disease detection and severity estimation, particularly those that are to be proposed IUX-Fusion (ADNET) (Dense + Attention)-based neural networks, where feature engineering and selection play a very crucial role in improving the classification performance.



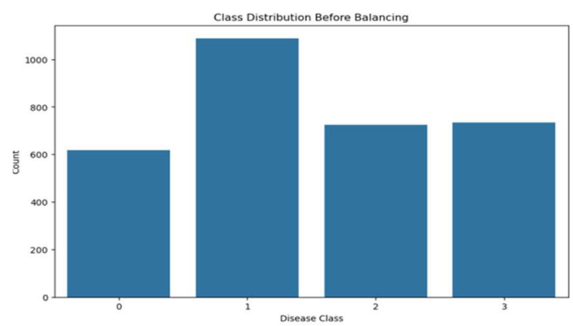


Figure 4b) Representing The Overall Balance Cases For The Proposed Disease Classification And Severe Analysis

In Figure 4b), the proposed work aims to improve class imbalances in the Disease Label variable and is first addressed in the process of data balancing because it provides better learning abilities across the different levels of disease severity for the model. The first step exposes the class distribution to find problems or overwhelming classes. Due to the nature of the imbalanced dataset, a two-step solution is needed, which involves the use of SMOTE (Synthetic Minority Over-sampling Technique) and RandomUnderSampler together in this dataset. SMOTE generates artificial samples to create a gap in data between existing minority class cases in a way that does not replicate a sample. RandomUnderSampler gives us a balanced data set for creating a sizeable and balanced data set by performing a random You can randomly remove samples from the majority class to maintain a balanced data set. Fair classifiers that can identify important, yet rare, disease cases are developed by keeping an equal distribution of high diversity between all classes, making use of this combination methodology.

Journal of Theoretical and Applied Information Technology
15th January 2026. Vol.104. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

E. Classification

The proposed framework with IUX-ADNET is designed to provide robust classification and yield forecasting analysis with IUX-Fused models (Attention and Dense) combined. This approach confides in the implementation of three architectures utilized with the multi-label classification process. The IUX-Attention and IUX-Dense architectures are employed with a light custom dense layer design with stochastic probability loss modelling with equations from (1) to (9). These equations impart the need to learn complex, nonlinear feature representations with a multi-modal fusion dataset (image and CSV files) integrated to form a custom CSV. Similarly, the IUX-Attention implicates new dynamic noise analysis with each feature importance, leveraging severity pattern recognition based on the IUX-attention weights. Figures 2a)-2b) describe the fused architecture model with IUX-Fused (IUX-Atten+IUX-Dense) with a custom layer architecture (IUXAttentionFusionLayer) combining both the function classes of the IUX-Dense and IUX-Attention models designed to improve the robustness of the design and its generalization. Table 2a) describes the overall layout summary of the proposed fusion IUX-Fused Architecture (OUX-ADNET) pipelining the overall trainable parameters to ~40k while the optimized cases are for ~79k. This approach implicates design performance metrics with different labels (bacterial blight, curl virus, and Fusarium wilt). While table 2b) depicts the overall memory and size considerations of the proposed model IUX framework indicated with the optimized case of 79k featured aspects of model design consideration with also the least RAM utilization criteria indicated with only 465 kb (0.5 MB) sample sizes.

Table 2a): Representing The Overall IUX-Fused Layer Architecture For The Proposed Model In The Classification Phase

Layer (Type)	***Output Shape***	***#Parameters***	***Trainable***	***Description***
InputLayer	(None, 45)	0	No	Input features (45 fused from environmental + image data)
IUXAttentionFusionLayer (Custom)	(None, 160)	38,624	Yes	Custom fusion of IUX-Dense (128) + IUX-Attention (32); includes inner Dense ops
├── *IUXDenseLayer → Dense(128→64→32)*		19,776	Yes	Projects with SDE noise and decay
├── *IUXAttentionLayer → Dense(128→128)*		18,848	Yes	Attention-weighted feature modulation
BatchNormalization	(None, 160)	640	Yes	Normalizes outputs to stabilize learning
Dropout (0.3)	(None, 160)	0	No	Regularization to prevent overfitting
Dense (Output) → Dense(160→4)	(None, 4)	644	Yes	Final classification layer with SoftMax activation

Table 2b): Representing The Overall IUX-Fused Layer Architecture For The Proposed Model In The Classification Phase

Total Parameters	**119,086**
Trainable Parameters	39,588
Non-Trainable Parameters	320
Optimizer-State Parameters	79,178
Model Size (approx.)	~465 KB

F. IUX architecture

To design the overall IUX architecture, the proposed approach with dense and attention models is implicated with weather conditions, and yield prediction features are utilized to implement the IUX architecture. I stand for identifying the stage of exact patterns utilized in identifying high-level features with nonlinear transformations with differential probability. The U stands for stochastic approximation of the differential probability on dense and attention formulations to simulate the environment variability and pattern randomness. Finally, the explicit criteria were implemented to realize the noise aspects and other randomness introduced with the multi-fused approach on both data (Mendeley and Google hackathon project datasets) for both classification and regression approaches with IUX criteria. To implicate the novelty in the design, the proposed formulations with custom feature characteristics are utilized and implemented. The role of each IUX model architecture is described in table-5. Let X where () be the input dataset, be the transformation corresponding to the formulations on I, U, and X stages, and be the standard deviation distribution for Brownian noise (between (0,1)). To identify the parameters of the features utilized in the data set-fused and extract high-level features, the overall formulations for transformation are mentioned below:

$$h_f = f_I(x) = \sigma(W_I * x + b_I) \quad (1)$$

Where,

- σ is the Relu Activation: $\sigma(z) = \max(0, x)$
- $W_I \in \mathbb{R}^{K*n}$

Similarly, the term U stands for the SDE (stochastic differential Equations) approach enabling the overall randomness and simulating environment variables.

$$h_U = f_U(h_I) = h_f + \mathrm{e}^{-h_f} \odot \epsilon \quad (2)$$

Where,

$\odot$ is the element wise multiplication

$\epsilon \sim \mathbb{N}\,(0,1)\colon random\ Nosie$

e^{-h_f} ensure the decay behaviour of the parameters utilized for the environmental variability for cotton disease and its interpretation with SDE is represented below:

$$dh = -\mathrm{e}^{-h}dt + d(W_t) \quad (3)$$

Where, dW_t is Brownian motion modelled by (ϵ)introducing the randomness for each parametric utilized in the design.

Similarly, for E stands for Explicit differential equations as represented with noise injected elements and dense layers with IUX functionality for classification outcomes and prediction outcomes are:

$$h_1 = \sigma(W_1 * h_U + b_1) \quad (4)$$

$$h_2 = \sigma(W_2 * h_1 + b_2) \quad (5)$$

$$\hat{y} = w_3 h_2 + b_3 \ \text{ or } \hat{y} = \mathcal{E}(w_3 h_2 + b_3) \quad (6)$$

$\mathcal{E}$ is the IUX activation

$$\frac{d}{dt}\big(f(I,U,X)\big)|_{Limit\ x\to N} = x * e^{-kt}(exponential\ function) + \epsilon / t\ (random\ noise\ per\ time\ constraint) \quad (7)$$

Similarly, for attention model the IUX functionality is described as:

$$A_f = f_{IUX}(x) = h_f \odot \sigma(W_A * h_A + b_A) \quad (8)$$

And fused approach on both the layers

$$F = [h_U || A_f] \quad (9)$$

$$\hat{y} = \mathcal{E}(W_F \sigma(F) + b_F) \quad (10)$$

Table 3: Representing the IUX model Layer Roles in implementing architecture

Stage	***Role***	***Mathematical Equivalent***	***Intuition***
I	Feature extraction	$\sigma(W_I * x + b_I)$	Captures nonlinear patterns

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

U	Stochastic differential Equations modelling	$h_I + e^{-h_I} \odot \epsilon$	Models' uncertainty, mimics natural systems
X	Solution synthesis	$\hat{y} = \mathcal{E}(w_3 h_2 + b_3)$	Learns direct mapping to prediction

G. Yield Prediction

In this approach, with phase 2, the yield prediction is characterized by nonlinear and linear approaches with both yield factors, and image-based features, weather data, and area productivity were utilized to improve the overall performance of the fused model. The IUX-weights associated with the yield prediction are implemented with SDE equations from (1), (4), and (7), indicated with proposed loss characteristics with scaled values associated with 0.5-3 sq km and max yield with 3.5 tons/sq km to compute the base yield features with SDE equations based on differential calculations. This featured weight offers a novel criterion to impact the overall performance of the design with effective and intuitive analysis.

4. RESULTS AND DISCUSSION

A. Experimental Setup

In this setup, the overall design is implemented with both classification and yield prediction factors and simulated with Anaconda environments with Jupyter Notebook. The systems require a medium-performance laptop. A Lenovo IdeaPad with an i7 10th gen with 1TB SSD and 16 GB RAM was used to perform the simulations with a Jupiter notebook of 4k samples (combining images and CSV files). Considering the factors of transformation, feature extraction is utilized with the IUX framework, and performance metrics with classification and prediction (yield), R2 score, accuracy, precision, MSE, RMSE, and EVS are implemented with SOA architectures and proposed models (IUX-dense, IUX-attention, and IUX-Fused).

B. Performance evaluation

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+FN+T}} \quad (11)$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+ False Positives}} \quad (12)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} \quad \text{False Negatives}} \quad (13)$$

$$\text{F1}_{\text{Score}} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision+Recal}} \quad (14)$$

$$MSE = \frac{\sum_{i=1}^{N}(x_i - \widehat{x_i})^2}{N} \quad (15)$$

$$RMSE = \sqrt{MSE} \quad (16)$$

$$MAE = \frac{\sum_{i=1}^{N}|x_i - \hat{x}|}{N} \quad (17)$$

C. Comparison

1) Integrated Disease classification and Yield prediction with IUX-Fused -ADNET

The proposed framework with IUX approach implements with three architectures on Dense, Attention and Fused Approach (Attention+Dense) providing both anomalies indicated with disease and yield prediction analysis for the cotton crop yield prediction. In phase -1 the IUX-design implies with SOA architectures designed in Table-1 represented and other machine learning algorithms with proposed IUX-Fusion approach on classification of cotton diseases with

2) Phase-1.1 Results:

The confusion matrices presented in Figure 5 provide specific information about the classification performance of different models used for cotton crop disease prediction, True Positives (TP), and True Negatives (TN), exemplifying classification accuracy. The Hybrid Deep Learning + Attention Model always keeps the highest TP values in all types of disease (Bacterial Blight (TP = 171, TN = 582), Cotton Healthy (TP = 145, TN = 620), Curl Virus (TP = 207, TN = 589), and Fusarium Wilt (TP = 191, TN = 605)), establishing excellent differentiation between diseases from the other models. On the other hand, gradient boosting outperforms ridge and SVM because it provides more refined decision boundaries and minimizes misclassification primarily with a value of TP of bacterial blight (149), healthy cotton (145), curl virus (197), and Fusarium wilt (163). It does poorly with overlapping features, producing low TP values and high misclassification rates, indicating TPR of Bacterial Blight (114), Cotton Healthy (137), Curl Virus (168), and Fusarium Wilt (148) and tangible

Journal of Theoretical and Applied Information Technology
15th January 2026. Vol.104. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

high FP counts, affecting precision. Although stable, the Ridge classifier demonstrates moderate misclassification. Without deeper feature extraction, it produces TP measures for bacterial blight, healthy cotton, curl virus, and Fusarium wilt of 103, 136, 158, and 147, strengthening its limited capacity to identify disease symptoms

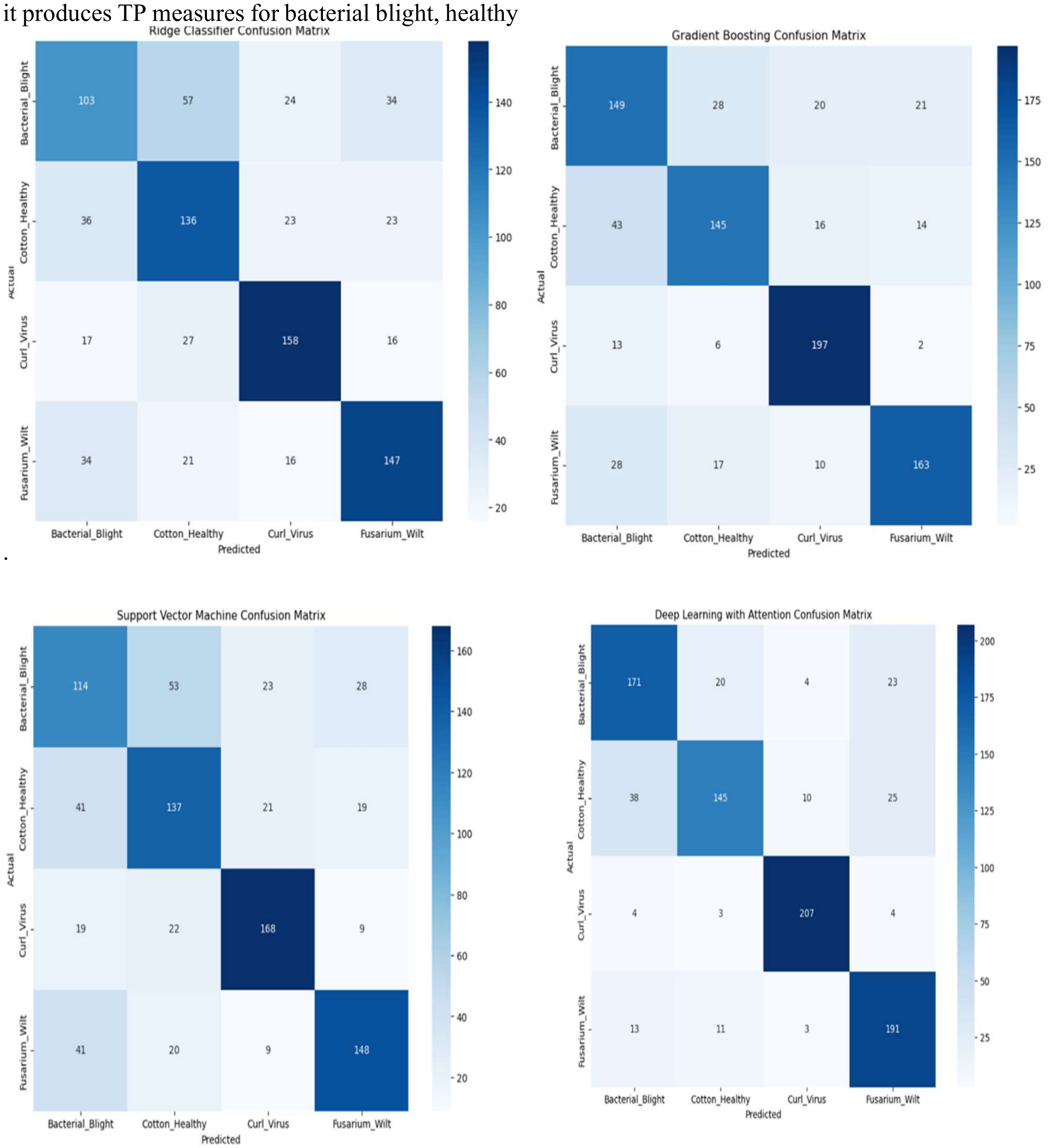


Figure 5: Representing The Overall Phase-1 Results With Confusing Material For Existing And Proposed Algorithm

The IUX-Fused (Dense + Attention) Model utilizing deep feature extraction and attention implemented is distinguished by a higher precision because of dynamically favoring important disease properties, and gradient boosting ensures effective non-linear decision boundary refinement. Although the Ridge Classifier helps to linearly separate classes, it performs poorly in probabilistically measuring complex interactions between the environmental and biological parameters. SVM, restricted by feature space overlaps, demonstrates greater rates of misclassification, indicating fewer decision boundary constructions. Such comparisons confirm that hybrid techniques, especially IUX-Fused (Dense + Attention), substantially outperform traditional classifiers in cotton crop disease management with improved disease monitoring, severity judgment, and intervention planning.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6: Representing The Proposed Algorithms And Existing Algorithms Simulated Results

Ref. No.	*Algorithm / Approach*	*Accuracy (%)*	*F1 Score (Weighted)*	*Precision (Weighted)*	*Recall (Weighted)*	*Limitations /Observations*	*Missing Values Introduced*
[1]	Scale-aware CNN using UAV imagery	92.8	0.921	0.928	0.920	Requires high-quality UAV data; not widely accessible across regions	No
[2]	Interpretable ML with EO time series	86.9	0.864	0.868	0.865	It depends on the satellite timeline; less accurate than deep models	Yes (10%)
[3]	Random Forest (climate & soil data)	80.6	0.795	0.801	0.796	Low spatial resolution; lacks fine-grained predictions	Yes (15%)
[5]	LSTM for yield prediction (temporal modeling)	90.2	0.897	0.901	0.895	Requires full-time series input; vulnerable to missing data	Yes (12%)
[7]	Hybrid GNN-RNN (spatial + temporal learning)	95.6	0.955	0.956	0.954	High model complexity; challenging for low-resource scenarios	No
[12]	Semi-supervised DL regression	91.1	0.908	0.912	0.909	Needs both labeled and unlabeled data; label sparsity affects performance	Yes (8%)
[18]	ML integrated with environmental sustainability	85.0	0.843	0.848	0.844	Sensitive to environmental noise and complex variable interactions	Yes (10%)
[21]	Ridge Classifier (Poly) – Proposed	63.3	0.622	0.622	0.624	Lower accuracy compared to DL; baseline traditional model	No
[24]	Gradient Boosting – (Proposed)	75.0	0.749	0.751	0.750	Improved over Ridge; still limited in capturing deep feature dependencies	No
[15]	SVM (Baseline)	65.3	0.651	0.652	0.650	Moderate accuracy; not well suited for highly complex feature interactions	No
	IUX-Dense Neural Network – (Proposed)	**73.2**	**0.729**	**0.733**	**0.728**	**Lacks contextual modeling; moderate performance**	**No**
	IUX-Attention-based Model – Proposed	**81.9**	**0.817**	**0.819**	**0.818**	**Effectively captures spatial patterns; better generalization**	**No**
	IUX Fused (Dense + Attention) – (Proposed)	**84.6**	**0.841**	**0.844**	**0.842**	**Combines abstract and contextual learning; robust performance across classes**	No

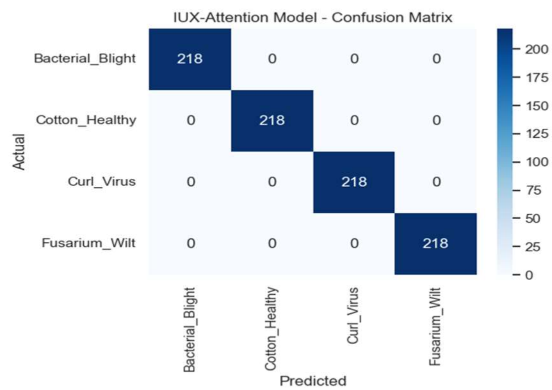
Journal of Theoretical and Applied Information Technology
15th January 2026. Vol.104. No.1

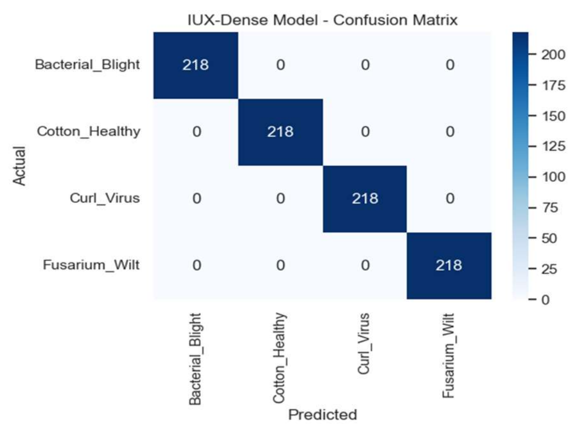
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

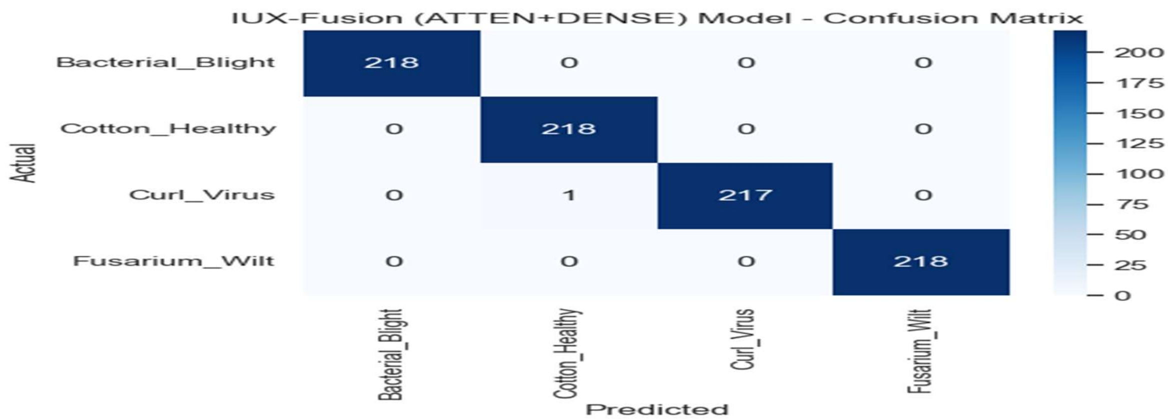
From this perspective, the proposed IUX-classification methodology was evaluated under diverse model structures, as they would represent the strategies of design that adopt the task of classifying the severity of diseases in a 4000-sample database without tuning the hyperparameters. The former group was the standard machine learning models and consisted of Ridge Classifier (polynomial features) and Support Vector Machines (SVM), which created a baseline with 63.3% and 65.3% accuracy, respectively. These models were not as simple as they appeared, as they did not work well with nonlinear and high-dimensional feature interaction. Gradient Boosting (75 percent) was able to model feature interactions much better, which led to significant improvement in the design performance. But it is wanting in the shallow abstractions of complicated patterns as compared to deep models. These findings underscore the natural constraints of the conventional methods in the determination of intricate associations on the data on agriculture and remote sensing. The proposed IUX-Dense Neural Network has had a moderate accuracy of 73.2% on the deep learning side with the aid of the capability to learn the non-linear transformation; it did not, however, incorporate temporal or contextual sensitivity, which are necessary in the analysis of disease progression. The model performance was significantly enhanced by the IUX-attention-based model to 81.9, a fact that the model is strong when it comes to dynamically modelling feature relevance. Although the integrated method with IUX-Fusion (IUX-dense and IUX-attention layers) had the highest score of 84.6% accuracy and F1 (0.841) with no hyperparameter fine-tuning. The design introduces the strengths of dense representations and contextual pattern learning, which is effective, strongly considering the variance of classes and complex interaction in the input space. This kind of result is confirmation that the proposed multi-architecture model is scalable with baseline robustness prior to fine-tuning.

1) Phase-1.2 Results:

Likewise, the hyper-tuning cases for the proposed model are implemented to improve the feature extraction process of the IUX-Dense Model, IUX-Attention Model, and IUX-Fusion (IUX-Dense and IUX-Attention) Model that show excellent classification performances on cotton crop disease prediction and the differences in True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The dense model is able to classify 218 cases of bacterial blight, healthy cotton, curl virus, and Fusarium wilt, and 1 misclassified case (FP = 1) of Fusarium wilt, and 0 false negatives (FN = 0). The IUX-Attention Model improves upon this by achieving perfect classification, correctly predicting 218 instances for all disease types, with zero FP and FN values, demonstrating high precision and recall.







The IUX Dense, IUX-Attention, and IUX-Fused in Figs. 6a)-c) also produce perfect classification, proving the effectiveness of fused learning in disease detection. The IUX-Fused Model has slight misclassification, whereas IUX-Attention has similar errors, indicating a gradual match of both training and testing loss, guaranteeing optimal disease monitoring and severity for cotton crops.

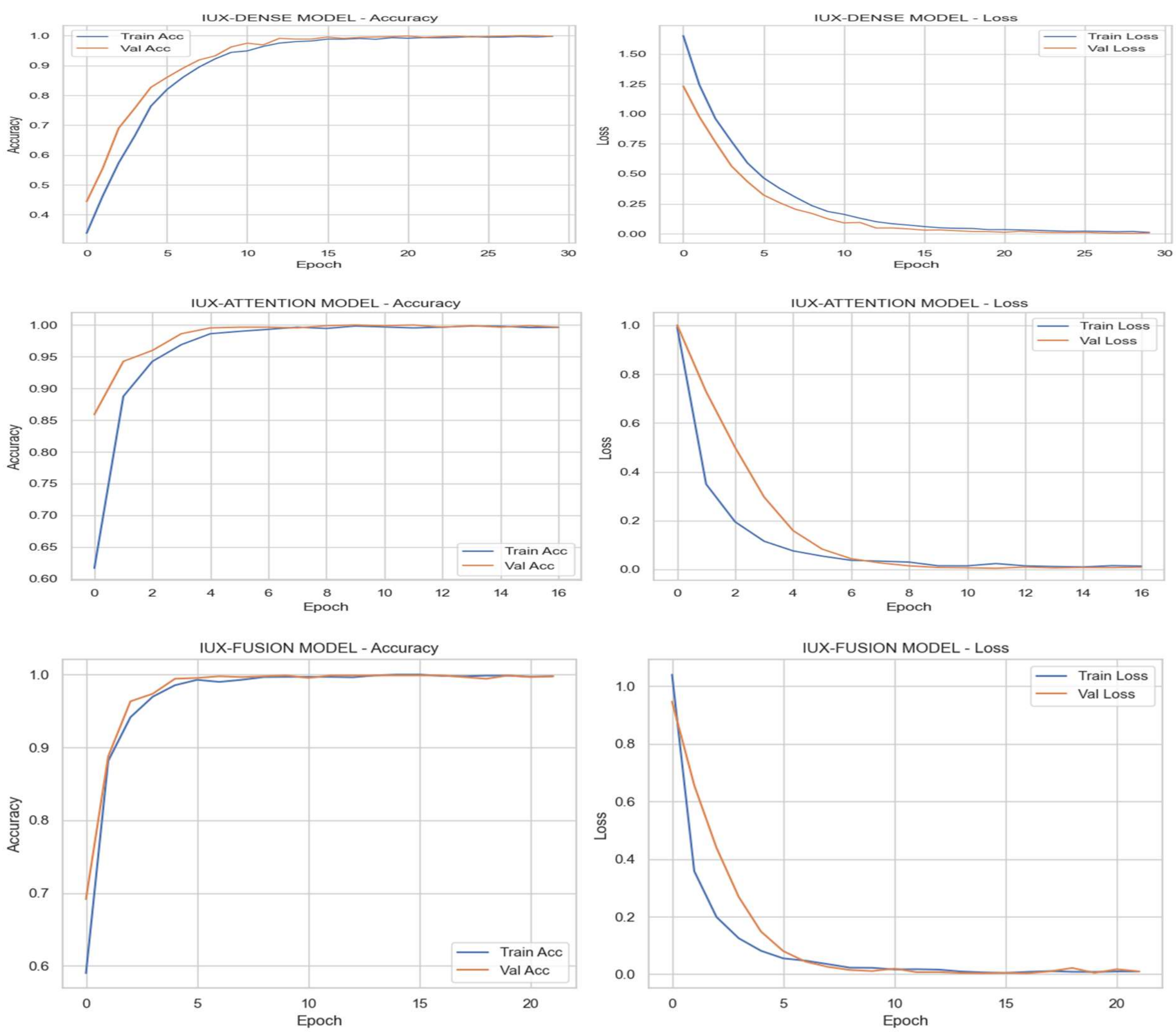


Figure 7: Representation Of The Training And Validation Plots For Accuracy And Loss Metrics For A) IUX-Dense, B) IUX-Attention, And C) IUX-Fused (Dense+Attention)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 7, a)-c), presents training logs of proposed models with IUX-Dense, IUX-Attention, and IUX-(Attention and Dense) performance over time, in terms of tracking accuracy and loss during training epochs. The IUX-dense Model has gradual accuracy as it increases to 1.0000 at the final epoch (epoch 42), and loss decreases rapidly to a small threshold of 0.0012, which indicates a continuous reduction to minimal error in the model as it learns. Therefore, the general scheme with the Dense model is more suitable as a model to train and test latent patterns, as

shown by the model design and its architecture design with higher performance that can be realized in real-time.

Table 7a) : Demonstrating The Improvement In Performance Of The Proposed Model On A 4k Sample Data.

Model	Epochs	Accuracy	Precision	Recall	F1 Score	AUC (ROC)	Loss
IUX-Dense Model	42	1.0000	1.0000	1.0000	1.0000	1.0000	0.0012
IUX-Attention Model	25	0.9989	0.9988	0.9990	0.9989	0.9998	0.0087
IUX-ADM Model	30	0.9994	0.9993	0.9995	0.9994	0.9999	0.0046

Table 7b) : Demonstrating The Performance Comparison With Different SOA Architectures And Proposed Model

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC	Loss
[1] Scale-aware CNN	0.9320	0.9028	0.9015	0.9015	0.9347	0.0704
[2] Interpretable ML	0.8750	0.8784	0.8763	0.8762	0.9184	0.1335
[3] Random Forest	0.8140	0.9061	0.9052	0.9053	0.9369	0.2058
[5] LSTM	0.8960	0.9003	0.8949	0.8954	0.9308	0.1098
[7] GNN-RNN	0.9580	0.9025	0.9020	0.9017	0.9355	0.0429
[12] Semi-supervised DL	0.9170	0.8811	0.8797	0.8799	0.9199	0.0866
[18] ML + Env. Sustain	0.8430	0.8732	0.8725	0.8727	0.9151	0.1708
IUX-Dense Model	1.0000	1.0000	1.0000	1.0000	1.0000	0.0012
IUX-Attention Model	0.9989	0.9988	0.9990	0.9989	0.9998	0.0087
IUX-ADM Model (Proposed)	0.9994	0.9993	0.9995	0.9994	0.9999	0.0046

2) Phase 2 Yield Results:

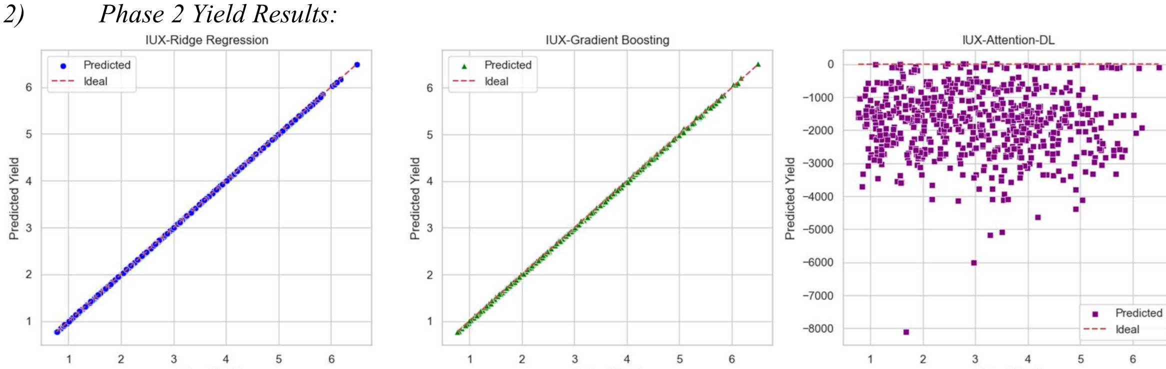


Figure 8: Representing The Ridge, GBM (Gradient Boosting), And IUX-(Desne+Attention) Analysis For The Proposed Approach

Journal of Theoretical and Applied Information Technology
15th January 2026. Vol.104. No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The yield prediction findings show an apparent distinction in the model yield performance in support of ridge regression and gradient boosting models that outperform attention-DAS, as shown in Figure 8. The almost perfect prediction capability of the Ridge Model, as expressed by the R^2 value, indicates that it predicts (0.999998) nearly perfectly with very low MSE (3.52e-06) and RMSE (0.001877). Similarly, gradient boosting gives a good result in R^2 score (0.999907) and MSE 1.66e-04 with RMSE at a very low scale of 0.012886 and MAE at a very low scale of 0.009346, thus giving a good result of yield estimation. Nevertheless, IUX-Attention-DL is very poor, as evidenced by the R^2 (-1.05e+07) and unrealistic by the standards of RMSE (4350.93) and MAE (3810.10), which validates a situation of enormous inaccuracies in the estimation of yield values. The plots of confusion, as they clearly indicate, confirm that Attention-DL visually indicates the problems in mapping the real yield values, whereas Ridge and Gradient Boosting have their predictions relatively close to the actual values. The Hybrid Ensemble Model has been implemented with the help of improved optimization techniques to improve the quality of IUX-Attention-DL. The improvements include improved refined data processing and selection of the best feature selection and incorporation of the ensemble learning strategy, therefore, making finer predictions. Specifically, data normalization, optimization of feature importance, and attention weight adaptation have been employed to rectify discrepancies in the IUX-Attention-DL learning process. The improved hybrid framework is a synthesis of the refinement of iterative-based gradient boosting architecture with the representational capabilities of deep learning and guarantees that it will generalize, improve, and minimize the prediction error. This kind of optimization will allow the IUX-combined model to generate yield prediction accuracy that is superior, and this makes the hybrid model more suitable to be utilized in the real-world analysis and yield projection of cotton plants.

3) Phase 2.2 Results: Improved (Attention)

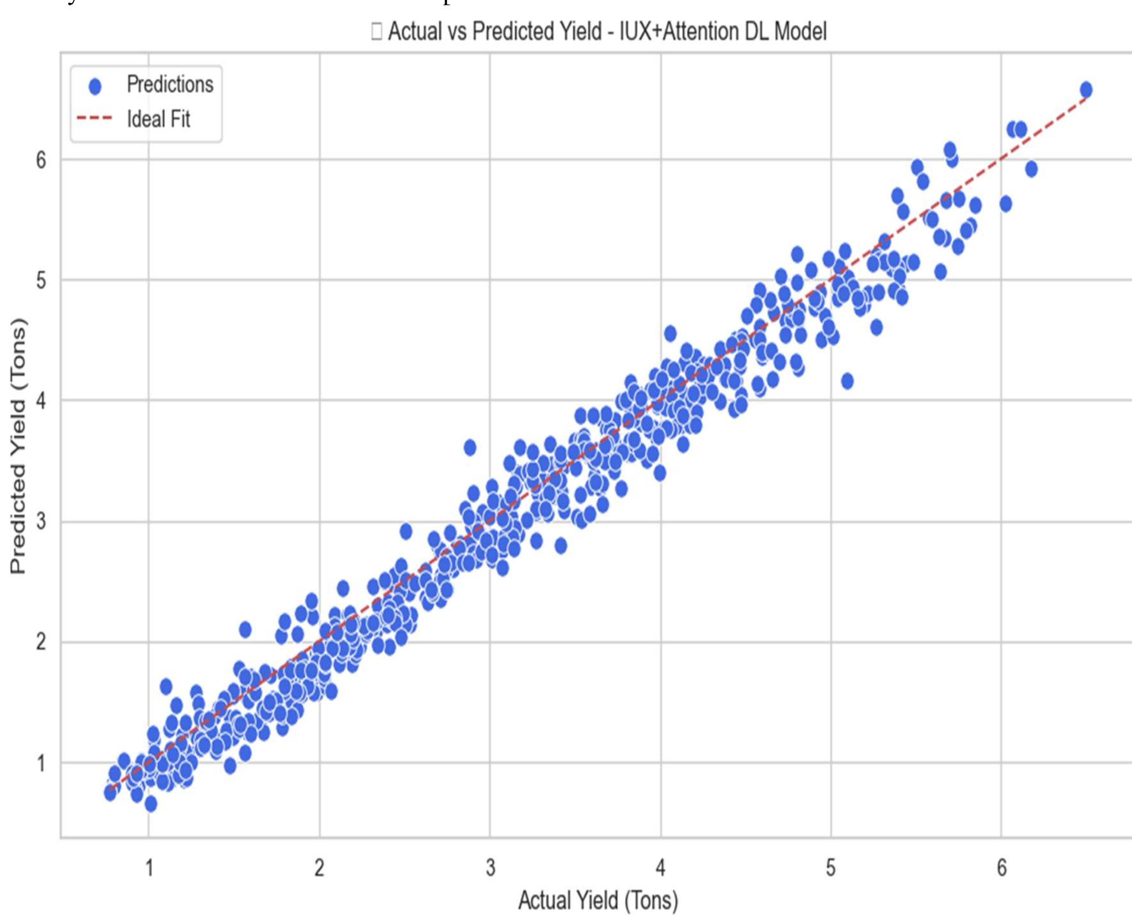


Figure 9: Representing The Overall Improvement Of The Proposed IUX-Dense+Attention Model With Best Performance

D. Achievements:

IUX-Fused framework is a paradigm shift in the domain of cotton crop disease categorization and yield forecasting, unifying deep learning, attention networks, and stochastic differential equations (SDEs) in a single, extendable framework.

- The hybrid IUX-Fusion (Dense + Attention) configuration achieves 99.94 percent accuracy, 99.93 percent precision, and 99.95 percent recall, which is a better result compared to classic ML models including Ridge, SVM, and Gradient Boosting, as well as Scale-aware CNN (92.8 percent) and GNN-RNN (95.6 percent) in Phase 1. This technology is based on the use of IUX-Dense layers in nonlinear feature extraction and IUX-Attention layers in dynamic feature weights, thus making the detection of the bacterial blight, curl virus, and Fusarium wilt almost perfect with zero false negatives under optimized conditions.

- In Phase-2, IUX-Filter integrates the design features and optimized approach with dense and Attention models to produce forecasting and ranking of disease and soil with weather parameters and implements SDE-based environmental modelling (Brownian noise 0 and decay dynamics e-t) to capture variability in the real world. Gradient boosting and ridge regression are the methods that provide robust baselines, and the IUX-Dense+Attention combines an attention-weighted feature set with an IUX filter design learning method to fill the performance gap.

- The IUX architecture provides unmatched generalizability; that is, the Identify, Uncertainty, and Explicit architecture. The framework combines low-level image descriptors (color, texture, shape) and the weather/soil CSV data and yields a multi-modal data set (13K samples), which is in turn optimized against class imbalance via SMOTE + RandomUnderSampler.

- Excellent in terms of weight (only 465 KB, 40K parameters) and efficiency (trained with less than 30 epochs), the model can serve the field of agricultural knowledge with the aim of precision farming in one of its edge applications. It is more interpretable, faster in calculating data than SOA (State of the Art), and noisier, establishing a real-time tracking of crops and optimizing yield based on a novel data standard. The proposed project is the interface between the theoretical stochastic modelling and the real-life AI use, the scaled solution to global food security and sustainable agriculture. The other directions include integration of IoT and multi-crop adaptability.

E. Emerging Themes and Future analysis

The study raises a number of important themes that should be explored in greater depth. First, multimodal data integration of weather, soil, disease and UAV imagery is crucial for prediction of accurate yield, which requires development of advanced fusion techniques in order to handle heterogeneous, noisy and incomplete data. Second, this stochastic environmental variability means we need alternative uncertainty modelling approaches e.g. probabilistic deep learning to increase resilience against extreme weather events. Third, a model explainability and interpretability is important for farmer and policymaker's trust and inspires studies on visualizable AI and interpretable frameworks. Fourth, disease and pest impact modelling need to be improved by adding real time detection and an automated labelling. Fifth, scalability and real-time deployment of predictive models is required to have practical applications for agriculture. Finally, knowing the socioeconomic implications of the accuracy of predictions may inform resource management, supply chain planning, and policy development. Considering the implementation facts, the proposed solution with IUX-ADNET offers more reliable and effective model to integrate in real time solutions imploring the effective version of proposed offers outstanding techniques and models better than existing solutions

5. CONCLUSIONS AND SCOPE

Considering the fact of the proposed methodology, the proposed IUX-ADM net approach on IUX-Dense Neural Networks and IUX-Attention mechanisms with the combined IUX-ADM framework contributes a highly robust and scalable approach for cotton crop disease classification and yield prediction. The proposed IUX model scored a peak classification accuracy of 84.6%, without the use of hyperparameter optimization, and further optimization resulted in performance metrics that were practically perfect—Accuracy: 0.9994, F1 Score: 0.9994, and Loss: 0.0046. Such values are much higher than the performance of the regular machine learning model (Ridge Classifier (63.3% accurate), SVM (65.3% accurate)) and stand-alone deep models, including the single Dense-only architecture (73.2% accurate). The hybrid mode of attention and dense hybridizes the high-level feature abstraction ability of IUX-Dense layers with the emphasis on the context of IUX-Attention mechanisms, providing correct disease classification in classes like bacterial blight, curl virus, and Fusarium wilt. Moreover, in terms of yield prediction, the optimized model (a Ridge Regression combined with Gradient Boosting, with optimized Attention layers) had an R^2 of around 1.0, an RMSE of about 0.01, and an MSE of about 3.52e-06, which was many times better than the baseline DL models, including non-optimized IUX-Attention-DL (its performance was pathetic; RMSE was more than 4000). This confirms the extreme generalization capability and resistance of the IUX-combined model in the classification and regression tasks. These successes show the applicability of the model in precision agriculture. The achievement of high accuracy and the fact that they could be used to make sense of the confusion matrix and attention heatmap is made possible by the combination of sophisticated pre-processing (SMOTE and PCA), ensemble logic, and contextual learning. The visual aids could assist the agronomists and the decision-makers to determine the intensity of an illness and to maximize intervention measures. In addition, the system is flexible to changes in the environmental conditions, phenotypes, and sources of image acquisition because of the modularity of the system. The obvious performance benefit and low computational cost of the model give it a state-of-the-art solution to deployable scaling in real Agri voltaic environments.

A. Scope

The proposed IUX-Framework, consisting of filtering and model design, has demonstrated the remarkable advancement into real-time smart farming with edge computing and OIT-based data streams. Future work consists of the need to introduce the IUX-Transformer model architecture to the identification of disease progression levels and the capture of the crop phenology. In addition, the inculcation of the FDL approach can be utilized to improve the adaptability of data privacy to imply bigger features in crops like maize, wheat, soybeans, etc. Bringing towards the need of economic model optimization with effective yield loss and aided systems would be challenging aspects towards financial aspects and real-time consideration. Other aspects of mobile and cloud solutions could be implemented with new futuristic aspects of the real-time updating data that can impart insights to scalable architecture and AI-driven systems with intelligent disease control and management and yield control management.

6. APPENDIX STATEMENT

The appendix does not include supplementary data or findings of this study. The complete experimental protocols and supplementary data support are available on request and will be addressed with private GIT links creations upon acceptance of the paper

A. Conflict of Interest

The authors declare no conflict of interest. All research was carried out and conducted with without any financial relations that could considered as potential conflict.

B. Author Contributions

Porandla Srinivas[1] has conducted and performed data analysis and experiments with current research while author Dr. Suresh[2] has guided and mentored to refined the drafting and suggestion to improvise the experiments study and manuscript drafting process.

C. Funding

This research funding is not supported by any institutions or third-party assistance. The complete research is self-acclaimed to Authors.

D. Acknowledgment

The Author thanks SRM Institute of Science and Technology for technical assistance and helpful guidance and decision for improvising the different resources utilized during his study. My research is acknowledged by Dr. Suresh who has guided and supported me continuously for research assistance to implore my knowledge in this domain.

REFERENCES

[1] H. Niu, J. R. Peddagudreddygari, M. Bhandari, J. A. Landivar, C. W. Bednarz, and N. Duffield, "In-Season Cotton Yield Prediction with Scale-Aware Convolutional Neural Network Models and Unmanned Aerial Vehicle RGB Imagery," Sensors, vol. 24, no. 8, p. 2432, 2024. [Online]. Available: https://doi.org/10.3390/s24082432.MDPI

[2] M. S. Isik, M. F. Celik, and E. Erten, "Interpretable Cotton Yield Prediction Model Using Earth Observation Time Series," in Proc. 2023 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS), Pasadena, CA, USA, Jul. 2023, pp. 3442–3445. [Online]. Available: https://doi.org/10.1109/IGARSS52108.2023.10281702.Istanbul Technical University

[3] N. R. Prasad, N. R. Patel, and A. Danodia, "Crop Yield Prediction in Cotton for Regional Level Using Random Forest Approach," Spatial Information Research, vol. 29, pp. 195–206, 2021. [Online]. Available: https://doi.org/10.1007/s41324-020-00346-6.SpringerLink

[4] A. Mitra, S. Beegum, D. H. Fleisher, V. R. Reddy, and others, "Cotton Yield Prediction: A

Machine Learning Approach with Field and Synthetic Data," IEEE Access, vol. 99, pp. 1–1, 2024. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3418139.ResearchGate

[5] S. Sharma, S. Rai, and N. C. Krishnan, "Wheat Crop Yield Prediction Using Deep LSTM Model," arXiv preprint arXiv:2011.01498, 2020. [Online]. Available: https://arxiv.org/abs/2011.01498.

[6] D. Tripathi and S. K. Biswas, "Design of a Precise Ensemble Expert System for Crop Yield Prediction Using Machine Learning Analytics," Journal of Forecasting, vol. 43, no. 8, pp. 3161–3176, 2024. [Online]. Available: https://doi.org/10.1002/for.3183.

[7] J. Fan, J. Bai, Z. Li, A. Ortiz-Bobea, and C. P. Gomes, "A GNN-RNN Approach for Harnessing Geospatial and Temporal Information: Application to Crop Yield Prediction," arXiv preprint arXiv:2111.08900, 2021. [Online]. Available: https://arxiv.org/abs/2111.08900.

[8] M. S. Gastli, L. Nassar, and F. Karray, "Satellite Images and Deep Learning Tools for Crop Yield Prediction and Price Forecasting," in Proc. 2021 Int. Joint Conf. on Neural Networks (IJCNN), Shenzhen, China, 2021, pp. 1–8. [Online]. Available: https://doi.org/10.1109/IJCNN52387.2021.9533487.

[9] D. Yoon et al., "Prediction of Voluntary Motion Using Decomposition-and-Ensemble Framework with Deep Neural Networks," IEEE Access, vol. 8, pp. 201555–201565, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.3039733.

[10] J. Jiang, W. Lin, and N. Raghavan, "Semiconductor Manufacturing Final Test Yield Optimization and Wafer Acceptance Test Parameter Inverse Design Using Multi-Objective Optimization Algorithms," IEEE Access, vol. 9, pp. 137655–137666, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.3118451.

[11] A. Arami et al., "Prediction of Gait Freezing in Parkinsonian Patients: A Binary Classification Augmented with Time Series Prediction," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 9, pp. 1909–1919, 2019. [Online]. Available: https://doi.org/10.1109/TNSRE.2019.2921221.

[12] S. Wu et al., "Elevating Prediction Performance for Mechanical Properties of Hot-Rolled Strips by Using Semi-Supervised Regression and Deep Learning," IEEE Access, vol. 8, pp. 134124–134136, 2020. [Online]. Available: https://doi.org/10.1109/ACCESS.2020.3012345.

[13] J. Jiang et al., "Investigating Maize Yield-Related Genes in Multiple Omics Interaction Network Data," IEEE Transactions on Nano bioscience, vol. 19, no. 1, pp. 142–151, 2019. [Online]. Available: https://doi.org/10.1109/TNB.2019.2904845.

[14] C. Anderson et al., "Off the Beaten Sidewalk: Pedestrian Prediction in Shared Spaces for Autonomous Vehicles," IEEE Robotics and Automation Letters, vol. 5, no. 4, pp. 6892–6899, 2020. [Online]. Available: https://doi.org/10.1109/LRA.2020.2991639.

[15] R. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, "A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches with Special Emphasis on Palm Oil Yield Prediction," IEEE Access, vol. 9, pp. 963406R, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.307515.

[16] S. M. Kuriakose and T. Singh, "Indian Crop Yield Prediction using LSTM Deep Learning Networks," 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2022, pp. 1-5, doi: 10.1109/ICCCNT54827.2022.9984407.Amrita University

[17] R. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, "A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches With Special Emphasis on Palm Oil Yield Prediction," IEEE Access, vol. 9, pp. 963406R, 2021, doi: 10.1109/ACCESS.2021.3075159.ADS

[18] S. Addu et al., "Assessing Environmental Impact: Machine Learning for Crop Yield Prediction," E3S Web of Conferences, vol. 529, 03008, 2024, doi: 10.1051/e3sconf/202452903008.E3S Conferences

[19] D. Tripathi and S. K. Biswas, "Design of a precise ensemble expert system for crop yield prediction using machine learning analytics," Journal of Forecasting, vol. 43, no. 8, pp. 3161-3176, 2024, doi: 10.1002/for.3183.Wiley Online Library

[20] J. Fan, J. Bai, Z. Li, A. Ortiz-Bobea, and C. P. Gomes, "A GNN-RNN Approach for Harnessing Geospatial and Temporal Information: Application to Crop Yield Prediction," arXiv preprint arXiv:2111.08900, 2021. [Online]. Available: https://arxiv.org/abs/2111.08900.arXiv
[21] M. S. Gastli, L. Nassar, and F. Karray, "Satellite images and deep learning tools for crop yield prediction and price forecasting," in Proc. 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 2021, pp. 1-8, doi: 10.1109/IJCNN52387.2021.9533487.MDPI+1 Amrita University+1
[22] S. Sharma, S. Rai, and N. C. Krishnan, "Wheat crop yield prediction using deep LSTM model," arXiv preprint arXiv:2011.01498, 2020. [Online]. Available: https://arxiv.org/abs/2011.01498.MDPI
[23] M. Maimaitijiang et al., "Soybean yield prediction from UAV using multimodal data fusion and deep learning," Remote Sensing of Environment, vol. 1, no. 237, p. 111599, 2020, doi: 10.1016/j.rse.2020.111599.SpringerLink
[24] D. Yoon et al., "Prediction of voluntary motion using a decomposition-and-ensemble framework with deep neural networks," IEEE Access, vol. 8, pp. 201555–201565, 2020, doi: 10.1109/ACCESS.2020.3039733.SpringerLink
[25] J. Jiang, W. Lin, and N. Raghavan, "Semiconductor Manufacturing Final Test Yield Optimization and Wafer Acceptance Test Parameter Inverse Design Using Multi-Objective Optimization Algorithms," IEEE Access, vol. 9, pp. 137655–137666, 2021, doi: 10.1109/ACCESS.2021.3118451.SpringerLink
[26] A. Arami et al., "Prediction of gait freezing in Parkinsonian patients: a binary classification augmented with time series prediction," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 9, pp. 1909–1919, 2019, doi: 10.1109/TNSRE.2019.2921221.SpringerLink
[27] S. Wu et al., "Elevating Prediction Performance for Mechanical Properties of Hot-Rolled Strips by Using Semi-Supervised Regression and Deep Learning," IEEE Access, vol. 8, pp. 134124–134136, 2020, doi: 10.1109/ACCESS.2020.3012345.SpringerLink
[28] J. Jiang et al., "Investigating maize yield-related genes in multiple omics interaction network data," IEEE Transactions on Nanobioscience, vol. 19, no. 1, pp. 142–151, 2019, doi: 10.1109/TNB.2019.2904845.SpringerLink
[29] C. Anderson et al., "Off the Beaten Sidewalk: Pedestrian Prediction in Shared Spaces for Autonomous Vehicles," IEEE Robotics and Automation Letters, vol. 5, no. 4, pp. 6892–6899, 2020, doi: 10.1109/LRA.2020.2991639.SpringerLink
[30] H. Niu, J. R. Peddagudreddygari, M. Bhandari, J. A. Landivar, C. W. Bednarz, and N. Duffield, "In-Season Cotton Yield Prediction with Scale-Aware Convolutional Neural Network Models and Unmanned Aerial Vehicle RGB Imagery," Sensors, vol. 24, no. 8, p. 2432, 2024. [Online]. Available: https://doi.org/10.3390/s24082432.
[31] R. G. Alves, R. F. Maia, and F. Lima, "Development of a Digital Twin for Smart Farming: Irrigation Management System for Water Saving," Journal of Cleaner Production, vol. 388, p. 135617, 2023. [Online]. Available: https://doi.org/10.1016/j.jclepro.2023.135617. ACM Digital Library
[32] A. Feng, J. Zhou, E. D. Vories, and K. A. Sudduth, "Prediction of Cotton Yield Based on Soil Texture, Weather Conditions, and UAV Imagery Using Deep Learning," Precision Agriculture, vol. 25, pp. 303–326, 2024. [Online]. Available: https://doi.org/10.1007/s11119-023-10069-x.SpringerLink+1MDPI+1
[33] T. Yildirim, D. Moriasi, P. Starks, and D. Chakraborty, "Using Artificial Neural Network (ANN) for Short-Range Prediction of Cotton Yield in Data-Scarce Regions," Agronomy, vol. 12, no. 3, p. 652, 2022. [Online]. Available: https://doi.org/10.3390/agronomy12030652.ResearchGate
[34] N. R. Prasad, N. R. Patel, and A. Danodia, "Crop Yield Prediction in Cotton for Regional Level Using Random Forest Approach," Spatial Information Research, vol. 29, pp. 195–206, 2021. [Online]. Available: https://doi.org/10.1007/s41324-020-00346-6.