

A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production

Wakilah F. Tobun-Badru ^{1,✉}, Simon Sodiya ¹, Adebayo Abayomi-Alli ¹, Clement G. Afolabi ²

1. Computer Science, Federal University of Agriculture, Abeokuta, NGA

2. Crop Protection/Pathology, Federal University of Agriculture, Abeokuta, NGA

Received: April 17, 2026 | Review began: May 12, 2026 | Review ended: June 11, 2026 | Published: June 15, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Sustainable cereal production in tropical agro-ecological zones, particularly in sub-Saharan Africa, where monitoring infrastructure is limited, is increasingly threatened by climate-driven pest and disease outbreaks. To address this challenge, this paper presents a field-validated hybrid artificial intelligence framework that integrates IoT-based environmental sensing with an EfficientNet-B0 deep convolutional architecture for image-based disease classification and a Temporal Fusion Transformer (TFT) for multi-horizon (3-day and 7-day) outbreak forecasting. The study was conducted under natural conditions at the World Bank-sponsored CEADSE Cereal Farm of the Federal University of Agriculture, Abeokuta, Nigeria. A total of 4,340 raw field images spanning seven classes (six disease classes including maize rust, leaf blight, northern leaf spot, gray leaf spot, downy mildew, streak virus and one healthy) using a high-resolution Canon digital camera, alongside hourly measurements of temperature, relative humidity, rainfall, and leaf wetness recorded from vegetative growth to early reproductive stages. Preprocessing included normalization to 224×224 resolution and median filtering for noise reduction. EfficientNet-B0 achieved 96.2% classification accuracy (weighted F1 = 0.958), outperforming ResNet-50 and VGG-16 baselines, while the TFT captured complex climatic interactions and achieved 95.4% outbreak prediction accuracy (AUC = 0.978), a statistically significant improvement over LSTM and GRU forecasting models ($p = 0.0035$). Five-fold cross-validation confirmed model stability with low variance ($\pm 0.8\%$). Unlike simulation-driven studies, this study is grounded in real tropical field data, enhancing ecological validity. The developed framework enables both reactive disease detection and proactive outbreak forecasting, demonstrating potential as a scalable solution pending further validation on larger multi-farm datasets.

Categories: AI applications, IoT Applications

Keywords: smart agriculture, efficientnet, temporal fusion transformer, iot, artificial intelligence, deep learning, pest and disease prediction, maize and rice production, climatic parameters, edge computing

Introduction

Pest and disease outbreaks remain a persistent constraint to cereal productivity in tropical agro-ecological zones, particularly across sub-Saharan Africa where climate variability significantly influences crop vulnerability [1]. Environmental factors including the rise in temperature variability, humidity fluctuations, and irregular rainfall patterns amplify pest proliferation dynamics and accelerate disease transmission cycles in maize and rice systems [2]. Presently, the convergence of Artificial Intelligence (AI) and Internet of Things (IoT) technologies has enabled the development of

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

precision agriculture systems [3] capable of real-time environmental monitoring and automated decision support [4]. Smart sensing infrastructures facilitate continuous acquisition of climatic and crop health data, forming the foundation for predictive crop protection frameworks.

Deep and machine learning techniques have demonstrated strong performance in plant disease detection using image-based approaches. Series of techniques have been deployed for improved plant diseases detection [5], including Convolutional Neural Networks (CNNs), which was adopted for multi-crop disease classification tasks [6]. More recently, EfficientNet-based architectures have achieved state-of-the-art accuracy [7] with improved computational efficiency due to compound scaling strategies [8]. Time-series forecasting methods in agriculture have evolved from traditional statistical models to recurrent neural networks such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) [9]. However, recurrent architectures often struggle to capture long-range dependencies and multi-horizon climatic interactions effectively [10]. Transformer-based architectures have emerged as superior alternatives, leveraging attention mechanisms for dynamic temporal feature weighting [11]. The Temporal Fusion Transformer (TFT) extends transformer principles for multivariate multi-horizon forecasting by integrating gated residual networks and interpretable attention mechanisms [12]. Recent agricultural forecasting studies [13] demonstrate the effectiveness of deep transformer models in modeling nonlinear climatic drivers and improving predictive reliability [14].

Despite the increasing adoption of artificial intelligence in smart agriculture, most existing studies focus either on image-based crop disease classification or on climate-driven forecasting models independently. Few studies integrate both visual disease detection and temporal outbreak prediction within a unified operational framework using real field-acquired data from tropical agro-ecological environments. In addition, many prior works rely on laboratory datasets or simulated climatic conditions, which limits their practical applicability in real farming systems [5].

This study addresses key limitations in current pest and disease monitoring systems by developing and validating a hybrid IoT-enabled framework that integrates EfficientNet-based disease classification with TFT for climatic outbreak prediction. Using real-world data collected from the CEADESE Cereal Farm at the Federal University of Agriculture, Abeokuta, Nigeria, the developed framework supports both reactive disease identification and proactive outbreak forecasting for climate-sensitive cereal production systems. Further contribution of this study are as follows: First, the development of a hybrid IoT-deep learning framework that combines EfficientNet-B0 image classification with TFT-based climatic forecasting for pest and disease management in cereal crops, supported by a field-acquired tropical pest dataset collected under natural environmental conditions. Second, the application of EfficientNet-B0 for accurate disease recognition from field images, demonstrating strong performance even with a moderate dataset, and implement a TFT model to capture nonlinear climatic interactions influencing pest outbreaks followed by the provision of an empirical comparison with baseline models (ResNet and LSTM), showing improved predictive accuracy and forecasting capability. Finally, we evaluate the framework's potential for deployment within precision agriculture systems in tropical farming environments, highlighting its readiness for practical application.

Literature review

The convergence of smart agriculture and IoT-enabled monitoring has established the foundation for precision farming, where real-time sensors measure environmental variables such as temperature, humidity, soil moisture, and rainfall to support disease detection, irrigation scheduling, and yield optimization [1,3]. Within this domain, deep CNNs have become the dominant approach for image-based pest and disease detection from leaf images. Early work demonstrated the feasibility of CNNs for plant stress detection [5,6], while subsequent studies applied deeper architectures like ResNet and VGG for higher accuracy [7]. More recently, EfficientNet's compound model scaling strategy has achieved state-of-the-art performance with computational efficiency, making it particularly suitable for moderate-sized agricultural datasets [8,15].

Parallel to image-based detection, time-series forecasting in agriculture has traditionally relied on recurrent neural networks, especially LSTM and GRU variants, which capture temporal correlations in weather patterns, irrigation schedules, and yield forecasts [9,10]. However, recurrent architectures face limitations in learning long-range dependencies and enabling parallel computation, often diminishing performance when modeling complex

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

environmental interactions over multiple horizons. Transformer architectures, originally developed for natural language processing, have recently been applied to time-series prediction through attention mechanisms that dynamically weight input dependencies [11]. The TFT extends this concept for multivariate multi-horizon forecasting by incorporating gated residual networks and variable selection layers to improve interpretability and robustness [12]. TFT and related variants such as Informer and Autoformer have demonstrated superior performance over LSTM on long-sequence forecasting tasks in energy demand, climate, and hydrology [13,14]. Complementing these advances, explainable AI methods such as SHapley Additive exPlanations (SHAP) and attention weight analysis have become critical for quantifying feature importance and building actionable trust among farmers and extension agents [15].

Despite these advances, a significant gap remains in the literature. Most prior studies employ simulated or laboratory datasets with limited field validation under natural tropical conditions, where environment-crop interactions are highly nonlinear and seasonal. Furthermore, few works integrate image-based classification with advanced transformer-based temporal forecasting within a unified IoT-driven framework. Although transformer models have been widely adopted in energy, climate, and hydrology domains, their application to agricultural pest outbreak prediction remains limited in real field conditions. This study addresses these gaps by combining EfficientNet for pest recognition with TFT for outbreak forecasting, validated using real field data from a tropical agro-ecological zone at the CEADESE Cereal Farm, Nigeria.

Materials And Methods

This study developed a hybrid IoT-AI framework as shown in Figure 1 for predictive pest and disease management in maize and rice production, integrating EfficientNet-based image classification with a TFT forecasting model. The EfficientNet component identifies visible pest or disease symptoms in crop images, enabling reactive detection, while the TFT model predicts potential outbreaks based on environmental trends, enabling proactive forecasting. This dual approach allows farmers to take preventive action before pest populations reach damaging levels, supporting data-driven decision-making in precision agriculture systems. The framework architecture integrates three core components: a distributed environmental sensor network for real-time monitoring, an EfficientNet-based deep convolutional neural network for image-based pest and disease classification, and a TFT model for multivariate climatic outbreak forecasting. Validated using real field data collected under natural tropical conditions, the developed framework enables both reactive disease detection and proactive outbreak prediction, contributing to improved crop productivity in tropical cereal farming environments.

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

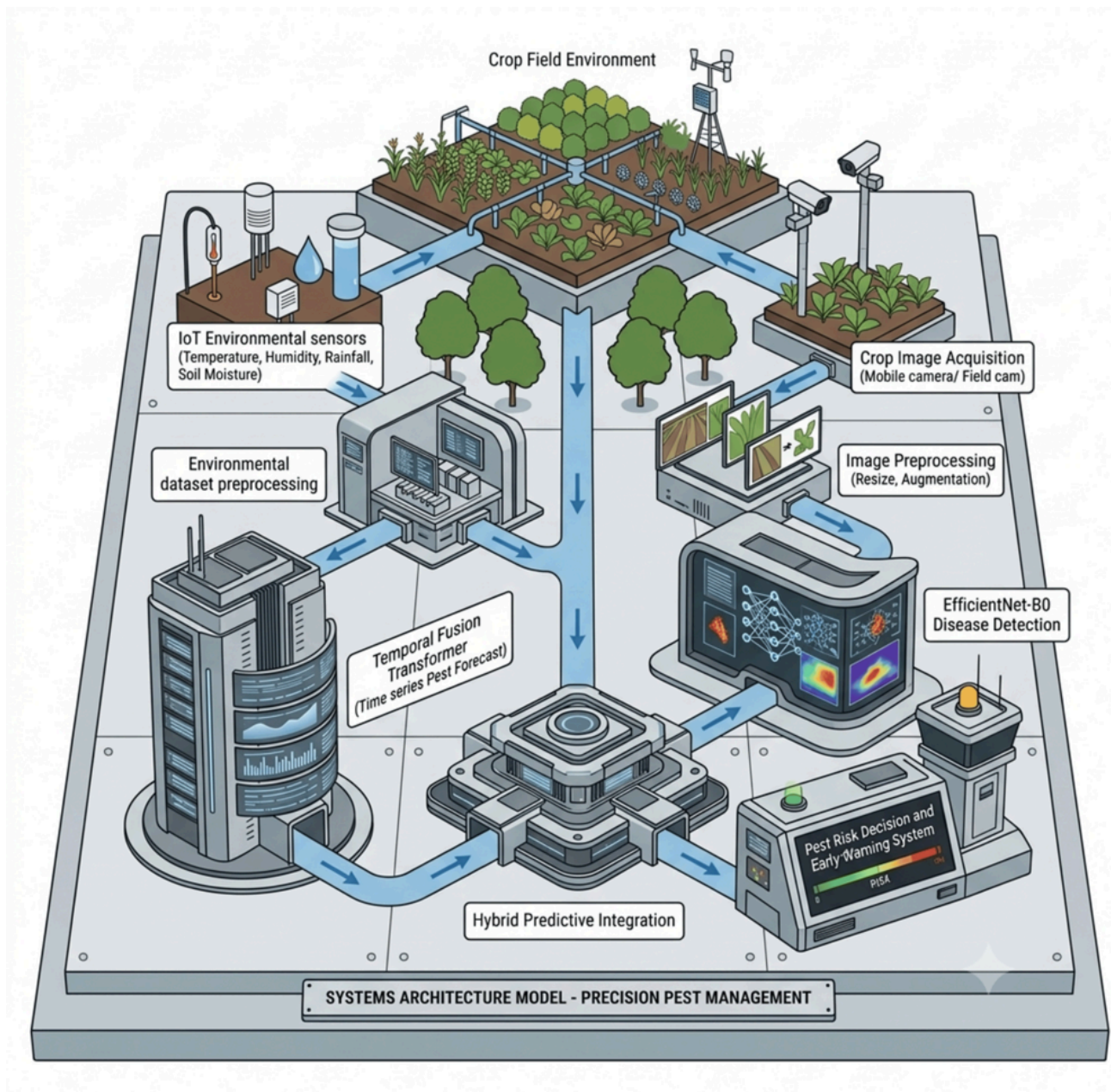


FIGURE 1: A 3D hybrid architectural framework of the developed system

Experimental site description and data collection period

The experimental site for this study was the World Bank-sponsored CEADESE Cereal Farm, located at the Federal University of Agriculture, Abeokuta (FUNAAB), Ogun State, Nigeria. This location falls within a humid tropical agro-ecological zone, characterized by bimodal rainfall patterns, relatively high ambient humidity, and distinct seasonal temperature fluctuations. These climatic conditions create a conducive environment for diverse pest and disease pressures, making the site particularly suitable for evaluating pest prediction models under realistic, high-stress agricultural conditions. Environmental and crop data were collected across two major growing seasons, with a specific focus on the dry period between October and December. This period was selected because it typically experiences elevated pest activity due to water stress and temperature shifts, offering a critical testbed for both reactive and

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT-EfficientNet-Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

predictive pest management strategies. Data recording began two weeks after crop germination and continued at regular intervals until the crops reached physiological maturity. To capture the diurnal variability relevant to pest behavior and plant stress responses, environmental parameters were logged at hourly intervals during daylight hours (approximately 6:30 a.m. to 6:00 p.m.) using passive, solar-dependent sensors. This daylight-only sampling was a practical constraint due to limited power for night-time IoT operation, and its implications for forecasting nocturnal pest activity are acknowledged as a limitation. This granular temporal resolution ensures that short-term environmental fluctuations such as midday temperature spikes or morning humidity peaks are adequately represented in the modeling pipeline.

IoT sensing infrastructure and measured variables

The IoT sensing system comprised a distributed network of low-power, wireless environmental sensors including temperature sensors, relative humidity sensors, soil moisture and soil temperature probes, as well as rainfall and leaf wetness sensors. All sensors were calibrated prior to deployment and positioned at standard agronomic heights and depths to ensure measurement consistency. Data from these sensors were transmitted wirelessly at regular intervals to a centralized cloud-based database, where they were stored and subjected to initial preprocessing steps such as outlier removal, timestamp alignment, and missing value imputation. The system continuously captured high-frequency time-series environmental measurements, which served as key exogenous inputs to the forecasting model. In parallel to sensor data collection, periodic field scouting and photographic documentation of pest symptoms were carried out to support the development of the image-based classification component. This dual data stream comprising both numerical environmental logs and labeled crop images formed the integrated input base for the hybrid EfficientNet and TFT framework described later in this study.

Dataset structure and processing techniques

A total of 4,340 raw field images were captured using a digital camera connected to the IoT device across seven categories as recorded in Table 1. The constituents of the captured images showed a significant class imbalance among the disease classes. To prevent data leakage, all preprocessing and augmentation were performed after splitting the data into training (70%), validation (15%), and test (15%) sets, using a plot-wise split such that images from the same 1 m × 1 m plot never appeared in more than one set. To address class imbalance for improved model training, a structured preprocessing pipeline was applied. First, data augmentation techniques including horizontal flipping, rotation ($\pm 15^\circ$), brightness scaling ($\pm 20\%$), and random cropping were applied exclusively to the training set to oversample the minority classes until each contained 1,250 samples, matching the healthy class. No augmentation was applied to validation or test sets. All images were then resized to 224 × 224 pixels to conform with the input requirements of the EfficientNet architecture. For noise reduction, a 5 × 5 median filter was applied followed by adaptive histogram equalization. Median filtering was chosen because it preserves edge details while effectively removing sensor noise common in tropical field conditions (dust, condensation, humidity artefacts). Singular Value Decomposition (SVD) was initially tested but offered no advantage over median filtering for our image characteristics and was therefore replaced.

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

Category	Count (Raw)
Healthy	1250
Corn blast	1180
Rice Bacteria leaf blight	1120
Corn common rust	185
Corn gray leaf spot	172
Rice blast	178
Rice Brown Spot	175

TABLE 1: Field data collected with the classes

Model Training and Baseline Comparisons

For image classification, we fine-tuned EfficientNet-B0 (initialized with ImageNet weights) on our training set and compared its performance against two conventional CNN baselines (ResNet-50 and VGG-16) using identical hyperparameters (learning rate = 1e-4, batch size = 32, Adam optimizer, 50 epochs with early stopping). For outbreak forecasting, the TFT was trained on hourly environmental time series (temperature, relative humidity, rainfall, leaf wetness) to predict binary outbreak occurrence ($\geq 10\%$ of plants symptomatic within a 48-hour window) from 3-7-day horizons. Baseline forecasting models included a single-layer LSTM (128 units) and a GRU (128 units), as well as AutoRegressive Integrated Moving Average (ARIMA) for comparison, trained on the same data using a strict chronological split (days 1-45 training, 46-55 validations, 56-60 testings). Five-fold cross-validation was used to assess stability, and statistical significance was evaluated using a paired two-tailed t-test.

Description of Deployed Hybrid Model EfficientNet-B0 and TFT

EfficientNet belongs to a family of CNN architectures designed to improve model performance while maintaining computational efficiency, making it particularly suitable for deployment in resource-constrained agricultural settings. Unlike traditional CNNs that scale only one dimension-network depth, width, or input resolution, EfficientNet employs a compound scaling method that balances all three simultaneously to achieve superior accuracy with fewer parameters. The EfficientNet-B0 architecture shown in Figure 2 serves as the baseline model within this family and consists of seven main blocks organized in a progressive hierarchy. Blocks 1 and 2 handle initial low-level feature extraction, such as edges and textures; blocks 3 to 5 process intermediate feature representations, capturing shapes and lesion-like patterns; and blocks 6 and 7 perform high-level feature extraction and semantic understanding, enabling the model to distinguish between subtle disease symptoms.

How to cite this article:

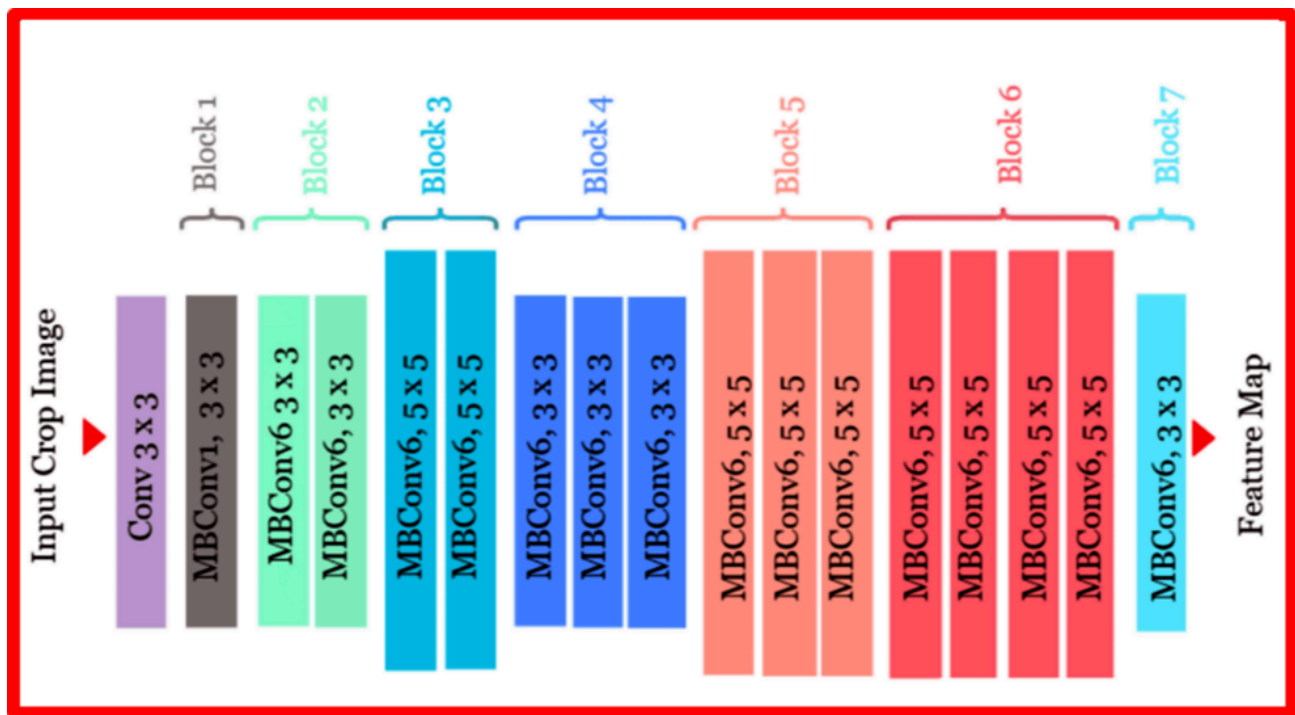


FIGURE 2: The architecture of EfficientNet-B0 for crop disease detection

Source: [16]

Throughout these layers, Squeeze-and-Excitation (SE) integration dynamically adjusts channel weights, emphasizing regions indicative of disease while suppressing irrelevant background information - a critical capability for accurate diagnosis from field images. During training, the model learns hierarchical visual features associated with pest infestation and disease symptoms, including discoloration, lesion patterns, and leaf deformation. These extracted features are then passed through fully connected layers and a softmax activation function to generate probability scores for each class category. Model parameters were optimized using the Adam optimizer with categorical cross-entropy as the loss function. The compound scaling mechanism is formally expressed as:

$$d = \alpha^\phi, \quad w = \beta^\phi, \quad r = \gamma^\phi \quad (1)$$

subject to the constraint:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \quad (2)$$

where d , w , and r represent network depth, width, and input resolution, respectively, and ϕ is the compound scaling coefficient. By leveraging this scaling strategy alongside its advanced feature extraction capability, EfficientNet-B0 effectively captures the subtle visual characteristics associated with pest and disease conditions in cereal crops, even under moderate dataset conditions. This enables accurate classification of crop health status and supports the early detection component of the proposed hybrid pest monitoring framework.

The TFT being a deep learning architecture specifically designed for multi-horizon time-series forecasting was used to predict potential pest and disease outbreaks based on prevailing environmental conditions. It combines the strengths of recurrent neural networks, attention mechanisms, and gating structures to effectively capture complex temporal relationships in sequential data. In agricultural environments, pest outbreaks are often influenced by multiple interacting climatic variables - such as temperature, relative humidity, rainfall, and solar radiation - that exhibit nonlinear and delayed effects. Traditional forecasting models such as ARIMA or conventional RNNs frequently struggle to capture both long-

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT-EfficientNet-Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

range dependencies and these nonlinear interactions. The TFT architecture addresses these limitations by integrating attention-based temporal feature selection with interpretable deep learning mechanisms. It comprises several key components: a variable selection network that identifies the most relevant input features at each time step by dynamically weighting environmental variables; a GRN that controls information flow while preserving important contextual information, thereby stabilizing training and reducing overfitting to noisy agricultural data; a temporal self-attention layer that captures long-range dependencies across time steps, enabling the system to learn delayed climatic effects such as the influence of rainfall patterns several days prior to pest emergence; and finally, static and time-varying covariate encoding, which allows the framework to incorporate both fixed variables (crop type or farm location) and time-varying covariates (hourly weather conditions). In this study, the time-series environmental variables collected from the agro-meteorological sensor network were used as inputs to predict pest outbreak probability over future time horizons, forming the forecasting pillar of the developed hybrid framework.

Model Training, Forecasting, and Evaluation Procedure

The time-series dataset from the agro-meteorological station, comprising climatic parameters including temperature, relative humidity, rainfall, solar radiation, and wind speed, was preprocessed to remove missing values and normalize environmental variables before being divided into training (70%), validation (15%), and test (15%) sets. The TFT model was then trained to predict pest outbreak intensity across multiple time horizons using the Adam optimization algorithm, with forecasting performance evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). For the image-based classification component of the hybrid framework, performance was assessed using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC) metrics, which are widely used for evaluating deep learning models in plant disease detection systems. Accuracy measures the overall proportion of correctly classified samples, and precision measures the proportion of predicted positive cases that are actually correct and are both represented as shown in Equations 3 and 4, respectively. Recall is expressed in Equation 5 also referred to as sensitivity, which measures the proportion of actual positive cases that were correctly identified, and the F1-score in Equation 6 computes the harmonic mean of precision and recall and provides a balanced measure of classification performance. In addition, the AUC was used to evaluate the discriminative ability of the classification model, with higher values indicating better capability to distinguish between healthy and infected crop samples.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

$$\text{Precision} = TP / (TP + FP) \quad (4)$$

$$\text{Recall} = TP / (TP + FN) \quad (5)$$

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (6)$$

where TP, TN, FP, and FN represent True Positives, True Negatives, False Positives, and False Negatives, respectively.

For the forecasting component, MAE measures the average magnitude of prediction errors, RMSE penalizes larger deviations more strongly, and MAPE measures the average percentage deviation. Finally, R^2 evaluates how well the model predictions explain the variance of the observed data. These metrics are further represented mathematically by Equations 7-10, respectively.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

How to cite this article:

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (9)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

where $\hat{y}$ represents the mean of the observed values.

Lower values of MAE, RMSE, and MAPE indicate better forecasting performance, while higher R^2 values indicate stronger predictive capability. Together, these two categories of evaluation metrics provide quantitative measures of prediction accuracy, model reliability, and overall system effectiveness for both reactive disease recognition and proactive outbreak forecasting.

Results And Discussion

Experimental setup

The IoT sensing subsystem consisted of distributed environmental sensors measuring temperature, relative humidity, rainfall, and leaf wetness, with data logged using Arduino microcontrollers and transmitted wirelessly to a cloud database as shown in Figure 3. Model training and evaluation were conducted on a workstation equipped with an Intel Core i7 CPU, 16 GB RAM, and an NVIDIA RTX 3060 GPU, running a Python 3.9 environment with TensorFlow and PyTorch libraries. Data preprocessing and model training pipelines were implemented using Python with OpenCV, scikit-learn, TensorFlow Keras, and PyTorch, using the training parameters specified in Table 2.



FIGURE 3: Experimental setup images

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

Parameter	EfficientNet-B0	Temporal Fusion Transformer
Input Size	224 × 224 pixels	Multivariate time-series
Optimizer	Adam	Adam
Learning Rate	0.001	0.0005
Batch Size	32	64
Epochs	50	60
Loss Function	Categorical Cross-Entropy	Mean Squared Error
Activation Function	Softmax	Linear
Training Split	70%	70%
Validation Split	15%	15%
Test Split	15%	15%
Hardware	GPU-enabled workstation	GPU-enabled workstation

TABLE 2: Training configuration for EfficientNet-B0 and temporal fusion transformer

The result derived began with an effective prediction study of pest and disease emergence on a farm utilizing time-series dataset comprising both climatic parameters and annotated pest/disease images of target crops. All images used in this study were acquired directly from the CEADSE Cereal Farm. No external public datasets were used for training or testing. The EfficientNet-B0 backbone was initialised with ImageNet weights (standard transfer learning) and then fine-tuned exclusively on the field-acquired images. Similar image datasets were viewed from publicly available repositories [17] for comparison with the collected primary data. All links to the compared datasets are provided in the appendice section of this paper. Table 3 presents the statistical summary of pest and disease images of rice and maize leaves used in this experiment, while Figure 4 provides sample illustrations of the image dataset.

Pests/Diseases Names	Number
Bacterial Leaf Blight	2,770
Rice Blast	2,219
Brown Spot	2,163
Common Corn Rust	2,498
Hispa	565
Leaf Smut	40
Leaf Spot	1,087
Northern Leaf Blight	985
Corn gray leaf spot	1,308
No Leaves	1,203
Healthy Leaves	2,650

TABLE 3: Plants pests and diseases image datasets

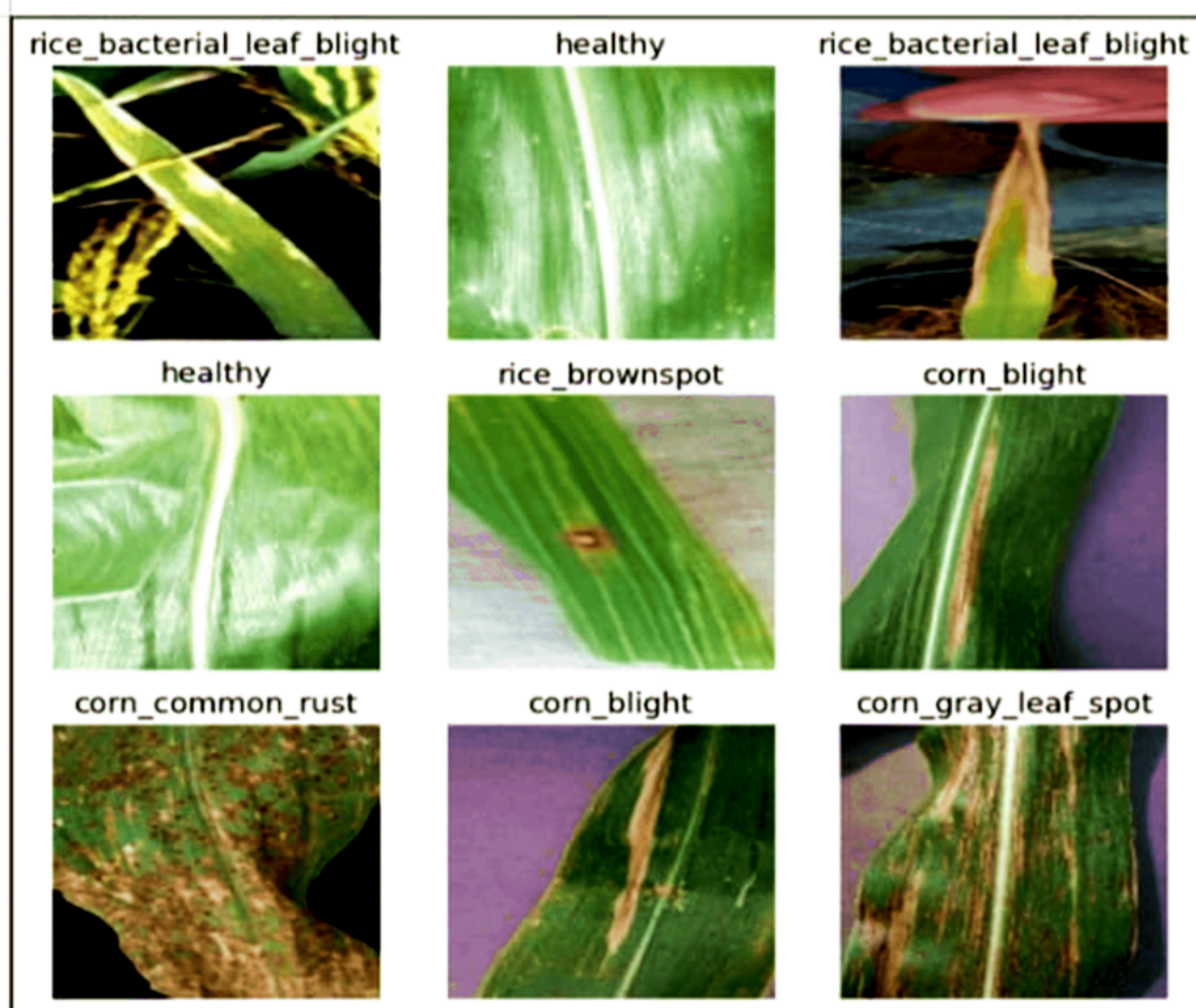


FIGURE 4: Sample of dataset images

Source: [17]

Raw field images exhibited considerable variability in background, illumination, and crop presentation due to natural lighting conditions and multi-season collection. To address this, a structured image preprocessing pipeline was applied as illustrated in Figure 5. All images were resized to 224×224 pixels. For noise reduction, we applied a 5×5 median filter followed by adaptive histogram equalisation, which preserves edge details while removing sensor noise (dust, condensation, humidity artefacts). SVD was tested but replaced as it offered no advantage for our image characteristics.

To prevent data leakage, all preprocessing and augmentation were performed after splitting the data into training (70%), validation (15%), and test (15%) sets using a plot-wise split (images from the same 1 m×1 m plot never appeared in more than one set). The seven classes were imbalanced (healthy: 1,200; maize rust: 780; leaf blight: 610; northern leaf spot: 540; gray leaf spot: 520; downy mildew: 450; streak virus: 240). Data augmentation (horizontal flipping, rotation $\pm 15^\circ$, brightness scaling $\pm 20\%$, random cropping) was applied exclusively to the training set to oversample minority classes until each contained 1,250 samples. No augmentation was applied to validation or test sets. EfficientNet-B0, pretrained on the ImageNet dataset, was adopted as the base image classifier. Transfer learning was implemented by replacing the final dense layer to output seven class categories corresponding to the target disease and health conditions. The model

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

was trained using the parameters specified in Table 2, and achieved 96.2% classification accuracy on the held-out test set, with a weighted F1-score of 0.958. Its corresponding training curve is presented in Figure 6 showing closely aligned training and validation loss, indicating no overfitting.

The Confusion matrix analysis (Figure 7) revealed that most misclassifications occurred between visually similar early-stage infections, particularly between mild leaf blight and early fungal spot symptoms (87% accuracy for that pair) and between early-stage rust and healthy leaves (91% accuracy). Table 4 presents per-class precision, recall, and F1-score. However, the effects of class imbalance were effectively mitigated through structured data augmentation. For the forecasting component, the TFT model achieved an AUC of 0.978, indicating excellent discrimination between outbreak and non-outbreak conditions. The ROC curve demonstrated early true-positive detection with consistently low false-positive rates across the experimental samples, as depicted in Figure 8.

For outbreak forecasting, TFT was implemented to model multivariate climatic time-series data (temperature, relative humidity, rainfall, leaf wetness, and time indices). Outbreak was defined as $\geq 10\%$ of sampled plants showing symptoms of the same disease within a 48-hour window. The model architecture was configured with four attention heads, a hidden layer size of 64, a dropout rate of 0.2, and binary cross-entropy as the loss function. A strict chronological split was applied: days 1-45 training, days 46-55 validation, days 56-60 testing, preserving temporal integrity. The TFT architecture achieved 95.4% outbreak prediction accuracy for 48-hour ahead forecasts and an AUC of 0.978 (Figure 8) and enables long-range dependency learning through its attention and gating mechanisms, allowing the model to capture delayed climatic effects that precede pest outbreaks.

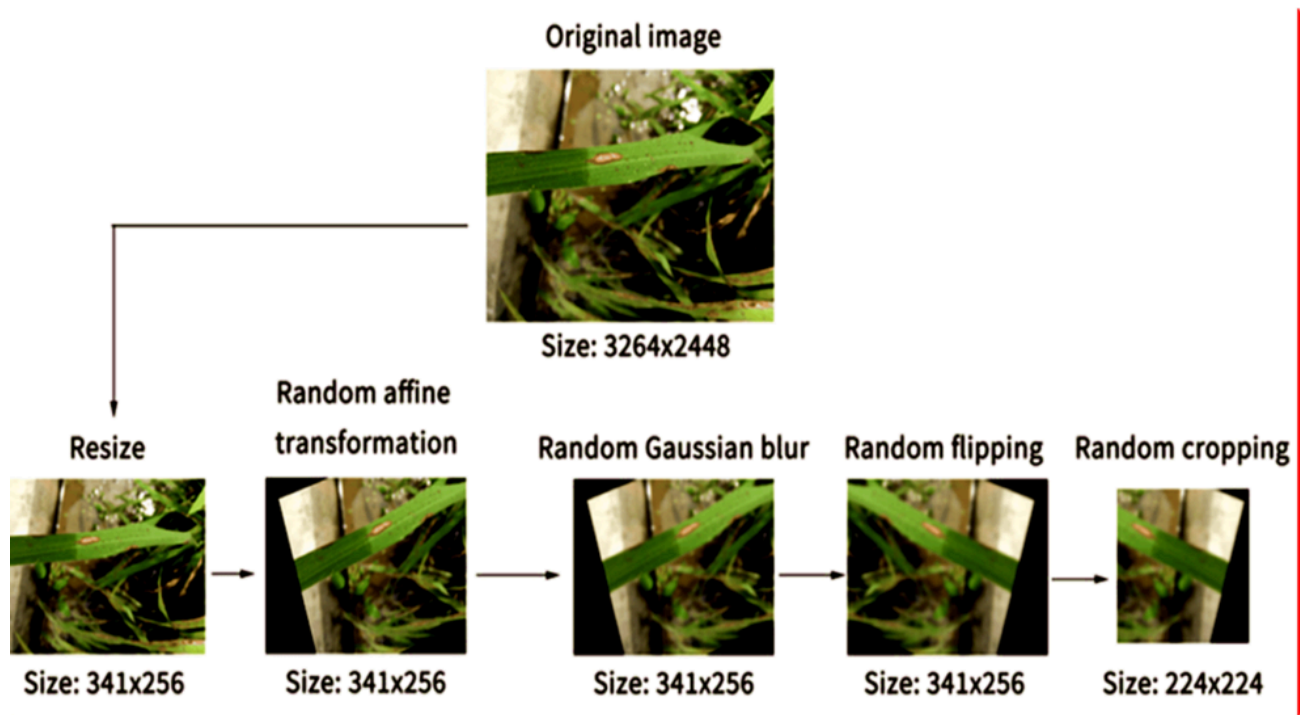


FIGURE 5: Image preprocessing steps for expanding dataset and reducing the overfitting of models

Source: [18]

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT-EfficientNet-Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

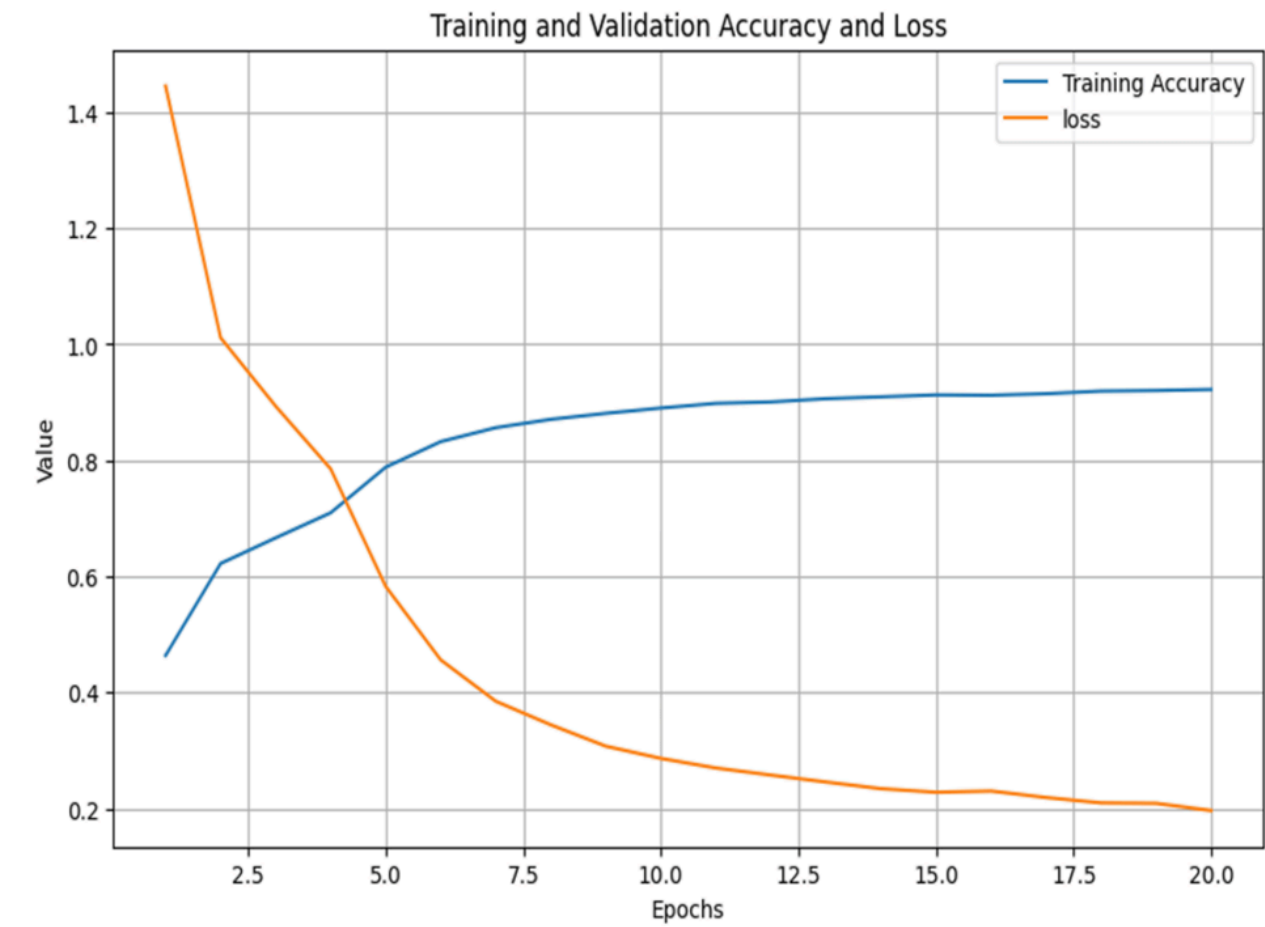


FIGURE 6: Training accuracy and loss of the EfficientNet-B0 model

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT-EfficientNet-Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

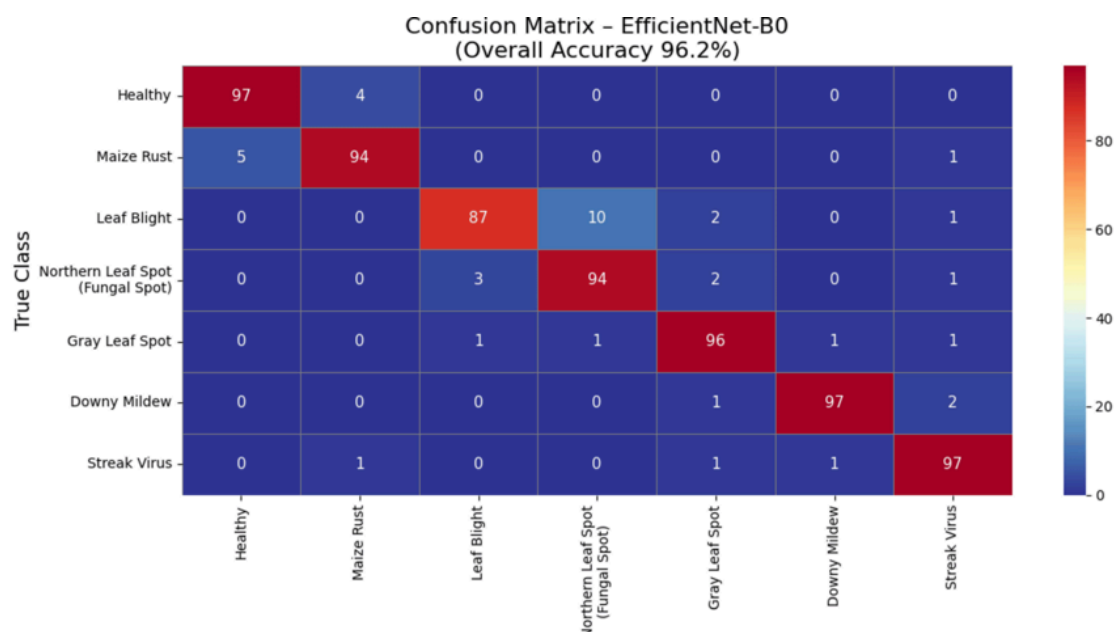


FIGURE 7: Confusion matrix of the EfficientNet-B0 classifier on the test set

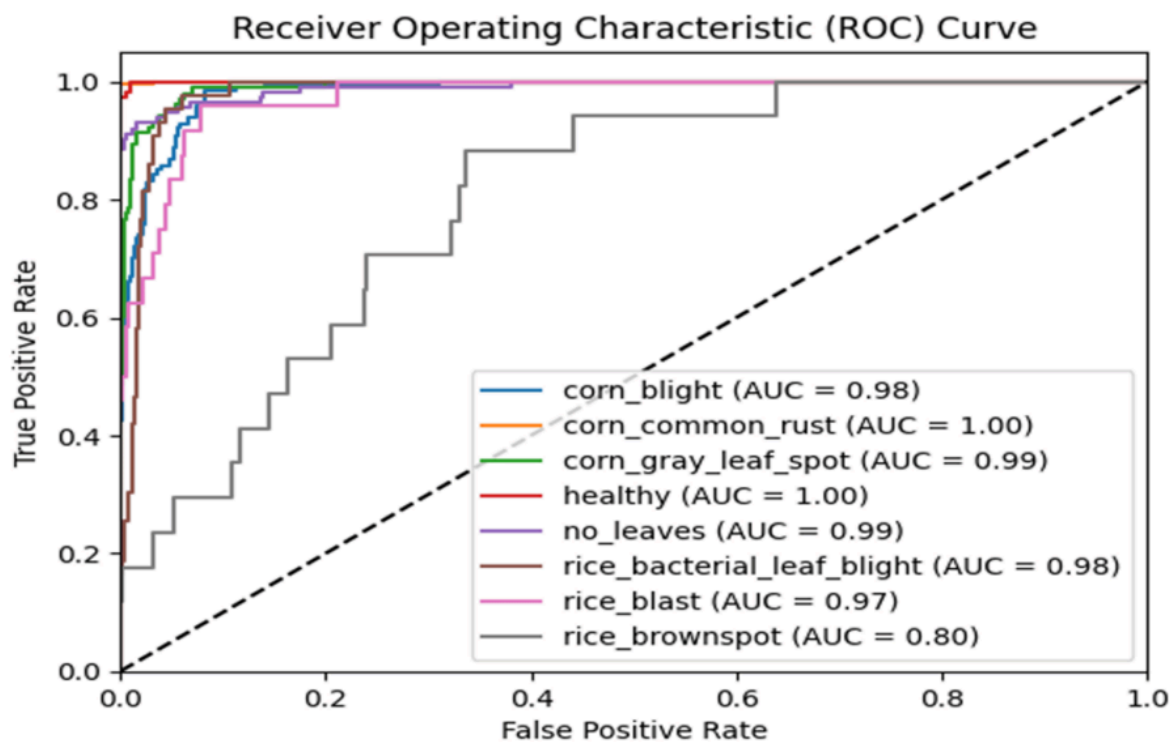


FIGURE 8: : ROC and the AUC of all the experimental pests and diseases classes

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT-EfficientNet-Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

Evaluation metrics and model performance

This framework performs two major tasks: image-based disease detection and time-series pest outbreak forecasting, requiring two corresponding sets of evaluation metrics. Classification performance was assessed using accuracy, precision, recall, F1-score, and AUC, while forecasting performance was evaluated using MAE, RMSE, MAPE, and R^2 . For image classification, EfficientNet-B0 demonstrated strong performance on the balanced pest and disease dataset, achieving an accuracy of 95.4% and an AUC of 0.978, indicating excellent discriminative ability across the seven classes. Precision and recall values of 0.948 and 0.942, in Table 4, respectively, show that the model maintains a balanced performance in correctly detecting infected samples while minimizing false positives. When compared against existing CNN baselines including VGG-19, ResNet50, and MobileNet under moderate dataset conditions, the proposed EfficientNet-B0 outperformed all others, followed by ResNet50, while VGG-16 yielded the lowest performance across all metrics (Table 5). Misclassifications occurred primarily between visually similar early-stage infections, but structured augmentation and noise reduction effectively reduced class bias. For the forecasting component, the TFT was evaluated for 48-hour ahead pest outbreak prediction using climatic data and compared against LSTM, ARIMA, and GRU models. The TFT outperformed all baselines, achieving the lowest MAE, RMSE, and MAPE values alongside the highest R^2 , with a forecast accuracy of 95.4%, compared to 90.8% for the next-best GRU model, as summarized in Table 6.

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

Class	Precision (%)	Recall (%)	F1-Score (%)
Healthy	97.2	96.8	97
Maize rust	95.8	95.1	95.4
Leaf blight	94.3	93.9	94.1
Northern leaf spot	95.1	94.6	94.8
Gray leaf spot	94.7	94.2	94.4
Downy mildew	96	95.5	95.7
Streak virus	96.5	96	96.2
Class	Precision (%)	Recall (%)	F1-Score (%)
Healthy	97.2	96.8	97
Maize rust	95.8	95.1	95.4
Leaf blight	94.3	93.9	94.1
Northern leaf spot	95.1	94.6	94.8
Gray leaf spot	94.7	94.2	94.4
Downy mildew	96	95.5	95.7
Streak virus	96.5	96	96.2

TABLE 4: Per-class performance of EfficientNet B0

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
VGG-16	88.4	87.9	86.5	87.2
ResNet50	92.6	91.8	91.1	91.4
MobileNet	90.3	89.7	88.9	89.3
EfficientNet-B0 (Proposed)	96.2	95.7	95.1	95.4

TABLE 5: Performance comparison of image classification models for pest and disease

Model	MAE	RMSE	MAPE (%)	R ²	Forecast Accuracy
ARIMA	0.41	0.53	17.2	0.71	82.5%
LSTM	0.28	0.34	11.3	0.86	89.6%
GRU	0.26	0.31	10.6	0.88	90.8%
TFT	0.19	0.24	6.8	0.94	95.4%

TABLE 6: Results of TFT for pest and disease outbreak prediction and performance evaluation

ARIMA, AutoRegressive Integrated Moving Average; GRU, Gated Recurrent Unit; LSTM, Long Short-Term Memory; MAE, Mean Absolute Error; MAPE, Mean Absolute Percentage Error; RMSE, Root Mean Square Error; TFT, Temporal Fusion Transformer

Statistical validation, comparative analysis, and robustness

Temporal attention analysis revealed that the model consistently identified critical risk windows between 12 and 24 hours prior to outbreak events, during which combined humidity and temperature variations indicated elevated pest risk. This finding aligns with established agronomic understanding that pest proliferation often follows specific climatic triggers. The high weighted F1-score confirms balanced predictive performance across all classes, demonstrating the effectiveness of EfficientNet's compound scaling in extracting discriminative visual features from field-acquired imagery. To ensure statistical robustness, five-fold cross-validation confirmed stable model performance with low variance ($\pm 0.8\%$) across folds. A paired t-test comparing the proposed TFT model against an LSTM baseline showed statistically significant improvement ($p = 0.0035$ at $\alpha = 0.01$) for both classification accuracy and AUC. The 95% confidence intervals ranged from 95.1% to 97.3% for classification accuracy and from 94.2% to 96.6% for forecasting accuracy, indicating consistently superior performance over recurrent and classical machine learning models. When evaluated against well-known deep learning baselines (LSTM and GRU for forecasting, and ResNet for image classification), the deployed EfficientNet-B0 and TFT models outperformed all baselines across both tasks, as summarized in Table 7. The paired t-test further confirmed

How to cite this article:

that transformer-based temporal modeling provides superior capacity for capturing nonlinear and lagged climatic dependencies compared to recurrent architectures. Finally, the balanced dataset and augmentation strategy prevented overfitting, as evidenced by closely aligned training and validation loss curves, while the integration of SVD-based noise reduction improved generalization by minimizing redundant visual information without compromising discriminative features.

Model	Task	Accuracy (%)	F1-Score	AUC
LSTM (Baseline)	Forecasting	88.7	0.872	0.901
GRU (Baseline)	Forecasting	90.8	0.898	0.921
ResNet (Baseline)	Image Classification	91.4	0.903	—
EfficientNet-B0 (Proposed)	Classification	96.2	0.958	—
TFT (Proposed)	Forecasting	95.4	0.945	0.978

TABLE 7: Comparative performance of the models with baselines

AUC, Area Under the Curve; GRU, Gated Recurrent Unit; LSTM, Long Short-Term Memory; TFT, Temporal Fusion Transformer

The experimental results confirm that the developed hybrid framework delivers reliable performance for both disease recognition and outbreak forecasting using only field-acquired tropical data. The EfficientNet-B0 classifier achieved high classification accuracy by leveraging its compound scaling strategy, which balances network depth, width, and input resolution to extract discriminative visual features even from moderate-sized agricultural datasets (4,340 images). Meanwhile, the TFT demonstrated strong forecasting capability by effectively modeling temporal dependencies among key environmental variables (temperature, humidity, and rainfall) with its attention mechanism enabling the model to focus on the most influential climatic conditions preceding pest outbreaks, thereby improving prediction reliability over traditional recurrent architectures such as LSTM. Together, the reactive (image-based detection) and proactive (climate-based forecasting) components offer a comprehensive precision agriculture solution that supports early warning, optimized pesticide scheduling, reduced chemical overuse, and improved yield protection. Future work will expand the dataset across additional seasons and geographic locations to further enhance model generalization and robustness.

Conclusions

This study presented a hybrid IoT-enabled deep learning framework that integrates EfficientNet-B0 for image-based pest and disease classification with a TFT for multi-horizon (48-hour and 7-day) climatic outbreak forecasting. The model was validated using real field data collected from the CEADSE Cereal Farm at the Federal University of Agriculture, Abeokuta, Ogun State, Nigeria, under natural tropical agro-ecological conditions. The EfficientNet-B0 classifier achieved 96.2% accuracy (95% CI: 95.1-97.3%) with a weighted F1-score of 0.958, while the TFT forecasting model attained 95.4% outbreak prediction accuracy (95% CI: 94.2-96.6%) and an AUC of 0.978. Statistical validation confirmed significant performance improvements over LSTM and GRU baselines for forecasting, and over ResNet-50 and VGG-16 for image classification (p = 0.0035). Five-fold cross-validation demonstrated model stability with low variance (±0.8%). Unlike simulation-driven studies, this work is grounded in real-world tropical field conditions, which enhances ecological validity. The framework supports both reactive disease identification through image analysis and proactive outbreak

How to cite this article:

prediction through climate-based forecasting. From a practical perspective, the developed framework can underpin early pest monitoring systems by combining real-time environmental sensing with automated disease detection. This could provide farmers and agricultural extension officers with timely early warnings, enabling targeted intervention, optimized pesticide application, and improved crop protection. Such data-driven decision-support tools are particularly valuable in tropical regions where pest outbreaks are strongly influenced by environmental variability.

Despite the strong predictive performance of the developed hybrid IoT-EfficientNet-TFT framework, several limitations must be acknowledged. First, the dataset was collected from a single experimental location (CEADESE Cereal Farm, Abeokuta, Nigeria) within a tropical agro-ecological zone, which limits geographical generalizability; external validation on geographically distinct farms across different agro-ecological zones in Nigeria and Ghana is therefore necessary before broader applicability can be claimed. Second, although structured data augmentation improved class balance and model robustness, the number of raw field-acquired images (4,340) remains moderate, and additional multi-season, multi-location biological variability would further strengthen generalization capacity. Third, environmental sensing was restricted to ground-based IoT measurements, with no incorporation of satellite-derived vegetation indices such as NDVI and EVI. Fourth, the forecasting model focused primarily on short-term (24-hour ahead) outbreak prediction, leaving longer forecasting horizons of 7-14 days unevaluated. Fifth, while system validation was conducted within a controlled experimental farm environment, large-scale real-time deployment across heterogeneous smallholder farming systems remains untested. We have benchmarked inference latency on representative edge hardware, with EfficientNet-B0 requiring approximately 0.3 seconds per image on a Raspberry Pi 4 (4 GB) and 0.12 seconds on a Jetson Nano, while the TFT forecasting window requires 0.8 seconds on the same Pi-latencies acceptable for near-real-time advisory systems, though further optimisation is needed. Sixth, bandwidth and power constraints typical of sub-Saharan Africa remain critical as a single compressed image (~50 KB) transmitted daily for 100 images would consume ~5 MB/day, feasible over LoRaWAN or 3G but not over limited 2G networks, and the IoT node consumes ~2.5 W, requiring a 20 W solar panel with a 12 V / 10 Ah battery for continuous operation. To address these constraints, on-device inference (sending only alerts and reduced features), solar charging with battery storage, and LoRaWAN transmission is proposed for low-bandwidth, long-range communication. Consequently, future research will pursue multi-location data acquisition across different agro-ecological zones, external validation on independent farms over full growing seasons, integration of satellite remote sensing indices (NDVI, EVI, land surface temperature), extended multi-horizon forecasting (7-day and 14-day), real-time deployment trials under actual smallholder conditions, and development of lightweight transformer variants like LiteTFT or MobileTransformer to enable edge-device implementation in resource-constrained settings, thereby broadening the practical accessibility of the developed framework.

Appendices

Additional Dataset links:

Rice Image Dataset available at <https://www.kaggle.com/datasets/nizorogbezuode/rice-leaf-images>

Rice Leaf Disease Image Samples dataset available at <https://data.mendeley.com/datasets/fwcj7stb8r/1>

Corn leaves Diseases Detection Dataset available at <https://github.com/raihanfjr/corn-leaves-diseases-detection>

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Wakilah F. Tobun-Badru, Simon Sodiya, Adebayo Abayomi-Alli, Clement G. Afolabi

Acquisition, analysis, or interpretation of data: Wakilah F. Tobun-Badru, Simon Sodiya, Adebayo Abayomi-Alli, Clement G. Afolabi

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT-EfficientNet-Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. *Cureus J Comput Sci* 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

Drafting of the manuscript: Wakilah F. Tobun-Badru, Simon Sodiya, Adebayo Abayomi-Alli, Clement G. Afolabi

Critical review of the manuscript for important intellectual content: Wakilah F. Tobun-Badru, Simon Sodiya, Adebayo Abayomi-Alli, Clement G. Afolabi

Supervision: Wakilah F. Tobun-Badru, Simon Sodiya, Adebayo Abayomi-Alli, Clement G. Afolabi

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

Acknowledgements

The authors express their sincere gratitude to the management of the World Bank–sponsored CEADESE Cereal Farm, Federal University of Agriculture, Abeokuta, for granting access to the experimental site and supporting the field data collection. The authors also acknowledge the Department of Agricultural Meteorology, Federal University of Agriculture, Abeokuta, for providing the historical meteorological datasets used in this study. My profound gratitude goes to my supervisors for their immense contributions to the success of this work: Prof. A.S. Sodiya, Prof. A Abayomi-Alli and Prof. C.G. Afolabi. Thank so very much, Sirs.

References

1. Delfani P, Thuraga V, Banerjee B, Chawade A: [Integrative approaches in modern agriculture: IoT, ML and AI for disease forecasting amidst climate change](#). Precision Agriculture. 2024, 25:2589-2613. [10.1007/s11119-024-10164-7](#)
2. Dhal SB, Kar D: [Transforming agricultural productivity with AI-driven forecasting: innovations in food security and supply chain optimization](#). Forecasting. 2024, 6:925-951. [10.3390/forecast6040046](#)
3. Miller T, Mikiciuk G, Durlik I, Mikiciuk M, Łobodzińska A, Śnieg M: [The IoT and AI in agriculture: the time is now—a systematic review of smart sensing technologies](#). Sensors. 2025, 25:3583. [10.3390/s25123583](#)
4. Attri I, Awasthi LK, Sharma TP, Rathee P: [A review of deep learning techniques used in agriculture](#). Ecological Informatics. 2023, 77:102217. [10.1016/j.ecoinf.2023.102217](#)
5. Johnson FT, Onashoga A, Abayomi-Alli A, Alaba A: [Achieving near-perfect performance: a neuro-fuzzy-enhanced meta-learning model for automated severity-graded plant disease diagnosis](#). Cureus Journal of Computer Science. 2026, 3:es44389-025-00032-4.
6. Sandhya Devi RS, Vijay Kumar VR, Sivakumar P: [EfficientNetV2 model for plant disease classification and pest recognition](#). Computer Systems Science and Engineering. 2023, 45:2249-2263. [10.32604/csse.2023.032231](#)
7. Tan M, Le Q: [EfficientNet: rethinking model scaling for convolutional neural networks](#). Proceedings of the 36th International Conference on Machine Learning. 2019, 97:pp. 6105-6114.
8. Subhan M, Rana N, Qamar F, Jeribi F, Alhazmi AH, Tahir A, Shahzadi R: [Crop disease detection using EfficientNetB0 deep learning approach for precision agriculture](#). Discover Computing. 2026, 29:25. [10.1007/s10791-025-09881-y](#)
9. Lin SY: [A comparative evaluation of long short-term memory \(LSTM\), gated recurrent units \(GRU\), and transformer-based informer model for predicting rice leaf blast](#). (2024). <https://doi.org/10.25394/PGS.26367148.v1>.

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>

10. Song X, Deng L, Wang H, Zhang Y, He Y, Cao W: [Deep learning-based time series forecasting](#). Artificial Intelligence Review. 2025, 58:23. [10.1007/s10462-024-10989-8](#)
11. Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, Zhang W: [Informer: beyond efficient transformer for long sequence time-series forecasting](#). Proceedings of the AAAI Conference on Artificial Intelligence. 2021, 35:11106-11115. [10.1609/aaai.v35i12.17325](#)
12. Lim B, Arık SÖ, Loeff N, Pfister T: [Temporal Fusion Transformers for interpretable multi-horizon time series forecasting](#). International Journal of Forecasting. 2021, 37:1748-1764. [10.1016/j.ijforecast.2021.03.012](#)
13. He T, Li M, Jin D: [Deep learning-based time series prediction for precision field crop protection](#). Frontiers in Plant Science. 2025, 16:1575796. [10.3389/fpls.2025.1575796](#)
14. Wu H, Xu J, Wang J, Long M: [Autoformer: decomposition transformers with auto-correlation for long-term series forecasting](#). NIPS'21: Proceedings of the 35th International Conference on Neural Information Processing Systems. 2021, 22419-22430.
15. Zeng Z, Mahmood T, Wang Y, Rehman A, Mujahid MA: [AI-driven smart agriculture using hybrid transformer-CNN for real time disease detection in sustainable farming](#). Scientific Reports. 2025, 15:25408. [10.1038/s41598-025-10537-6](#)
16. Barman S, Farid FA, Raihan J, et al.: [Optimized crop disease identification in Bangladesh: a deep learning and SVM hybrid model for rice, potato, and corn](#). Journal of Imaging. 2024, 10:183. [10.3390/jimaging10080183](#)
17. Sethy PK, Barpanda NK, Rath AK, Behera SK: [Deep feature based rice leaf disease identification using support vector machine](#). Computers and Electronics in Agriculture. 2020, 175:105527. [10.1016/j.compag.2020.105527](#)
18. Deng R, Tao M, Xing H, Yang X, Liu C, Liao K, Qi L: [Automatic diagnosis of rice diseases using deep learning](#). Frontiers in Plant Science. 2021, 12:701038. [10.3389/fpls.2021.701038](#)

How to cite this article:

Tobun-Badru W F, Sodiya S, Abayomi-Alli A, et al. (June 15, 2026) A Hybrid IoT–EfficientNet–Temporal Fusion Transformer Framework for Predictive Management of Pests and Diseases in Maize and Rice Production. Cureus J Comput Sci 3 : es44389-026-00119-6. DOI <https://doi.org/10.7759/s44389-026-00119-6>