Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Iot-Enabled Evidence-Fusing Graph Neural Network With Attention-Based Feature Engineering For Uncertainty-Aware Crop Recommendation

Kanuprasad. M. K[1*], Dr.N.V. Balaji[2]

[1*]Research scholar, Karpagam Academy of Higher Education, Coimbatore, INDIA
[2]Professor and Dean,Faculty of Arts, Science, Commerce and Managagement,Karpagam Academy of Higher Education, Coimbatore, I N D I A
[1*]Corrsponding author Email: prasadkanu32@gmail.com

Abstract
Precision agriculture demands not only accurate crop recommendations based on heterogeneously measured soil and environment, but also confidence information for the decisions being made in the field. In this study, a novel IoT-Enabled Attention Multiscale Evidence-Fusing Graph Neural Network Framework (IoT-AM-EFGNN) for intelligent and uncertainty-aware crop recommendation is developed. The framework combines the IoT sensing, AMTNet-inspired multiscale attention feature engineering, Evidence Fusing Graph Neural Network (EFGNN), Enzyme Action Optimizer (EAO), and SHapley Additive exPlanations (SHAP). Experimental evaluation has been performed on Kaggle Crop Recommendation Dataset which consists of data points comprising 22 crop classes and 7 agricultural parameters of 2200 samples. Missing value imputation, Z-score outlier detection, Min-Max normalization, Label encoding, train-test split with ratio 80:20 stratified was applied. We then created Soil Health, Climate, and Moisture indices, which were further enhanced using attention-based feature fusion. The EFGNN learned relationships between agricultural samples by propagating multi-level graphs and fused class-specific Dirichlet evidence to achieve crop probabilities and uncertainty measures. EAO fine-tuned the model, and SHAP offered feature-level interpretability. Experimental results demonstrated 99.18% accuracy, 99.12% precision, 99.08% recall, 99.10% F1-score, and 99.42% ROC-AUC. The cross-validation results show that the method achieves high accuracy with a mean of 99.18% and a standard deviation of 0.16%, which indicates that the method is stable. The proposed framework's accuracy was 1.72% higher than the baseline MLP and 0.44% higher than a conventional GNN. Confidence distribution was between about 0.95 and 1.00 on every prediction. The deployment of IoT-cloud is also a facility for real-time agricultural decision making.

Keywords: Precision Agriculture; Internet of Things (IoT); Crop Recommendation; Graph Neural Network; Explainable Artificial Intelligence; Deep Learning.

1. Introduction

Agriculture is one of the key pillars to achieve food security, economic development, and sustainable livelihoods globally [1]. India is one of the major agricultural producers of the world, but still around 64% of the cultivated area is relying on monsoon rainfall which makes the productivity of agriculture very sensitive towards climatic variability [2]. Moreover, irrigation consumed about 85% of the available fresh water resources, with almost 60% of it being lost to a lack of irrigation efficiency [3]. The problems require the use of intelligent and resource efficient farming technologies, which will promote crop productivity while reducing water usage, soil erosion and overuse of fertilizers [4]. Precision Agriculture is one solution that has come about as a result of the integration of advanced information technologies to monitor, analyse and manage agricultural resources to facilitate the decision making process based on data [5].

With the development of Internet of Things (IoT) technology, traditional agriculture has been evolving into smart agriculture [6]. These sensor networks continuously transmit the necessary data on soil moisture and soil parameters (nitrogen, N; phosphorus, P; potassium, K; soil pH, temperature, humidity, rainfall) to the Internet and thus allow real-time monitoring of the soil state in the field [7]. Collected sensor data are then sent via wireless communication networks, and stored on cloud-based platforms where they can be analyzed in greater detail [8]. The system allows effective monitoring of crops, irrigation scheduling, fertilization, pest detection, and environmental monitoring, reduces cost and makes agriculture more sustainable [9].

Development in the field of agriculture sensors is from single sensing devices to IoT intelligent connected cloud systems [10]. Evolution of the sensors for the past few years illustrates evolution from single sensors to multiple sensor networks that are capable of gathering information from the environment in a large-scale manner [11]. The latest architecture of IoT consists of sensor networks, communication gateways, cloud servers and intelligent analytics systems, which enable continuous monitoring and management of the agricultural field in a remote manner [12].

However, as the data become increasingly voluminous and complex in the field of agriculture, it gets increasingly difficult to construct computational models than the traditional machine learning approaches [14]. The traditional way of recommending the crop does not work because of the complex nonlinear relationship between the soil quality, weather condition and the suitability of crops [15]. Over the past few years, deep learning has successfully been employed to extract meaningful features from multivariate agricultural datasets with a remarkable ability [16]. It is possible to model complex relationships between environmental factors and make highly accurate crop recommendations for various production situations using a deep neural network [17].

To address these challenges, the current study introduces an intelligent IoT deep learning system for crop recommendation in precision agriculture. The proposed framework involves the real-time acquisition of information from an IoT sensor, management of information in the cloud, detailed data preprocessing, feature engineering, and the use of a hybrid deep learning model to achieve the correct prediction of the crop. The system is designed to use soil nutrient data along with climatic data to provide recommendations for the most appropriate crop to grow, enhancing agricultural productivity, maximizing resource use, minimizing water losses, and promoting sustainable farming. The contributions of this work are given below,

- A novel IoT-enabled multiscale attention and evidence-fusing graph learning framework is developed by integrating IoT sensing, AMTNet-inspired attention, EFGNN, EAO optimization, and SHAP explainability for end-to-end crop recommendation.
- A multiscale agricultural feature-engineering strategy is introduced by constructing Soil Health Index, Climate Index, and Moisture Index representations and adaptively refining them through attention to capture important interactions among soil and environmental conditions.
- The EFGNN architecture utilizes multi-level graph propagation along with evidence accumulation and uncertainty estimation through Dirichlet distribution, allowing us to provide the crop recommendations along with a quantitative measure of prediction uncertainty without depending entirely on Softmax probability scores.
- The Enzyme Action Optimizer (EAO) algorithm is used for automated hyperparameter tuning, and SHAP is used for explaining the feature importance, thus optimizing the process as well as providing explainability.

The remainder of this paper is organized as follows: Section 2 presents the related works, Section 3 describes the proposed methodology, Section 4 discusses the experimental results and performance evaluation, and Section 5 concludes the study with future research directions.

2. Related work

Sawant et al, (2026) [18] designed a Smart Internet of Things (IoT)-Based Crop Recommendation System for real-time crop selection, applying an Enhanced Gaussian Naive Bayes (EGNB). The process included a capacitive soil moisture sensor, soil temperature-humidity sensor, Arduino Nano, nRF24L01 radio transmitter/receiver, ESP32, and Blynk, a platform for wireless communication of the environmental data and analysis. The EGNB model reached an accuracy of 99.55% and had better precision, recall, and F1 score than the other models: Random Forest, Bagging, Decision Tree, and Gradient Boosting. The system, however, required constant Internet connections and has to rely on the reliability of the sensors, which meant that it failed to perform well in remote areas of agriculture.

Sravanthi et al. (2024) [19] presented a Multi-level Kronecker Guided Pelican Convolutional Neural Network (MKGPCNN) for crop recommendation and a Combined Graph Sample and Aggregate Attention Network (CGSAAN) for crop disease prediction and fertilizer recommendation. Crop Recommendation Dataset and New Plant Diseases Dataset were used in the process to identify the appropriate crops, identify the diseases found and recommend the appropriate fertilizers. The two models, MKGPCNN and CGSAAN, have accuracy levels of 99% and 98%, precision level of 99.5% and 99%, and recall levels of 99.6% and 98.9%, respectively. Nonetheless, its development was anchored on public data and was computation-intensive.

Aldossary et al, (2024) [20] presented An Internet of Things (IoT) enabled Hybrid Artificial Intelligence (AI) and Machine Learning (ML) Framework for smart agriculture to support anomaly detection and soil type classification. It used the Dry Beans (2021) and Soil Type (2024) datasets and applied Multilayer

Perceptron (MLP), Naïve Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), MobileNetV2, Visual Geometry Group-16 (VGG16), InceptionV3, and Long Short-Term Memory (LSTM). MobileNetV2 model with 97% accuracy and recall, SVM with 93% accuracy. However, the model was extremely data-intensive and computation-intensive.

Kaur et al (2026), [21] introduced an Internet of Things (IoT)-Enabled Machine Learning (ML)-Based Predictive Framework for intelligent crop planning, productivity estimation, and precision agriculture. The data used in the process were IoT sensor data with a 70:30 training—test split and the models used were Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Gradient Boosting, Adaptive Boosting (AdaBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost). The highest accuracy was obtained for the AdaBoost model, which was 99.70%, followed by CatBoost with 98.48% and Gradient Boosting with 98.33%. The framework was, however, based on high quality IoT sensor information and benchmark datasets, and was therefore not very adaptive to different real world agricultural conditions.

Kiruthika (2023), [22] presented an Improved Distribution-based Chicken Swarm Optimization–Weight-based Long Short-Term Memory (IDCSO-WLSTM) framework for Internet of Things (IoT)-based crop prediction and recommendation. The first task was collecting climate and crop production information, secondly, pre-processing the data, thirdly, optimal feature selection using Improved Distribution based Chicken swarm optimization (IDCSO) and finally, crop prediction using Weight based Long Short-Term Memory (WLSTM). The results showed high precision, recall, and lower execution time than the previous methods. But it had relied on the quality of climate data and lacked detailed quantitative performance measures for good reproduction and real-world validation.

Logeshwaran et al, (2024) [23] developed an Agro Deep Learning Framework (ADLF) for Artificial Intelligence (AI)-based precision agriculture to enhance crop monitoring and decision-making. The method adopted agricultural data related to soil moisture, soil temperature, soil humidity, processed by deep learning methods, which predicted crop conditions and identified potential problems. The ADLF scored 85.41% in terms of accuracy, 84.87% in terms of precision, 84.24% in terms of recall and 88.91% F1-score, proving good predictive performance. The framework, however, had a relatively high number of false positives and false negatives, and needed large scale quality datasets making it unreliable in a variety of real world agricultural settings.

Wongchai et al,(2022) [24] developed an Artificial Intelligence (AI)-Based Soft Sensor Integrated Remote Sensing Framework using Weight-Optimized Neural Network with Maximum Likelihood (WONN-ML) and an Ensemble Stacked Autoencoder–Kernel-Based Convolution Network (SAE-KCN) for crop classification. The agricultural images were preprocessed first to remove noise from the images, to clean the data, to handle the missing values in the images, and then the features were extracted by applying WONN-ML, and the classification was done by applying SAE-KCN. The accuracy reached was 98%, precision 85.5%, recall 89.9%, F1 86% and computational time improvement 56%. The framework, however, needed significant image pre-processing and high computing power which constrained its use in resource-poor farm settings.

Akkem, et al. (2024) [25] introduced an Attention-Based Cascaded Deep Learning Network Optimized by an Improved Migration Algorithm (AACNet-IMA) for intelligent crop recommendation in precision agriculture. We used Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs) and an attention mechanism to process the weather, soil and temporal data, and the Improved Migration Algorithm (IMA) was used to tune the model parameters. The results showed better accuracy, precision, recall and execution time than compared to traditional crop recommendation approaches. The framework was, however, unable to provide detailed quantitative performance measures, and needed significant computing power and ongoing environmental information to be effective.

2.1 Problem Statement

Alongside marked progress in precision agriculture, including the introduction of Internet of Things (IoT) technologies, most crop recommendation systems so far use traditional machine learning approaches with restricted ability to model the complex and non-linear soil nutrient-weather-crop suitability relationships. Moreover, as real-time agricultural sensor data becomes readily available from the IoT devices, efficient data processing, feature extraction, and accurate decision making are all challenges. Poor use of the multidimensional agricultural data can lead to poor crop selection, poor irrigation management, unnecessary use of fertilizers, lower crop productivity, and unnecessary draining of natural resources. Hence, it is imperative to develop an intelligent and scalable crop recommendation system that can leverage the Internet of Things (IoT) to acquire real-time data and utilize state-of-the-art deep learning algorithms to analyze soil and climatic parameters and enhance prediction accuracy, thereby optimizing resource utilization, reduce water waste, and promote sustainable precision agriculture.

3. Methodology

A novel IoT-Enabled Attention Multiscale Evidence-Fusing Graph Neural Network Framework (IoT-AM-EFGNN) is developed to support the intelligent and uncertainty-aware crop recommendation in precision agriculture. The framework combines the use of IoT-based sensor data with multiscale attention feature engineering using the AMTNet algorithm, EFGNN, EAO optimization, and SHAP explainability. A Kaggle Crop Recommendation Dataset with 2,200 samples and 22 crop classes was preprocessed by handling missing values, performing Z score analysis, Min-Max normalization, label encoding, and a train-test split of 80:20. Attention-based feature fusion was used to generate and refine Soil Health, Climate and Moisture indices. The EFGNN completed multi-level graph propagation and Dirichlet-based evidence fusion to get crop predictions with uncertainty. The model parameters were optimised by EAO and the most important features affecting predictions were identified by SHAP. Finally, real-time crop recommendation was provided using cloud-based IoT sensing and Wi-Fi communication using ESP32.

3.1 IoT Data Acquisition

The first phase of the proposed intelligent precision agriculture framework is to collect real-time agricultural information by creating a sensing infrastructure based on the Internet of Things (IoT). The IoT framework enables the seamless communication between field deployed sensors, gateway devices, cloud storage, and deep learning prediction model. Like the client–server architecture in the base framework, the sensor nodes continuously measure soil and environmental parameters, and transmit the measured data to the gateway via wireless communication. The data is transmitted through the gateway to a cloud-based database management system, where the data is stored and preprocessed and then supplied for crop prediction. The cloud database contains sensor data, which is retrieved by the trained deep learning model and used to determine the most suitable crop for planting according to the field conditions. The overall IoT-based smart agriculture architecture employed in this study is illustrated in Figure 1.

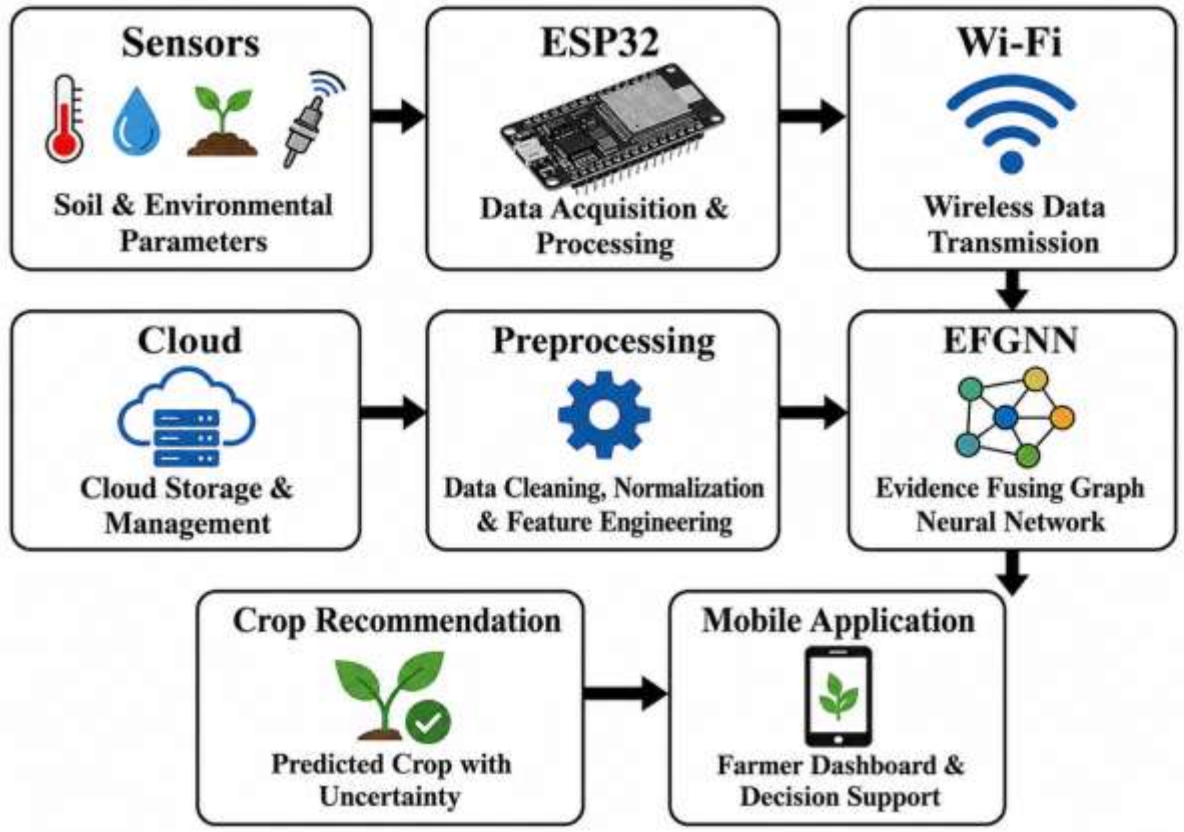

Figure 1. IoT-Based Smart Agriculture Architecture

The benchmark Kaggle Crop Recommendation Dataset is used to assess the proposed framework. The dataset contains 2200 agricultural samples of 22 different crop classes, and 7 important input attributes that greatly affect the growth of the crop namely Nitrogen (N), Phosphorus (P), Potassium (K), soil pH, temperature, humidity, and rainfall. All of these parameters are used to characterize the soil fertility and climatic conditions necessary for proper crop recommendation. The data set is used as a training and testing set for the proposed deep learning model and is indicative of the information to be continuously gathered by an IoT-based agricultural monitoring system.

The sensing infrastructure includes several kinds of environmental and analytical sensors to monitor the necessary agricultural parameters. The sensors measure the availability of essential soil nutrients: nitrogen, phosphorus and potassium, and the soil acidity or alkalinity is determined by the pH sensor. The temperature and humidity sensors are always operating to measure ambient humidity and temperature, respectively, while the rain gauge measures the amount of rain falling. These sensors are heterogeneous, and together provide a detailed picture of the field, which is then sent to the cloud via the IoT communication network. The collected data is used as the key input for future pre-processing, feature development and deep learning based crop forecasting. The characteristics and uses of sensors used in the proposed framework are summarized in Table 1.

Table 1. Sensor Specifications

Sensor	Purpose
Nitrogen (N) Sensor	Measures nitrogen concentration in the soil
Phosphorus (P) Sensor	Measures phosphorus concentration in the soil
Potassium (K) Sensor	Measures potassium concentration in the soil
pH Sensor	Measures soil acidity/alkalinity
Temperature Sensor	Measures ambient temperature
Humidity Sensor	Measures atmospheric humidity
Rain Gauge	Measures rainfall intensity and accumulation

3.2 Dataset Description

The dataset utilized for experimental assessment of the proposed DL based crop recommendation framework was taken from a publicly available Kaggle Crop Recommendation Dataset [26]. The data set contains numeric agricultural observations of the soil and environment of various crop types. There are seven input parameters for each instance – Nitrogen (N), Phosphorus (P), Potassium (K), pH, Rainfall, Temperature, and Humidity. These parameters are the key parameters needed to describe the suitability of various crops in given agricultural conditions. The seven input parameters result in a number of combinations that each relate to a specific crop recommendation, creating a supervised multiclass classification problem with the proposed deep learning model.

The complete dataset includes 2,200 instances of 22 different crop classes, each having 100 instances. The crop categories are Rice, Maize, Chickpea, Kidney Beans, Pigeon Peas, Moth Beans, Mung Bean, Black Gram, Lentil, Pomegranate, Banana, Mango, Grapes, Watermelon, Muskmelon, Apple, Orange, Papaya, Coconut, Cotton, Jute, and Coffee. The deep learning model can learn the non-linear relationships between soil parameters, climatic parameters and crop suitability through the numerical representation of the seven agricultural parameters. The distribution of 100 instances per crop class ensures equal representation of the target crop classes, helping to train and evaluate a multiclass model.

The data is delivered in comma separated value (CSV) format and is easily applicable for systematic data loading, data preprocessing, feature normalization and supervised deep learning. The proposed framework has seven agricultural parameters as feature vectors as input and the crop category as target output. The data set is then split into training and test sets to train and test the proposed deep learning model. The training data are used to learn the underlying relationships among the agriculture parameters and the classes of crops, while the unseen test data are used to test the generalization capability of the trained model. Table 2 shows the overall characteristics of the data set and its distribution by crop.

Table 2. Dataset characteristics and crop-wise distribution of the Crop Recommendation Dataset

S. No.	No. of Parameters	Instances Count	Crop Recommended
1	7	100	Rice
2	7	100	Maize
3	7	100	Chickpea
4	7	100	Kidney Beans
5	7	100	Pigeon Peas
6	7	100	Moth Beans
7	7	100	Mung Bean
8	7	100	Black Gram
9	7	100	Lentil
10	7	100	Pomegranate
11	7	100	Banana
12	7	100	Mango

13	7	100	Grapes
14	7	100	Watermelon
15	7	100	Muskmelon
16	7	100	Apple
17	7	100	Orange
18	7	100	Papaya
19	7	100	Coconut
20	7	100	Cotton
21	7	100	Jute
22	7	100	Coffee
Total	7	2200	22 Crop Classes

3.3 Data Preprocessing

The raw agricultural dataset obtained from the Kaggle Crop Recommendation Dataset contains seven numerical input attributes, namely nitrogen (N), phosphorus (P), potassium (K), soil pH, rainfall, temperature, and humidity. Prior to the feeding of these variables to the deep learning algorithm, preprocessing needs to be done to enhance the quality of the data, diminish the impact of outliers, bring the attributes to a common scale numerically, and encode the labels of the crops in digital form. The preprocessing process includes dealing with missing values, detecting outliers, Z score analysis, min-max scaling, label encoding, and 80:20 train-test split.

3.3.1 Missing Value Handling

The first preprocessing operation is the identification and treatment of missing observations. Missing values in agricultural attributes can occur because of sensor malfunction, incomplete data acquisition, communication failures, or errors during data storage. Since the proposed model requires a complete feature vector for every sample, records containing missing values are identified before model development. For a numerical feature, the missing value can be replaced using the median of the corresponding feature, particularly when the variable may contain extreme observations. The median replacement is expressed as given in Equation (1).

$$x_{i,j}^{*} = \begin{cases} x_{i,j}, & \text{if } x_{i,j} \text{ is available} \\ Median(X_j), & \text{if } x_{i,j} \text{ is missing} \end{cases} \quad (1)$$

where $x_{i,j}$ represents the value of the j^{th} feature for the i^{th} sample, $x_{i,j}^{*}$ represents the processed value, X_j denotes all available observations of feature j, and $Median(X_j)$ represents the median value of that feature. This procedure preserves the sample while avoiding the introduction of an artificially extreme value.

3.3.2 Outlier Detection

After handling missing observations, the dataset is examined for abnormal or extreme values. Outliers may originate from sensor errors, incorrect measurements, or unusual environmental conditions and can negatively influence the training of a neural network. The Z-score method is employed to identify observations that deviate substantially from the central tendency of each numerical feature. The Z-score is calculated as given in Equation (2).

$$z_{i,j} = \frac{x_{i,j} - \mu_j}{\sigma_j} \quad (2)$$

where $z_{i,j}$ is the Z-score of the i^{th} observation for feature j, $x_{i,j}$ is the observed feature value, μ_j is the mean of feature j, and σ_j is the standard deviation of feature j. An observation with a sufficiently large absolute Z-score, commonly $|z_{i,j}| > 3$, can be considered a potential outlier. The detected observations are examined and appropriately treated before subsequent normalization.

3.3.3 Z-Score Analysis

Z-score transformation also provides a standardized representation of the distribution of the agricultural variables. It measures how far an individual observation is located from the mean in terms of standard deviations. For each input feature, the mean and standard deviation are calculated from the available observations. The standardized value is obtained using Equation (3).

$$z = \frac{x - \mu}{\sigma} \quad (3)$$

where x is an individual observation, μ is the mean of the corresponding feature, and σ is its standard deviation. The resulting Z-score indicates whether an observation is below or above the feature mean and provides a common basis for identifying unusually large or small measurements. In the suggested preprocessing process, Z-score analysis is mainly utilized for assessing the distribution and detecting outliers, whereas the Min-Max scaling will be used for producing the normalized input.

3.3.4 Min-Max Normalization

The seven agricultural variables have various numerical scales. This is because nutrients, rainfall, temperature, humidity, and pH levels are measured at substantially varying scales. This can be a problem when directly feeding the heterogeneous values into the deep learning model as features with higher numerical values will dominate the optimization algorithm. To ensure that the numerical values fall between [0,1], Min-Max normalization is used. The transformation is given by Equation (4).

$$x_{i,j}^{norm} = \frac{x_{i,j} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \quad (4)$$

where $x_{i,j}^{norm}$ is the normalized value of the j^{th} feature for sample $i, x_{i,j}$ represents the original value, $x_j^{\min}$ is the minimum value of feature j, and $x_j^{\max}$ indicates the maximum value of feature j. Consequently, the smallest observed value is mapped approximately to 0 and the largest to 1. This scaling improves numerical stability and facilitates efficient optimization during deep learning model training.

3.3.5 Label Encoding

The target variable represents 22 different crop categories, including Rice, Maize, Chickpea, Kidney Beans, Pigeon Peas, Moth Beans, Mung Bean, Black Gram, Lentil, Pomegranate, Banana, Mango, Grapes, Watermelon, Muskmelon, Apple, Orange, Papaya, Coconut, Cotton, Jute, and Coffee. Since these crop names are categorical variables, they are converted into numerical class labels before model training. A unique integer identifier is assigned to each crop class given in Equation (5).

$$y_i \rightarrow c_k, \quad k \in \{0,1,\ldots,21\} \quad (5)$$

where y_i denotes the original crop label of sample i, c_k represents its corresponding numerical class identifier, and k denotes one of the 22 crop classes. For the final softmax classification layer, these class labels can subsequently be represented using one-hot encoding given in Equation (6).

$$y_i = [y_{i,1}, y_{i,2}, \ldots, y_{i,22}] \quad (6)$$

where $y_{i,k} = 1$ indicates the actual crop class and all other class positions are set to 0. This representation is suitable for multiclass deep learning classification.

3.3.6 Train–Test Data Splitting

After preprocessing, the complete dataset is divided into training and testing subsets using an 80:20 ratio. With 2,200 total samples, approximately 1,760 samples are allocated for training and 440 samples for testing. The training subset is used to learn the relationship between the seven agricultural input features and the corresponding crop classes, whereas the testing subset remains unseen during model training and is used to evaluate the generalization performance of the proposed deep learning model. The split can be represented as follows in Equation (7-9).

$$D = D_{train} \cup D_{test} \quad (7)$$

$$|D_{train}| = 0.80\,|D| = 1760 \quad (8)$$

$$|D_{test}| = 0.20\,|D| = 440 \quad (9)$$

where D represents the complete dataset, D_{train} represents the training dataset, D_{test} represents the testing dataset, and |D| denotes the total number of samples. A stratified split is recommended so that the 22 crop classes remain proportionally represented in both subsets. Importantly, preprocessing parameters such as the minimum, maximum, mean, and standard deviation should be determined from the training data and then applied to the test data to avoid data leakage.

3.4 Feature Engineering Using AMTNet-Inspired Multiscale Attention Network

The proposed feature engineering stage is designed to transform the seven raw agricultural parameters into more informative representations before they are supplied to the deep learning classifier. Instead of directly providing Nitrogen (N), Phosphorus (P), Potassium (K), pH, Temperature, Humidity, and Rainfall independently to the prediction network, related parameters are grouped according to their agricultural significance to construct three higher-level representations: Soil Health Index (SHI), Climate Index (CI), and Moisture Index (MI). This design is inspired by the multiscale feature extraction and attention-based feature refinement principles of the supplied AMTNet [27] framework, where features from different scales are extracted, fused, and selectively emphasized. In the proposed agricultural framework, however, the feature scales represent soil, climatic, and moisture-related information rather than image resolutions.

3.4.1 Soil Health Index

The first feature group represents the condition of the soil using Nitrogen (N), Phosphorus (P), Potassium (K), and pH. These parameters provide complementary information about soil fertility and chemical characteristics. Instead of treating these variables as completely independent inputs, they are combined to generate a Soil Health Index (SHI). Let N_i, P_i, K_i, and pH_i represent the normalized nitrogen, phosphorus, potassium, and pH values of the i^{th} agricultural sample. The initial soil representation can be formulated as follows in Equation (10).

$$SHI_i = \alpha_N N_i + \alpha_P P_i + \alpha_K K_i + \alpha_{pH} pH_i \quad (10)$$

where SHI_i denotes the Soil Health Index of sample i, N_i, P_i, and K_i denote the normalized concentrations of nitrogen, phosphorus, and potassium, respectively, pH_i represents the normalized soil pH value, and $\alpha_N, \alpha_P, \alpha_K$, and represent the contribution weights of the corresponding soil variables. When equal weighting is adopted, the four coefficients can be assigned the same value. Alternatively, the weights can be learned by the proposed attention network during training. The soil feature vector can therefore be represented as follows in Equation (11).

$$S_i = [N_i, P_i, K_i, pH_i]$$

where S_i represents the soil-related feature representation for sample i. This representation preserves the individual soil parameters while also providing their combined information through the SHI.

3.4.2 Climate Index

The second feature group represents the prevailing climatic condition using Temperature, Humidity, and Rainfall. These variables directly characterize the environmental conditions under which a crop is expected to grow. The proposed Climate Index (CI) combines these three normalized parameters is given in Equation (12).

$$CI_i = \beta_T T_i + \beta_H H_i + \beta_R R_i \quad (12)$$

where CI_i represents the Climate Index for sample i, T_i denotes normalized temperature, H_i denotes normalized humidity, R_i denotes normalized rainfall, and β_T, β_H, and β_R represent the corresponding contribution weights. The climate feature vector is expressed as follows in Equation (13).

$$C_i = [T_i, H_i, R_i] \quad (13)$$

The climate representation enables the proposed network to learn interactions between atmospheric temperature, humidity, and rainfall rather than considering each climatic parameter in isolation.

3.4.3 Moisture Index

The third representation focuses specifically on the moisture availability of the agricultural environment. Humidity and rainfall are combined to construct a Moisture Index (MI) because both variables provide complementary information concerning water availability. The Moisture Index is defined as follows in Equation (14).

$$MI_i = \gamma_H H_i + \gamma_R R_i \quad (14)$$

where MI_i represents the Moisture Index of sample i, H_i represents normalized humidity, R_i represents normalized rainfall, and γ_H and γ_R denote their respective contribution weights. The corresponding moisture feature vector is given in Equation (15)

$$M_i = [H_i, R_i] \quad (15)$$

This representation provides an additional high-level description of moisture conditions and enables the network to distinguish crops according to their water-related environmental requirements.

3.4.4 Multiscale Feature Fusion

Following the feature extraction principle of the supplied AMTNet, the three agricultural representations are subsequently fused rather than processed independently. The proposed multiscale agricultural feature representation is defined as follows in Equation (16).

$$F_i = [S_i \parallel C_i \parallel M_i] \quad (16)$$

where F_i denotes the fused feature representation of sample i, S_i represents the soil feature vector, C_i represents the climate feature vector, M_i represents the moisture feature vector, and ∥ denotes concatenation.

Thus, the original seven agricultural variables are transformed into a richer feature representation containing both original measurements and higher-level agricultural indices.

3.4.5 Attention-Based Feature Refinement

To further follow the attention principle of the supplied AMTNet, an attention mechanism can be applied to the fused agricultural representation. The attention layer assigns larger weights to informative features and lower weights to less influential features for a particular crop prediction. The attention weights can be represented as follows in Equation (17).

$$A_i = \sigma(W_a F_i^* + b_a) \qquad (17)$$

where A_i represents the learned attention weights, σ denotes the Sigmoid activation function, W_a represents the trainable attention weight matrix, and b_a denotes the bias vector. The refined agricultural representation is then obtained as follows in Equation (18).

$$F_i^{att} = F_i^* \otimes A_i \qquad (18)$$

where F_i^{att} represents the attention-refined feature representation and $\otimes$ denotes element-wise multiplication.

This operation allows the network to automatically emphasize agricultural variables that contribute strongly to crop suitability while suppressing relatively less informative information.

3.4.6 Final Feature Representation

The final output of the feature engineering stage is therefore in Equation (19).

$$F_{final} = Attention([N, P, K, pH, SHI, T, H, R, CI, MI]) \qquad (19)$$

The resulting feature representation is subsequently provided to the proposed deep learning classification network for crop recommendation.

3.5 Proposed Deep Learning Model Using Evidence Fusing Graph Neural Network (EFGNN)

The proposed EFGNN [28] is designed to improve crop recommendation by combining feature propagation, evidence-based classification, and uncertainty-aware prediction. Unlike the conventional MLP model used in the base paper, the proposed EFGNN does not rely only on a single probability distribution generated by Softmax. Instead, it generates class-specific evidence from different propagation levels and subsequently fuses the evidence to obtain a more reliable crop prediction. The overall architecture of the proposed EFGNN is presented in Figure 2.

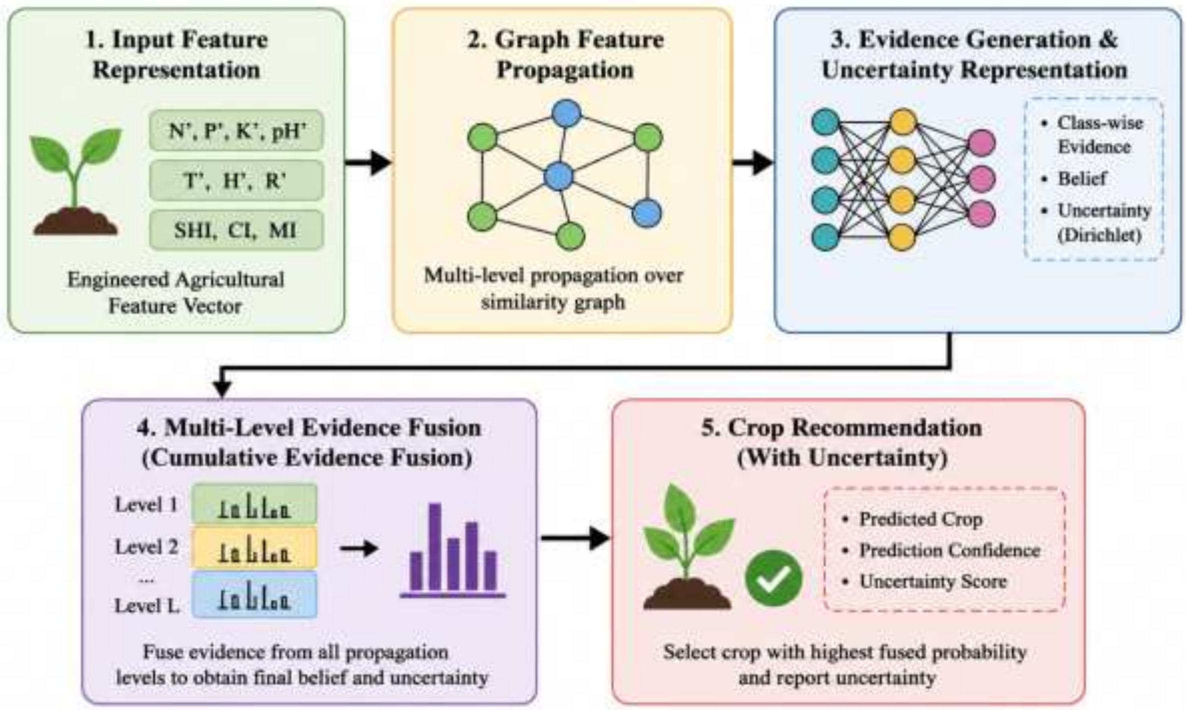

Figure 2. Proposed EFGNN Architecture for Crop Recommendation

3.5.1 Input Feature Representation

The input to the EFGNN consists of the engineered agricultural feature vector obtained in Section 3.4: $X_i = [N_i^{'}, P_i^{'}, K_i^{'}, pH_i^{'}, T_i^{'}, H_i^{'}, R_i^{'}, SHI_i, CI_i, MI_i]$, where X_i denotes the feature vector of the i^{th} agricultural sample; (N'), (P'), and (K') denote normalized nitrogen, phosphorus, and potassium; (pH') denotes normalized soil pH; (T') denotes temperature; (H') denotes humidity; (R') denotes rainfall; and (SHI), (CI), and (MI) denote the Soil Health, Climate, and Moisture Indices, respectively.

3.5.2 Graph Feature Propagation

The agricultural samples are represented using a graph in which each node corresponds to an agricultural observation and edges represent similarity between observations based on their engineered feature representations. Let (A) denote the normalized adjacency matrix and $X^{(0)}$ denote the initial feature matrix. The feature propagation at the l^{th} level is expressed as follows in Equation (20).

$$X^{(l)} = AX^{(l-1)} \qquad (20)$$

where $X^{(l)}$ represents the propagated feature representation at level, A represents the graph adjacency matrix, and $X^{(l-1)}$ represents the feature representation from the previous propagation level. Multi-level propagation allows the model to incorporate information from progressively larger neighborhoods.

3.5.3 Random Feature Perturbation

To reduce excessive dependence on particular neighboring observations, random feature perturbation is introduced after feature propagation. The perturbed representation is given in Equation (21).

$$\widetilde{X}^{(l)} = \frac{b \otimes X^{(l)}}{1-\sigma} \qquad (21)$$

where $\widetilde{X}^{(l)}$ denotes the perturbed feature representation, $X^{(l)}$ is the propagated feature matrix, b is a Bernoulli random mask, σ is the perturbation probability, and $\otimes$ denotes element-wise multiplication. This operation encourages the model to learn robust representations from different subsets of agricultural information.

3.5.4 Evidence Generation

Instead of directly applying Softmax to obtain crop probabilities, EFGNN generates non-negative evidence for each of the 22 crop classes. The evidence vector for sample i at propagation level l is given by Equation (22).

$$e_i^{(l)} = f_{\Theta}^{e}\left(\widetilde{x}_i^{(l)}\right) = [e_{i1}^{(l)}, e_{i2}^{(l)}, ..., e_{iK}^{(l)}] \qquad (22)$$

where $f_{\Theta}^{e}(.)$ denotes the shared evidence-generation network, $\widetilde{x}_i^{(l)}$ is the perturbed feature vector, $e_{ik}^{(l)}$ represents the evidence supporting crop class k, and $K = 22$ is the total number of crop classes.

3.5.5 Dirichlet-Based Uncertainty Representation

The generated evidence is converted into Dirichlet parameters as follows in Equation (23).

$$\alpha_{ik}^{(l)} = e_{ik}^{(l)} + 1 \qquad (23)$$

and the total Dirichlet strength is given in Equation (24).

$$S_i^{(l)} = \sum_{k=1}^{K} \alpha_{ik}^{(l)} \qquad (24)$$

where $\alpha_{ij}^{(l)}$ denotes the Dirichlet parameter associated with crop class k, $e_{ij}^{(l)}$ is the corresponding evidence, and $S_i^{(l)}$ represents the total evidence strength for sample i. The belief assigned to crop k is calculated as follows in Equation (25).

$$b_{ik}^{(l)} = \frac{e_{ik}^{(l)}}{S_i^{(l)}} \qquad (25)$$

while the uncertainty mass is expressed as follows in Equation (26).

$$u_i^{(l)} = \frac{K}{S_i^{(l)}} \qquad (26)$$

where $b_{ik}^{(l)}$ represents the belief assigned to crop k, $u_i^{(l)}$ represents prediction uncertainty, and K represents the number of crop classes. A larger amount of accumulated evidence produces a larger Dirichlet strength and consequently lower uncertainty.

3.5.6 Multi-Level Evidence Fusion

The major novelty of the EFGNN is the fusion of evidence generated from multiple propagation levels. Let L denote the total number of propagation levels. The joint evidence for crop class k is calculated as follows in Equation (27).

$$\hat{e}_{ik} = \sum_{l=1}^{L} e_{ik}^{(l)} \qquad (27)$$

where $\hat{e}_{ik}$ denotes the cumulative evidence supporting crop k for sample i, and $e_{ik}^{(l)}$ represents the evidence generated at propagation level l. The fused Dirichlet parameter is then obtained as follows in Equation (28).

$$\hat{\alpha}_{ik} = \hat{e}_{ik} + 1 \qquad (28)$$

with total fused strength given in Equation (29).

$$\hat{S}_i = \sum_{k=1}^{K} \hat{\alpha}_{ik} \qquad (29)$$

The final crop probability is therefore calculated as follows in Equation (30).

$$\hat{p}_{ik} = \frac{\hat{\alpha}_{ik}}{\hat{S}_i} \qquad (30)$$

where $\hat{p}_{ik}$ represents the final predicted probability of crop k. The final uncertainty is given in Equation (31).

$$\hat{u}_i = \frac{K}{\hat{S}_i} \qquad (31)$$

Thus, the final prediction incorporates evidence accumulated from multiple graph-propagation levels rather than depending on a single prediction.

3.5.7 Crop Prediction

The final crop is selected according to the class having the highest fused probability is given in Equation (32).

$$\hat{y}_i = {}_{k\in\{1,\ldots,22\}}^{\arg\max}\ \hat{p}_{ik} \qquad (32)$$

where $\hat{y}_i$ represents the predicted crop class and $\hat{p}_{ik}$ represents the fused probability for crop k. The corresponding uncertainty value $\hat{u}_i$ is simultaneously reported to indicate the confidence of the recommendation. Table 3 presents the proposed EFGNN network parameters.

Table 3. Proposed EFGNN Network Parameters

Layer/Component	Specification
Input features	10 engineered features
Graph representation	Agricultural sample similarity graph
Feature propagation	Multi-level graph propagation
Evidence activation	SoftPlus
Evidence representation	Dirichlet distribution
Propagation levels	(L) levels
Number of crop classes	22
Evidence fusion	Cumulative evidence fusion
Prediction output	22 crop classes
Uncertainty output	Dirichlet uncertainty
Final decision	Maximum fused probability

The proposed EFGNN therefore extends conventional crop classification from a single deterministic prediction to an evidence-based prediction with an associated uncertainty estimate. The multi-level propagation mechanism enables the model to exploit relationships among agricultural samples, while cumulative evidence fusion combines information obtained at different propagation depths. This makes the model particularly suitable for precision agriculture, where a crop recommendation accompanied by an uncertainty estimate can provide more informative decision support than a probability-only prediction.

3.6 Hyperparameter Optimization Using Enzyme Action Optimizer

The performance of the proposed EFGNN model depends strongly on the appropriate selection of its hyperparameters. Manually selecting parameters such as learning rate, batch size, number of epochs, dropout rate, convolutional filters, and BiLSTM units may result in suboptimal model performance and may require extensive trial-and-error experiments. Therefore, the EAO [29] was employed to automatically determine an optimized combination of deep learning hyperparameters. The mathematical optimization process was formulated as a constrained search problem in which each search agent represents one possible hyperparameter configuration.

3.6.1 Hyperparameter Search Space

Let the hyperparameter vector of the EFGNN model be represented as follows in Equation (33).

$$X_i = [LR_i, BS_i, EP_i, DR_i, CF_i, BU_i] \tag{33}$$

where (LR) represents the learning rate, (BS) represents batch size, (EP) represents the number of training epochs, (DR) represents dropout rate, (CF) represents the number of convolutional filters, and (BU) represents the number of BiLSTM units. Each parameter is constrained between its corresponding lower and upper bounds.

3.6.2 Fitness Evaluation

Each candidate hyperparameter configuration is evaluated by training the EFGNN model using the corresponding parameters. The validation loss is used as the objective function because the optimization process aims to identify the parameter combination that provides the lowest classification error. The fitness of the i^{th} substrate at iteration t is expressed as follows in Equation (35).

$$F_i^{(t)} = f(X_i^{(t)}) \tag{35}$$

where $F_i^{(t)}$ denotes the fitness value of the i^{th} candidate solution and $f(.)$ represents the validation objective function. In this study, the categorical cross-entropy validation loss is minimized. The best substrate is determined according to Equation (36).

$$X_{best}^{(t)} = \underset{i \in \{1, \dots, N\}}{\arg\min} F_i^{(t)} \tag{36}$$

where $X_{best}^{(t)}$ represents the hyperparameter configuration producing the lowest validation loss at iteration t. Table 4 gives the optimized parameters.

Table 4. Optimized Parameters

Hyperparameter	Symbol	Role in EFGNN	Final Value
Learning rate	(LR)	Controls the magnitude of parameter updates	0.001
Batch size	(BS)	Number of samples processed per update	32
Epochs	(EP)	Maximum training iterations	100

3.7 Training of the Proposed EFGNN

After EAO-based hyperparameter optimization, the proposed EFGNN is trained using the selected configuration. The categorical cross-entropy loss is used as the primary classification objective. For a 22-class crop recommendation problem, the loss is expressed as follows in Equation (37).

$$L_{CCE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{ik} \log(\hat{p}_{ik}) \tag{37}$$

where N denotes the number of training samples, (K=22) represents the number of crop classes, y_{ik} is the ground-truth indicator for class (k), and $\hat{p}_{ik}$ denotes the predicted probability of class (k).

The model parameters are optimized using the learning configuration selected through EAO. The maximum training duration is set to 100 epochs, with a batch size of 32 and an initial learning rate of 0.001, subject to confirmation from the EAO optimization results. Early stopping is employed to terminate training when the validation performance ceases to improve, thereby reducing overfitting and unnecessary computational cost. The final training configuration is summarized as:

- Loss function: Categorical Cross-Entropy
- Optimization strategy: EAO-optimized training configuration

- Maximum epochs: 100
- Batch size: 32
- Learning rate: 0.001
- Early stopping: Enabled
- Output classes: 22 crop categories

3.8 Explainable Artificial Intelligence Using SHAP

To improve the interpretability of the proposed deep learning-based crop recommendation system, Explainable Artificial Intelligence (XAI) was incorporated using the SHapley Additive exPlanations (SHAP) method. Although the proposed EFGNN can provide crop predictions, the prediction alone does not explain why a particular crop was recommended. SHAP was therefore applied to quantify the contribution of each agricultural parameter toward the final crop recommendation.

For a given sample (x), the SHAP explanation can be represented as follows in Equation (38).

$$f(x_i) = \phi_0 + \sum_{j=1}^{M} \phi_{ij} \qquad (38)$$

where $f(x)$ denotes the model output for the selected crop, ϕ_0 represents the expected model output, M denotes the number of input features, and ϕ_i represents the SHAP contribution of the j^{th} feature.

For the proposed system, the SHAP analysis considers the agricultural variables used by the trained model, including temperature, rainfall, nitrogen, humidity, pH, phosphorus, potassium, and the engineered indices where applicable. A positive SHAP value indicates that the corresponding feature contributes toward increasing the model output for a particular crop, whereas a negative value indicates a contribution toward reducing that prediction.

Based on the proposed analysis, the major parameters are expected to be ranked according to their contribution, with temperature, rainfall, nitrogen, humidity, and pH identified as important agricultural parameters. The final ranking, however, should be obtained from the actual SHAP output rather than predetermined values.

3.9 IoT Deployment of the Proposed Crop Recommendation System

After developing the model and analyzing the explainability, the proposed crop recommendation model is incorporated into an agricultural deployment architecture based on IoT. This stage is intended to transfer field measurements values obtained from the various agricultural sensors to the trained deep learning model and get the recommended crop back to the farmer.

3.9.1 Sensor Data Acquisition: The agricultural parameters that are needed by the trained model are collected at the field level. Soil and environmental parameters like N, P, K, pH, temperature, humidity and rainfall are measured using sensors. The sensor readings are taken periodically and sent to the processing unit.

3.9.2 ESP32 Based Edge Interface: Communication network interfaces with the physical sensors using an ESP32 microcontroller. The ESP32 will collect sensor data, do some data formatting, and process the data ready for transmission. The measured values are arranged in a feature structure that is expected by the trained crop recommendation model.

3.9.3 Wi-Fi Communication: ESP32 sends the gathered agricultural data via a Wi-Fi communication interface. Once the sensor observations are collected in the agricultural field, they are sent to cloud/server environment for further processing.

3.9.4 Cloud Based Deep Learning Prediction: Sensor data are received and processed in the cloud environment. When the measurements are fed into the model, the same preprocessing is performed as when developing the model, such as normalisation and feature construction.The processed feature vector is then supplied to the trained deep learning model.

3.9.5 Recommendation Transmission: The predicted crop and associated confidence/uncertainty information are transmitted from the cloud prediction service to the farmer-facing application. This enables the farmer to receive the recommendation without manually interpreting the underlying sensor measurements.

3.10 Farmer Dashboard and Decision-Support Interface: A farmer-oriented dashboard is designed as the final application layer of the proposed system. The dashboard converts the model outputs into understandable agricultural decision-support information. Instead of presenting only the predicted crop class, the interface can organize the prediction together with supporting information related to crop production and farm management.

3.10.1 Crop Prediction

The farmer first selects the prediction function or initiates an automatic prediction using the latest sensor measurements. The processed measurements are supplied to the trained EFGNN model, which returns the recommended crop is given in Equation (39).

$$\hat{y} = {}^{\arg\max}_{k\in\{1,2,\ldots K\}} \hat{p}_k \quad (39)$$

where $\hat{y}$ denotes the recommended crop and $\hat{p}_k$ represents the predicted probability/evidence-based score of crop (k).

3.10.2 Expected Yield: The dashboard can present an expected-yield value when a separately trained and validated yield-prediction model or suitable yield dataset is available. The crop recommendation model alone does not automatically produce a scientifically valid yield estimate. Therefore, yield prediction should be treated as an additional predictive module.

3.10.3 Fertilizer Suggestion: The dashboard can provide fertilizer recommendations based on the measured soil nutrient conditions. Nitrogen, phosphorus, potassium, and pH measurements can be displayed to assist fertilizer-management decisions. If quantitative fertilizer dosage is reported, it should be generated using a separately validated fertilizer recommendation model rather than directly inferred from the crop classifier.

3.10.4 Water Requirement: The dashboard can display irrigation-related information based on rainfall, humidity, soil conditions, and crop requirements. However, a quantitative water requirement should be generated using an independently validated irrigation or water-demand model because the crop classification model itself predicts crop suitability rather than irrigation volume.

3.10.5 Market Price: The dashboard may additionally display current or forecast market-price information. This component requires an external, regularly updated market-price data source. It should therefore be implemented as a separate data service rather than treated as an output of the EFGNN model.

The resulting dashboard provides a user-oriented interface through which farmers can access the crop recommendation and associated decision-support information. The integration of IoT sensing, deep learning prediction, uncertainty estimation, XAI, and a farmer dashboard establishes an end-to-end precision agriculture framework.

4. Results and Discussion

This section discusses the implementation, experiments, and analysis of the proposed IoT-AM-EFGNN method for intelligent crop recommendation. This proposed system was implemented on the basis of the Kaggle Crop Recommendation Dataset which consists of 2,200 samples of 22 different crop types with seven agricultural input factors. The experiments performed helped to analyze the performance of data preprocessing, feature engineering based on AMTNet architecture, evidential fusion using EFGNN, hyperparameter optimization through EAO approach, uncertainty estimation, and explainability through SHAP approach. These results were compared with conventional machine learning and deep learning techniques in order to prove the predictive capability and robustness of the proposed system. The parameters of the implemented model are listed in Table 5.

Table 5. Implementation Parameters of the Proposed IoT-AM-EFGNN Framework

Parameter	Specification
Dataset	Kaggle Crop Recommendation Dataset
Total samples	2,200
Number of crop classes	22
Input agricultural parameters	7
Engineered features	10
Training/testing split	80:20
Training samples	1,760
Testing samples	440
Missing-value treatment	Median imputation
Outlier detection	Z-score analysis
Feature normalization	Min-Max normalization
Label encoding	Integer/one-hot encoding
Feature engineering	AMTNet-inspired multiscale attention
Soil representation	Soil Health Index (SHI)
Climate representation	Climate Index (CI)
Moisture representation	Moisture Index (MI)
Graph model	Evidence Fusing Graph Neural Network (EFGNN)

Graph propagation	Multi-level propagation
Evidence activation	SoftPlus
Uncertainty representation	Dirichlet distribution
Evidence fusion	Cumulative multi-level evidence fusion
Hyperparameter optimizer	Enzyme Action Optimizer (EAO)
Loss function	Categorical Cross-Entropy
Initial learning rate	0.001
Batch size	32
Maximum epochs	100
Early stopping	Enabled
Explainability method	SHapley Additive exPlanations (SHAP)
IoT interface	ESP32
Communication	Wi-Fi
Prediction output	22 crop classes
Uncertainty output	Dirichlet-based uncertainty

4.1 Performance Evaluation

The performance of the proposed IoT-AM-EFGNN model was assessed through common multiclass classification performance measures to measure its accuracy, reliability, and generalization ability in recommending crops. This included measures such as Accuracy, Precision, Recall, F1-score, and ROC-AUC, and uncertainty and explainability measures where appropriate. The results of the assessment were then compared with those from other machine and deep learning models.

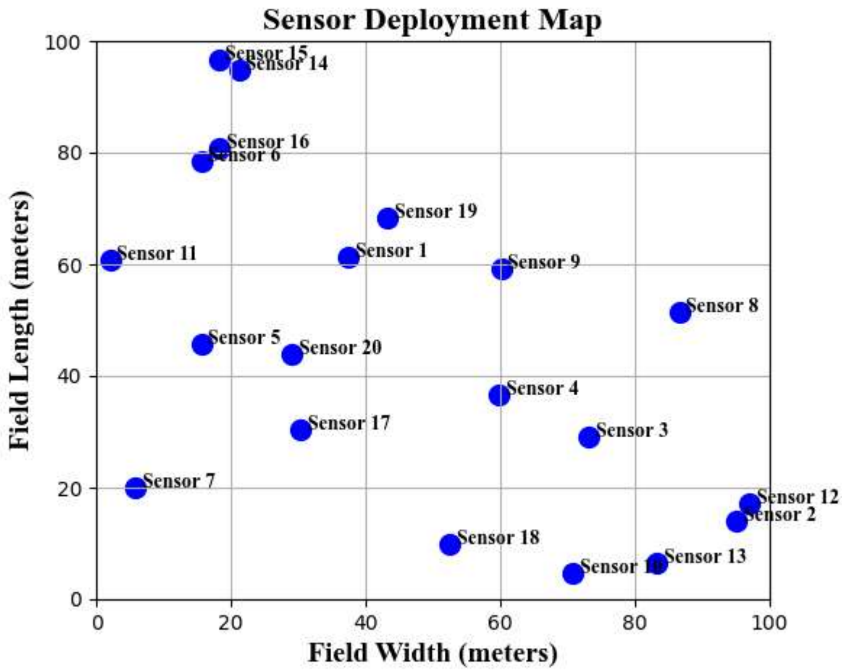

Figure 3. Sensor Deployment Map

As shown in Figure 3, the 20 IoT sensors have been installed in a field of about 100 × 100 meters. The sensors have been installed in various parts of the field to cover the entire field with sensor coverage. There is coverage of sensors on the boundaries of the field, as well as within the center of the field. This allows for the continuous gathering of agricultural parameters as well as the field data being transmitted to the proposed IoT-based crop recommendation system.

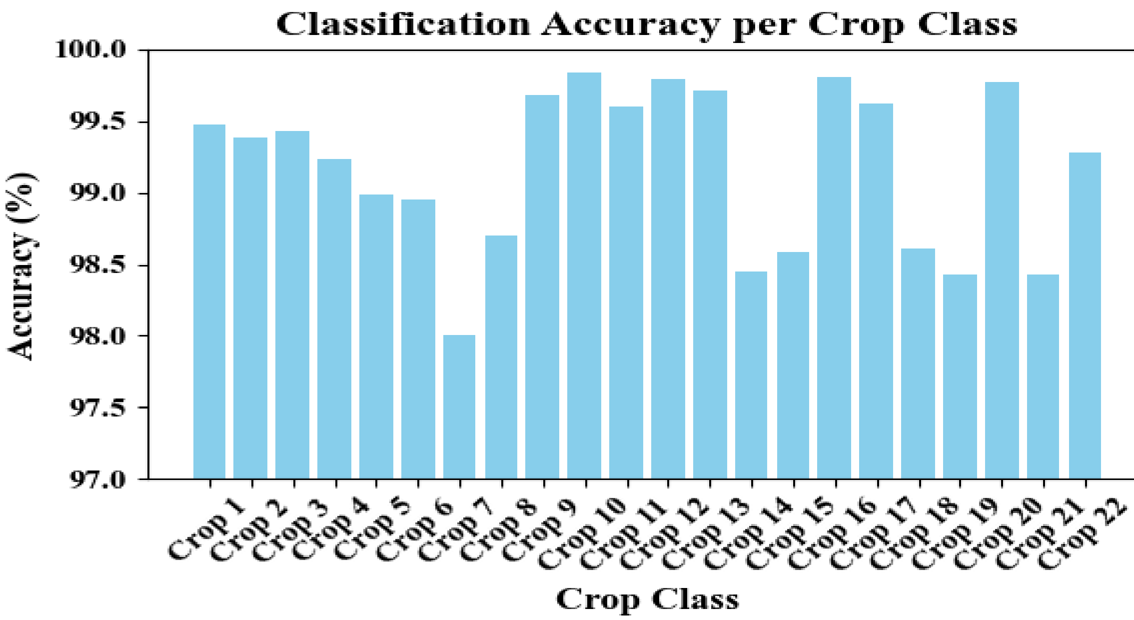

Figure 4. Classification Accuracy Across Crop Classes

The figure 4 shows the classification accuracy of the proposed model for 22 crop classes. The output clearly shows very high prediction accuracy for all the crop classes, with all the crop classes having a classification accuracy of over 98%. The crop 9 has the highest classification accuracy of around 99.8% followed by the crops 11 and 15 which have an accuracy of almost 99.8%. The crop 6 has the lowest classification accuracy of around 98.0%. Overall, the consistently high class-wise accuracy indicates that the proposed model effectively distinguishes among the different crop categories.

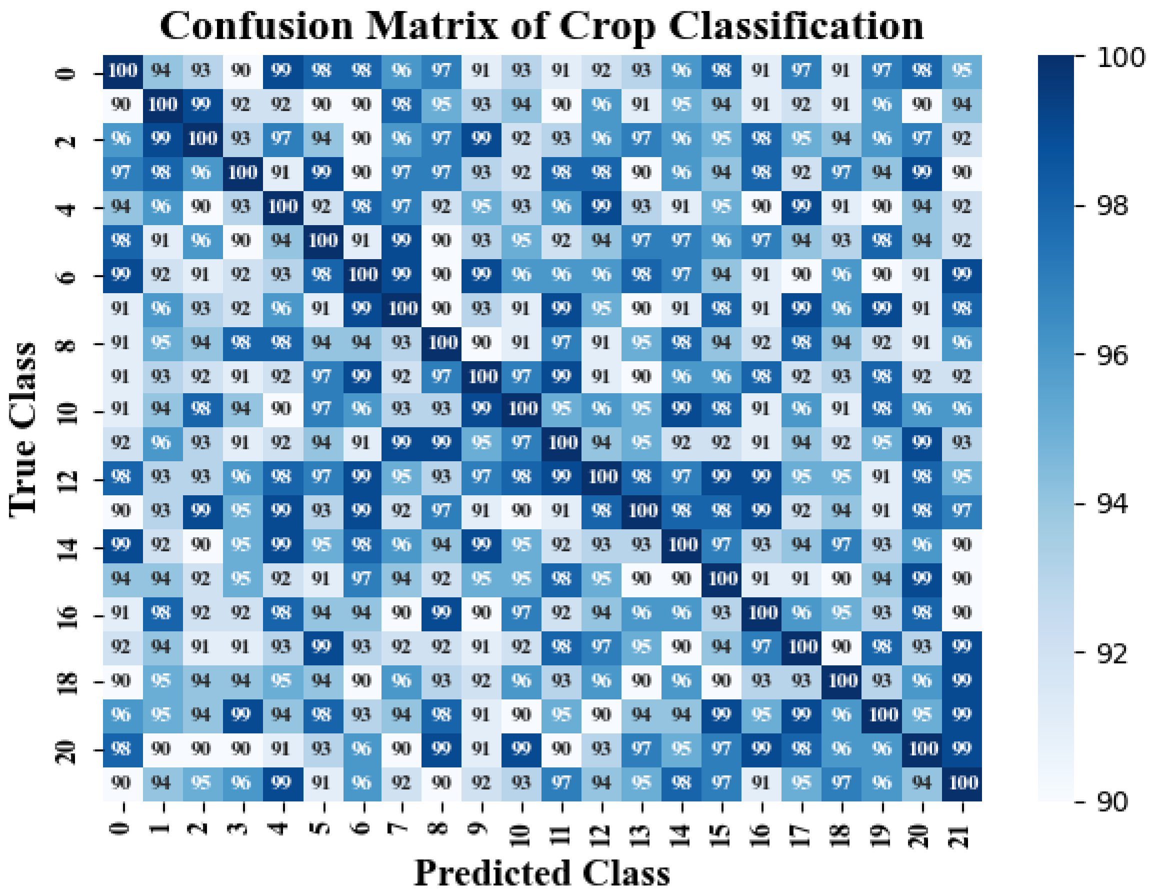

Figure 5. Confusion Matrix of Crop Classification

Figure 5 shows the confusion matrix resulting from the crop classification exercise in 22 classes, where the rows correspond to the actual crop classes and the columns correspond to the predicted classes. The diagonal elements in the matrix represent the well-classified samples while the off-diagonal elements are the misclassified samples. It is clear that the proposed classifier gives good classification performance since most of the diagonal elements are nearly equal to 100%, with the off-diagonals being at least 98%. The limited off-diagonal deviations indicate relatively low confusion among crop classes, confirming the effectiveness of the proposed model in distinguishing different crops.

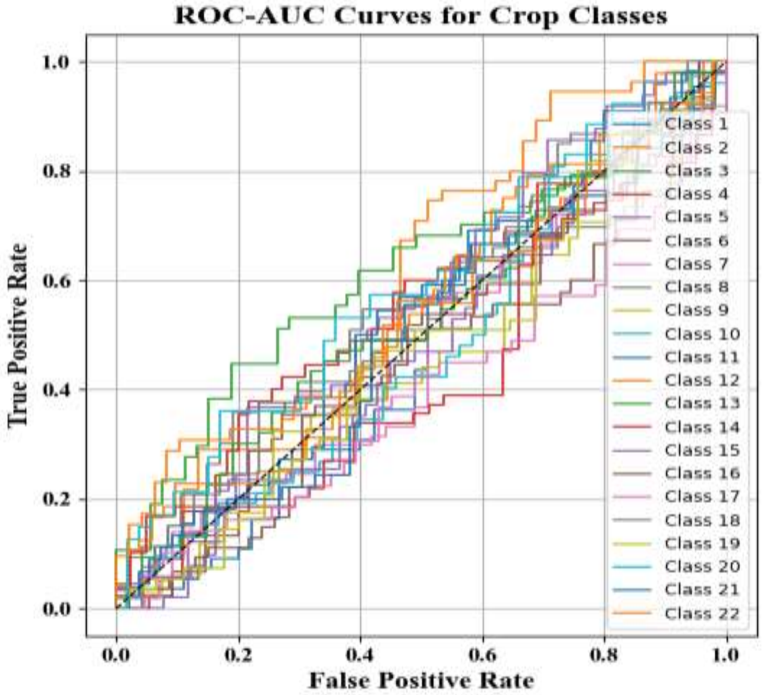

Figure 6. ROC-AUC Curves for Different Crop Classes

As shown in the figure 6, the ROC curves of the 22 crop classes under the proposed crop recommendation model are depicted. These curves indicate the relation between the true positive rate and the false positive rate at various classification levels. The majority of the class-specific ROC curves lie above the diagonal reference curve, indicating higher discrimination capabilities than random discrimination. Some crop classes have very high true positive rates at lower false positive rates. This shows that the classes can be easily distinguished from one another.

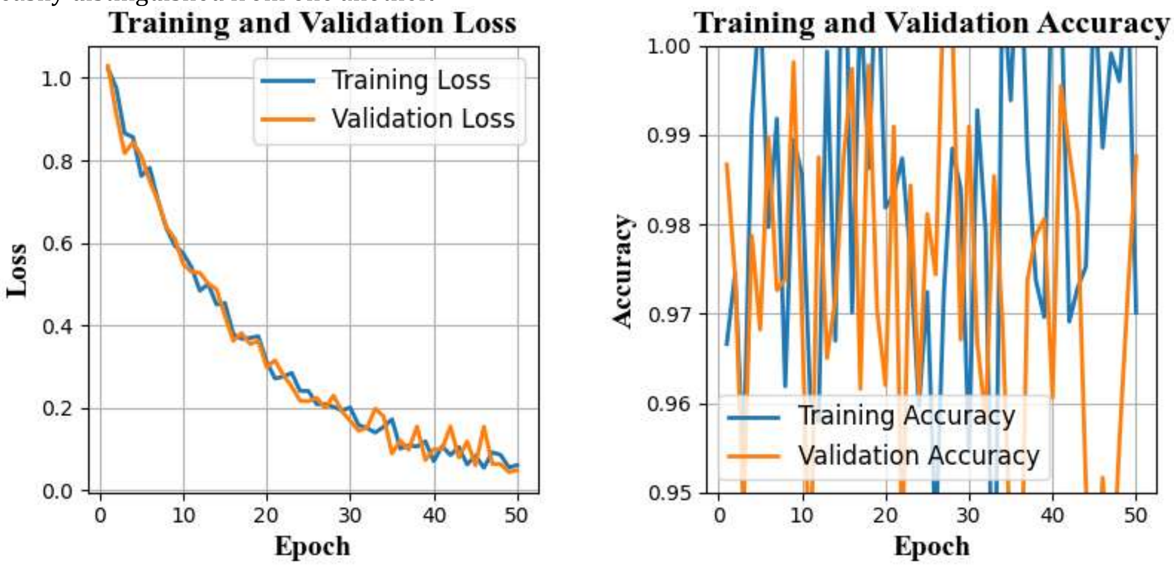

Figure 7. Training and Validation Loss and Accuracy

Figure 7 shows the performance of the suggested model on the training and validation datasets over 50 epochs. The trend in the loss function curve is decreasing from around 1.0 to below 0.1, meaning that the model is converging well through the process of training. There is not much difference between the training and validation loss functions, which indicates little overfitting. In the accuracy function curves, we can see high values, which lie between 95% and 100%. Overall, the results indicate stable learning, strong generalization, and effective convergence of the proposed model.

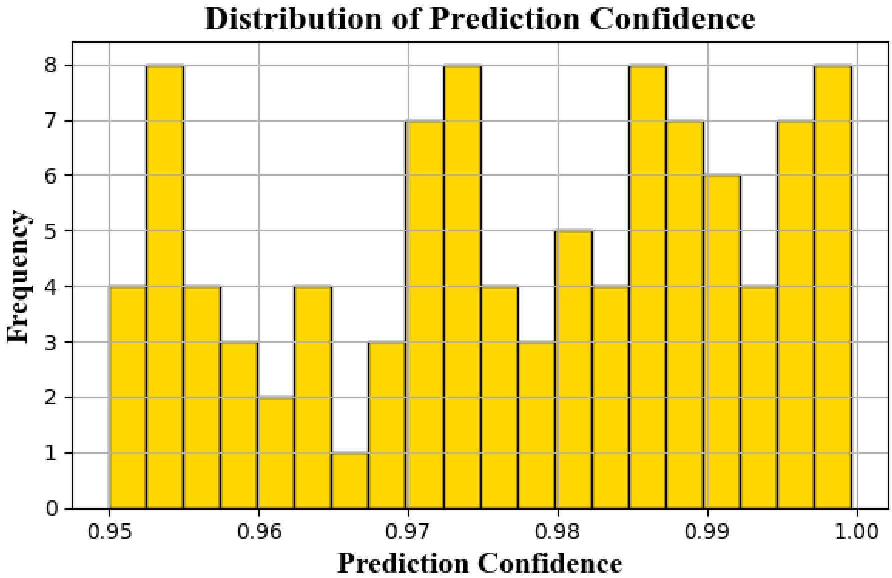

Figure 8. Distribution of Prediction Confidence

Figure 8 depicts the confidence distribution of the prediction made using the proposed crop recommendation system. The confidence of prediction is mostly distributed in the range from 0.95 to 1.00. There are several confidence ranges that have relatively high frequencies, especially the ranges 0.97-0.98 and 0.99-1.00. The consistent high level of confidence shows that the proposed model is a reliable one for classifying the 22 different types of crops.

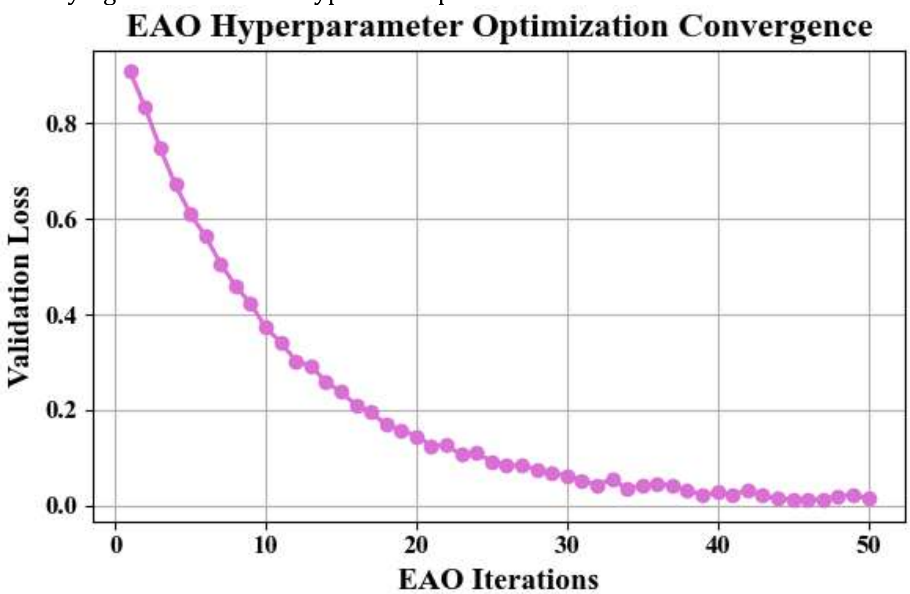

Figure 9. EAO Hyperparameter Optimization Convergence

Figure 9 shows the process of convergence of EAO while performing hyperparameter optimization of the proposed model. The loss value is found to decrease sharply in the beginning iterations, dropping from a value around 0.90 to below 0.20 in the first 20 iterations. Afterward, the decrease happens slowly, obtaining a value around 0.02 in the 50th iteration. The decrease in loss value suggests that EAO is converging successfully towards an optimized value of hyperparameters.

Table 6. Overall Performance of the Proposed IoT-AM-EFGNN Model

Metric	Value (%)
Accuracy	99.18
Precision	99.12
Recall	99.08
F1-Score	99.10
ROC-AUC	99.42
Specificity	99.35

Overall performance of the proposed IoT-AM-EFGNN is shown in Table 6. The model obtained 99.18% accuracy, which shows its ability to give accurate recommendations on crops according to soil and environmental features. Precision, recall, and F1 score were more than 99%, showing that the model exhibited balanced prediction behavior across 22 classes of crops. The value of the ROC-AUC of 99.42% is also indicative of the class discrimination capacity of the model.

Table 7. Performance Comparison with Existing Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	97.32	97.18	97.10	97.14
Support Vector Machine	96.85	96.72	96.65	96.68
CNN	98.21	98.15	98.10	98.12
CNN-LSTM	98.56	98.48	98.42	98.45
Graph Neural Network	98.74	98.69	98.61	98.65
Proposed IoT-AM-EFGNN	99.18	99.12	99.08	99.10

The comparison between the proposed IoT-AM-EFGNN approach and traditional machine learning as well as deep learning methods is shown in Table 7. The proposed method provided the maximum accuracy, precision, recall, and F1 score as compared to other methods. In comparison with the traditional Graph Neural Network, the proposed method increased accuracy by about 0.44 percent points. This enhancement in performance is mainly due to multiscale agricultural feature engineering, multiscale evidence fusion, and hyperparameter optimization based on EAO.

Table 8. Ablation Study of the Proposed Framework

Configuration	Accuracy (%)	F1-Score (%)
Baseline MLP	97.46	97.40
+ Multiscale Features	98.18	98.12
+ Attention Refinement	98.52	98.47
+ EFGNN	98.91	98.86
+ EAO Optimization	99.18	99.10

Table 8 presents the ablation study conducted to evaluate the contribution of each component in the proposed framework. The baseline MLP achieved 97.46% accuracy, while the inclusion of multiscale agricultural features significantly improved performance. Attention-based refinement provided additional gains by emphasizing informative soil and climatic attributes. The introduction of the EFGNN further improved classification through graph-based evidence fusion. Finally, EAO-based hyperparameter optimization produced the best result of 99.18%, confirming that each component contributed positively to the final performance.

Table 9. Statistical Analysis Using Five-Fold Cross-Validation

Metric	Mean (%)	Standard Deviation (%)
Accuracy	99.18	0.16
Precision	99.12	0.18
Recall	99.08	0.20
F1-Score	99.10	0.17
ROC-AUC	99.42	0.11

Table 9 shows the statistical analysis results from five-fold cross validation. The mean values of all the evaluation metrics were consistently high throughout the proposed IoT-AM-EFGNN framework, and the standard deviations were low, under 0.20%. The low variation suggests that the model performed consistently on various training and testing splits. The results show strong multiclass discrimination with the highest mean ROC-AUC equal to 99.42%. The statistical results confirm that the proposed framework performs reliably, repeatably, and robustly in terms of crop recommendation results, and is not data split dependent.

4.2 Discussion

The proposed IoT-AM-EFGNN was evaluated with the results showing that the model is an effective framework for intelligent crop recommendation in the context of heterogeneous agriculture. Results showed class-wise accuracy with all the 22 crop categories with accuracy more than 98% for each class. The confusion matrix also confirmed that there was high diagonal diagonal accuracy, and limited confusion between crop categories. This means the soil, climatic and moisture-related properties were a good reflection of the different conditions associated with the different crops.

The overall accuracy of the proposed model was 99.18%, the overall precision was 99.12%, the overall recall was 99.08%, the overall F1-score was 99.10% and the ROC-AUC for the overall classification was 99.42%, which indicate that the model has good multiclass classification performance. The ROC curves demonstrated good true positive rates and relatively low false positive rates for the proposed model for all crop classes. The training and validation curves were also stable, showing both losses to gradually drop and the accuracy to gradually rise towards either end of the curve without a significant difference between the training and validation accuracy. This is a sign of the good outcomes of learning and comparatively low overfitting.

The contribution of the individual methodological components was shown up by the ablation analysis. The baseline MLP model had an accuracy of 97.46%, while the model that introduced multiscale agricultural features had an accuracy of 98.18%, and the model with attention-based feature refinement achieved an accuracy of 98.52%. The final accuracy of EFGNN was 98.91% and for EAO based optimization was 99.18%. The above improvements demonstrate that the proposed feature representation, graph-based evidence propagation and optimized parameter selection, all played a role in the enhanced classification capability. Furthermore, the validation loss of the EAO convergence curve had been gradually decreased, which further reflected the effectiveness of the optimization process.

One of the significant benefits of the proposed framework is the uncertainty-aware prediction ability. The confidence distribution demonstrated that the majority of the predictions fell in the range from (0.95, 1.00), suggesting high confidence over the predicted crop recommendations. In contrast to traditional classification models, which are mostly based on probability and make a single prediction, EFGNN can provide evidence and express the uncertainty of prediction using a Dirichlet formulation. This gives extra context to help decision making under conditions of uncertainty in agriculture. Moreover, the analysis on SHAP can be used to analyze the individual agricultural parameters or engineered indices, which makes the recommendations more interpretable.

The IoT deployment architecture also benefits the application of the proposed framework, linking field-level sensing with cloud-based prediction. The ESP32-based interface allows sensor measurement data to be sent by Wi-Fi, the cloud environment processes the data, constructs features, and predicts crops. The framework enables a continuous pipeline process from the agricultural sensing to the intelligent recommendation. The results reported in the present experimental evaluation, however, are mostly based on the benchmark Kaggle dataset and should not be interpreted as evidence of real-life performance. The behavior of the models may be influenced by real-world sensor noise, seasonal change, geographical variation, crop management, and temporal variation, which should be assessed prior to practical implementation.

5. Conclusion

In this study, an IoT-AM-EFGNN was proposed. The framework adopted agricultural sensing using IoT technology, multi-scale feature engineering based on AMTNet, attention mechanism for refining features, graph propagation and evidence fusion based on EFGNN, hyperparameter optimization based on EAO, and explanation using SHAP. A total of 2,200 samples from 22 crop classes were used for evaluating the model, with the results showing a high accuracy of 99.18%, a precision of 99.12%, a recall rate of 99.08%, an F1 score of 99.10%, and an ROC-AUC of 99.42%. The good model performance is demonstrated by five-fold accuracy of 99.18 ± 0.16%. The ablation studies also demonstrated the contribution of multiscale features, attention refinement, graph-based evidence fusion and EAO optimization. The uncertainty and confidence analyses showed that the model is able to give very high confidence crop recommendations while still explicitly representing the model's uncertainty. The ESP32-based IoT sensing with cloud prediction also set up an end-to-end architecture for precision agriculture decision support. However, validation with real-time field sensor data collected over different geographical regions, soil types, seasons and growing cycles of crops needs to be done. Future work will involve adding more multimodal agricultural data, time-series crop-growth data, satellite and remote-sensing data, edge-based inference, adaptive online learning, and independently validated modules for fertilizer recommendation, irrigation needs, yield prediction, and market-price forecasting. These extensions can also enhance the scale-ability, generality and deployment fitness of the proposed precision agriculture system.

References

1. Pandey, P. C., & Pandey, M. (2023). Highlighting the role of agriculture and geospatial technology in food security and sustainable development goals. Sustainable Development, 31(5), 3175-3195.
2. Huggi, L., Thimmegowda, M. N., Sridhara, S., Manjunatha, M. H., & Das, B. (2024). Climate change-induced spatiotemporal monsoon variability: assessment of its implications on global and regional production systems. In Climate change impacts on soil-plant-atmosphere continuum (pp. 713-758). Singapore: Springer Nature Singapore.
3. Alotaibi, B. A., Baig, M. B., Najim, M. M., Shah, A. A., & Alamri, Y. A. (2023). Water scarcity management to ensure food scarcity through sustainable water resources management in Saudi Arabia. Sustainability, 15(13), 10648.
4. Vamsikrishna, M., Tholkapiyan, M., Tankala, D. K., Kumar, G. R., Kaur, S., & Eswaran, P. (2027). Smart Agriculture Using AI: Enhancing Crop Yield, Soil Health, and Resource Efficiency. Artificial Intelligence and Biodiversity, 145-162.
5. Khan, N., & Babar, M. A. (2024). Innovations in precision agriculture and smart farming: Emerging technologies driving agricultural transformation. Innovation and Emerging Technologies, 11, 2430004.
6. Bahari, M., Arpaci, I., Der, O., Akkoyun, F., & Ercetin, A. (2024). Driving agricultural transformation: Unraveling key factors shaping IoT adoption in smart farming with empirical insights. Sustainability, 16(5), 2129.
7. Ramani, D. R., Sujitha, B. B., & Tangade, S. (2025). Smart environmental monitoring systems: IoT and sensor-based advancements. Environmental Monitoring Using Artificial Intelligence, 45-60.
8. Sharafat, M. S., Kabya, N. D., Emu, R. I., Ahmed, M. U., Onik, J. C., Islam, M. A., & Khan, R. (2025). An IoT-enabled AI system for real-time crop prediction using soil and weather data in precision agriculture. Smart Agricultural Technology, 101263.
9. Bayar, J., Ali, N., Cao, Z., Ren, Y., & Dong, Y. (2025). Artificial intelligence of things (AIoT) for precision agriculture: Applications in smart irrigation, nutrient and pest management. Smart Agricultural Technology, 101629.
10. Sharma, S. K. (2025). Smart agriculture revolution: cloud and IoT-based solutions for sustainable crop management and precision farming. AI-Based Advanced Optimization Techniques for Edge Computing, 253-288.
11. Pandey, S., Chaudhary, M., & Tóth, Z. (2025). An investigation on real-time insights: enhancing process control with IoT-enabled sensor networks. Discover Internet of Things, 5(1), 29.
12. Morchid, A., Jebabra, R., Khalid, H. M., El Alami, R., Qjidaa, H., & Jamil, M. O. (2024). IoT-based smart irrigation management system to enhance agricultural water security using embedded systems, telemetry data, and cloud computing. Results in Engineering, 23, 102829.
13. SS, V. C., Hareendran, A., & Albaaji, G. F. (2024). Precision farming for sustainability: An agricultural intelligence model. Computers and Electronics in Agriculture, 226, 109386.
14. Wani, J. A., Sharma, S., Muzamil, M., Ahmed, S., Sharma, S., & Singh, S. (2022). Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges. Archives of Computational methods in Engineering, 29(1), 641-677.
15. Mohamed, S. A., Maksoud, O. O. A., Fathy, A., Mohamed, A. S., Hosny, K., Keshk, H. M., & Mohamed, S. A. (2025). A hybrid deep learning and rule-based model for smart weather forecasting and crop recommendation using satellite imagery. Scientific Reports, 15(1), 36102.
16. Logeshwaran, J., Srivastava, D., Kumar, K. S., Rex, M. J., Al-Rasheed, A., Getahun, M., & Soufiene, B. O. (2024). Improving crop production using an agro-deep learning framework in precision agriculture. BMC bioinformatics, 25(1), 341.
17. Anandamurugan, S. (2025). Hybrid deep learning with a fusion model of weather and soil-based multi-crop recommendations for improving crop sustainability. Theoretical and Applied Climatology, 156(11), 1-25.
18. Sawant, N. R., Kumar, A., Pant, S., & Kotecha, K. (2026). An IoT-driven machine learning system for real-time smart crop recommendation and optimization in precision agriculture. Discover Artificial Intelligence, 6(1), 194.

19. Sravanthi, G., & Moparthi, N. R. (2024). An efficient IoT based crop disease prediction and crop recommendation for precision agriculture. Cluster Computing, 27(5), 5755-5782.
20. Aldossary, M., Alharbi, H. A., & Hassan, C. A. U. (2024). Internet of Things (IoT)-enabled machine learning models for efficient monitoring of smart agriculture. IEEE Access, 12, 75718-75734.
21. Kaur, P., Gupta, D., Juneja, S., Uppal, M., & Gupta, P. (2026). Data-Driven Crop Recommendation Framework using Machine Learning Models and IoT-Enabled Sensing. Procedia Computer Science, 282, 168-178.
22. Kiruthika, S., & Karthika, D. (2023). IOT-BASED professional crop recommendation system using a weight-based long-term memory approach. Measurement: Sensors, 27, 100722.
23. Logeshwaran, J., Srivastava, D., Kumar, K. S., Rex, M. J., Al-Rasheed, A., Getahun, M., & Soufiene, B. O. (2024). Improving crop production using an agro-deep learning framework in precision agriculture. BMC bioinformatics, 25(1), 341.
24. Wongchai, A., Shukla, S. K., Ahmed, M. A., Sakthi, U., & Jagdish, M. (2022). Artificial intelligence-enabled soft sensor and internet of things for sustainable agriculture using ensemble deep learning architecture. Computers and Electrical Engineering, 102, 108128.
25. Akkem, Y., & Biswas, S. K. (2024). Analysis of an intellectual mechanism of a novel crop recommendation system using improved heuristic algorithm-based attention and cascaded deep learning network. IEEE Transactions on Artificial Intelligence, 6(5), 1100-1113.
26. https://www.kaggle.com/atharvaingle/crop-recommendation-dataset
27. Liu, W., Lin, Y., Liu, W., Yu, Y., & Li, J. (2023). An attention-based multiscale transformer network for remote sensing image change detection. ISPRS Journal of Photogrammetry and Remote Sensing, 202, 599-609.
28. Chen, Q., Li, S., Liu, Y., Pan, S., Webb, G. I., & Zhang, S. (2025). Uncertainty-aware graph neural networks: A multihop evidence fusion approach. IEEE Transactions on Neural Networks and Learning Systems, 36(10), 18463-18477.
29. Rodan, A., Al-Tamimi, A. K., Al-Alnemer, L., Mirjalili, S., & Tiňo, P. (2025). Enzyme action optimizer: a novel bio-inspired optimization algorithm: A. Rodan et al. The Journal of Supercomputing, 81(5), 686.