Proceedings of the Pakistan Academy of Sciences: A Physical and Computational Sciences 60(3): 37-48 (2023)

ISSN (Print): 2518-4245; ISSN (Online): 2518-4253
https://doi.org/10.53560/PPASA(60-3)825

Pakistan Academy of Sciences

Research Article

An Intelligent Decision Support System for Crop Yield Prediction Using Machine Learning and Deep Learning Algorithms

Maryum Bibi[1], Saif-Ur-Rehman[1*], Khalid Mahmood[2], and Rana Saud Shoukat[1]

[1]University Institute of Information Technology,
PMAS - Arid Agriculture University Rawalpindi, Pakistan
[2]Institute of Computing and Information Technology, Gomal University, D.I. Khan, Pakistan

Abstract: Agriculture is crucial to economic growth and development. Crop yield forecasting is critical for food production which includes vegetables, fruits, flowers, and cattle. Artificial Intelligence (AI) is rising in agriculture, providing farmers with real-time or long-term insights about their fields. It allows us to identify the areas that require irrigation, fertilization, or pesticide treatment. Statistical models struggle to track complex relationships in crop yields due to numerous factors. Machine Learning (ML) and Deep Learning (DL) algorithms can solve this problem by training themselves in these relationships, enabling accurate predictions in agricultural yield prediction methods. Predicting product performance in agriculture is challenging due to various factors, but profit forecasting improves decision-making, production, economics, and food safety. The present study focuses on the use of ML and DL algorithms to suggest a novel decision support system for crop yield prediction with the objectives to develop a robust, accurate model, investigate algorithm effectiveness, and create a user-friendly system for informed crop production decisions. According to the results, the developed system is capable of making precise predictions, which can support farmers in making better decisions about how to manage their crops. The simulation results demonstrate that the intelligent decision support system proposed for crop yield prediction using ML and DL algorithms is capable of achieving high accuracy and precision. The system can be used to help farmers make better decisions about crop planting and management, which can lead to increased crop yields and profits. The results of our experiment show that our model is better than the others and it achieves an accuracy of 99.82 %. Additionally, we utilized ML to condense the input space while preserving high accuracy.

Keywords: Machine Learning Algorithm, Deep Neural Network, Deep Learning Algorithm, Crop Yield Forecasting, Artificial Intelligence, Agricultural Productivity.

1. INTRODUCTION

Digitization is having a major impact on many different areas of life including medicine, agriculture, consensus platforms, and weather forecasting, etc., [1]. Weather affects agrarian yield, food security, GDP, and environmental protection. ML and DL techniques improve the prediction of agricultural production by capturing complex correlations between crop output and environmental parameters [2]. Remote sensing improves agricultural yield prediction, but noisy data challenges accuracy [3]. ML enhances agricultural planning and production, but challenges remain in dataset quality, algorithms, and decision-making integration [4, 5]. AI aids industrial sectors, while agriculture faces climate change risks. Climate change impacts agricultural industry, affecting livelihoods and food security [6]. Post-hoc methods clarify trained predictions, process methods improve interpretation. Agricultural planning optimizes land use and yields using machine learning algorithms [7, 8]. Precision agriculture uses GPS, remote sensing, and internet technologies to manage crops, reduce fertilizers, pesticides, and water usage [9]. As such the precision agriculture ensures crop-specific product quality control. Maximizing resources and predicting yields using ML algorithms and DL techniques. The aim is to make the most of what we have while conserving resources. Yield predictions

Received: June 2023; Accepted: September 2023
*Corresponding Author: Saif-Ur-Rehman <Saif@uaar.edu.pk>

were made using ML algorithms and DL such as linear regression and multiple regressions. ML techniques such as RF, SVM, multiplexer, logistic regression, and DL techniques such as DCNN and LSTM can provide quick and accurate solutions to this problem [10]. Agricultural sustainability, food availability, productivity, and farmers' cultural familiarity are imperative for food safety and food security [11-13]. The UN Sustainable Development Goals (SDGs) for 2030 include zero hunger and sustainable agriculture. DL techniques predict crops using Convolution Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) [14, 15]. Smart farming consists of advance precision in agriculture with intelligent, remote solutions [16-18]. The rubber market grows because deep learning aids rubber yield forecasts with accuracy and robustness [19]. Traditional methods overlook complex factors affecting yields [20]. The parameters that have had a major impact on crops are water, ultraviolet (UV) radiations, pesticides, fertilizers, and the area of land covered by the area. A proposed ML model illustrating the use of NN and related ANN algorithms was evaluated. The dataset consists of 140 data points that represent the effect of the attribute on crop yield. Predicting crop yields is crucial for informed decisions in agriculture, but traditional methods rely on limited statistical models that are unable to capture the complex relationship between yield and factors like weather, soil conditions, and pests [21, 22]. Traditional methods struggle with predicting crop yields due to limited statistical models. DL effectively performs image classification, speech recognition, and crop yield forecasting using CNN and weather parameters [23]. US soybean yield prediction using neural networks and CNN outperforms remote sensing by 15 % on average MAPE (mean absolute percentage error), incorporating spatiotemporal features. [24, 25]. IoT technology offers diverse applications in smart homes, cities, traffic management [26, 27]. Technology integrates agricultural equipment for optimal planting and fertilization decisions [28]. Smart machines enhance plant and animal growth monitoring accuracy [29, 30].

Satellite missions, remote sensing sensors, big data, artificial intelligence, and machine learning offer new opportunities for understanding crop processes and monitoring yield using remote sensing. Figure 1 shows that a thorough site survey is crucial

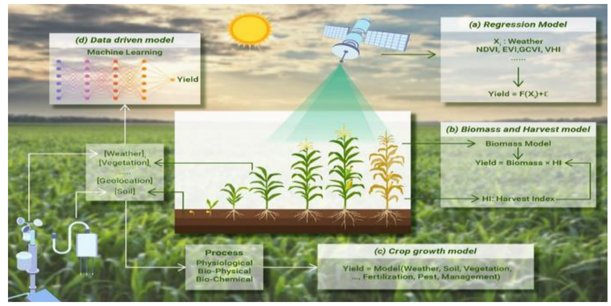

Fig. 1. IoT sensors on the ground and in farm equipment [31]

for construction plans, with drones simplifying the process and achieving impressive results [31-33]. Satellite remote sensing enhances monitoring efficiency for large, multi-scale applications [34, 35]. In recent development, satellite remote sensing has been successfully used for crop monitoring to forecast output, accurately describe location, weather, and temporal changes, and estimate yields per pixel [36-39]. Machine vision technology has gained importance in agricultural automation [40].

Figure 2 shows the use of remote sensing for monitoring and yield estimation. The plant material used is reddish-orange *Vitis vinifera L. cv.* Bhopal was vaccinated at 110 degrees Richter. The land was planted in 2002 with dimensions of 2.5 × 1.4 m (2857 vines ha-1) with ropes connected to the north-south vertical trusses. Remote sensing, big data, AI, and ML tools were adopted for sustainable agriculture [41]. AI advancements and Graphics Processing Unit (GPU) and Deep Belief Network (DBN) technologies have significantly improved plots with 1.2 drip nozzles [42]. Computer vision technology enhances resource efficiency in agricultural production through decision support [43, 44]. The approach challenges noise and distortion in underwater photographs by creating a

Fig. 2. Remote Sensing for Crop Monitoring and Yield Estimation [41]

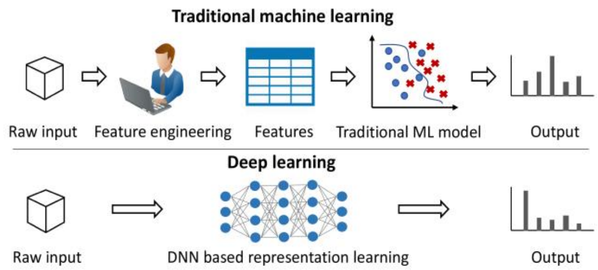

Fig. 3. Techniques for ML and DL model [46]

3D model using depth maps, overlapping tiles, and mosaic images [45]. Figure 3 categorizes ML models into feature engineering-based and end-to-end Deep Neural Network (DNN) pipelines, highlighting similarities in interpretation paradigms, focusing on understanding neural representations [46].

The typical objectives and contributions of the present study are as follows:

1) Estimation of winter wheat yields using data from multiple sources at both district and pixel levels in large areas by comparing multiple ML and DL methods, including RNN and Random Forest (RF) algorithms.
2) To explore factors such as soil, weather, and crops that are important in predicting yield.
3) Finally, we propose a scalable, simple, and cost-effective operating modeling approach for accurate and fast yield estimation.

The present study also reviews the literature on crop forecasting, analyzes planning strategies, and discusses experiments, results, and ongoing research. Table 1 depicts a summary of different studies on agricultural yield using various techniques, approaches, and models, utilizing diverse datasets. The research utilized Support Vector Machine (SVM), CNN, Long Short-term Memory Networks (LSTM), and other ML algorithms. The performance and accuracy of each algorithm is being varied.

2. MATERIALS AND METHODS

The proposed method uses data from the Kaggle database to estimate crop yield predictions [56]. Precipitation, temperature, air pressure, vapor pressure, and the frequency of rainy days are all examples of climatic parameters. The information in this document is geographically organized by latitude and county. Random Forest (RF) and Artificial Neural Network (ANN) algorithms are powerful ML tools for crop yield analysis, combining ensemble learning and AI networks to make decisive decisions and recognize complex relationships between inputs and outputs. The random forest-based crop mapping framework utilizes various data sources and remote sensing data to enhance crop classification accuracy and efficiency. This method aids in land use planning, precision agriculture, environmental monitoring, advancing agriculture, and remote sensing. RF and ANN are chosen for crop yield analysis due to their efficiency in handling large data sets, ability to learn complex relationships, and easy training, making them suitable for time-consuming tasks. ML approaches like Bayes and Decision Trees are not suitable for this task due to their probabilistic nature and limited handling of large datasets [57]. Decision Trees are unsupervised learning algorithms for classification and regression tasks, but they lack complex relationship learning capabilities. The use of existing datasets and various ML and DL approaches for crop yield prediction at different measures in high-yield agricultural manufacturing locations requires increased limited attention [58].

DL, which uses neural networks to learn features directly from the data, is the basis of present work. The DL approach is more flexible and enables to achieve better results for a series of tasks [59]. In terms of robustness, scalability and interpretation ability, the present work is better than the other because it is based on DL techniques, which are much more powerful and flexible than traditional methods of ML.

Figure 4 shows the framework for the present study using DL approach. Modules for feature extraction, Decision Support System (DSS), and data preprocessing are all included in the framework. Performance measurement and forecasting are also included in the DSS module. With ML and DL, predictions can be made, and performance can be evaluated by looking at the DSS's accuracy.

2.1 Dataset

The dataset used in the present study is from Kaggle, hypothetical data is used which is a public data repository [56]. The MSMD feature selection method improves agricultural classification efficiency and accuracy by reducing redundancy

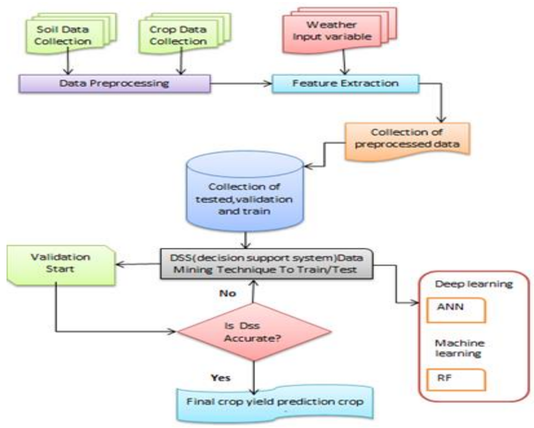

Fig. 4. The framework for the proposed methods

and focusing on optimal features. It enhances precision in farmland mapping using multi-source imagery, making it a significant addition to remote sensing and land cover categorization. The dataset contains information on crop yield, climate, and other factors.

Figure 5 depicts a visual representation of the dataset in the combined bi-temporal optical radar data for crop distribution in pictorial form recorded by Rapid Eye (optical) satellites and polarization radar data captured by UAVSAR (Unmanned Aerial Vehicle Synthetic Aperture Radars) in a rural area close to Winnipeg, Canada and are used in the present study. At harvest

Table 1. Prediction algorithm as applied in ML and DL

Ref No	Algorithm	Dataset Used	Results
[47]	Deep neural network	Each model was first trained with 900,000 data sets.	The 10-neuron, 5-layer Bayesian DNN model is the same as the original 400-neuron 10-layer DNN model, although the number of neural networks is reduced by about 80.
[48]	Hybrid machine learning, (ANN-ICA), (ANN-GWO)	Wheat, barley, potato, sugar beet	ANN-GWO showed better prediction results than the ANN-ICA model with R = 0.48, RMSE = 3.19, and MEA = 26.65.
[49]	Convolutional neural network with 1-D Convolutional operation	Meteorological, soil, and plant phenology data from 271 German districts over 21 years (1999–2019).	RMSE 7-14% lower, MAE 3-15% lower, and correlation coefficient 4-50% higher than the best-performing reference factor on all test data.
[50]	(ML): XGBoost Algorithm, (CNN), Deep (DNN), CNNXGBoost, (RNN), and CNN (LSTM).	The soybean dataset contains 25345 samples and 395 factors, such as climate and soil parameters.	CNN and DNN hybrid models have rmse 0.276, mse 0.071, mae 0.199, and R2 0.87; The XGBoost models are better than other models.
[51]	BPA with FFNN and ANN	regional soil parameter	Regional soil factors may have a critical role in enriching the CYP.
[52]	Deep Recurrent Q-Network model	Vellore district in the southern	Accuracy of 93.7%.
	Wheat crop simulation model (CSM), remote sensing (RS)	The use of Sentinel 2A and Landsat 8 imagery and in-person LAI measurements is used for verification	NE increases by 2%, 5%, 3%, and 1% more on simulated days until flowering.
[53]	Imagery satellite, (NDVI), (SAVI).	East Java with various spatial and remote sensing datasets	NDVI (R2 = 77.81%) and SAVI (R2 = 72.8%).
[54]	Backpropagation Neural Networks (BPNN) and Genetic Algorithms (GA)	When combining the three main tobacco growth metrics (plant density, nitrogen fertilization, and leaf count), prepare to plant goals (yield or QC), weather, and soil information.	77.66 kg/mu of smoke is produced, and the CQ is 81.02. The main objective is to smoke QC with a wait of 80.
[55]	Information and communication technologies (ICTs), DSSPIM	Southern Spain's greenhouses are home to orange and tomato trees.	Orange plants demonstrate how 20 % less water is used when implementing a water management method for tree crops.

Fig. 5. The visual illustration of the sample images used in this study

time, seven crops were grown in the region: corn, peas, canola, soybeans, oats, wheat, and hardwoods.

Predict crop yields for five Gulf-grown crops: potatoes, melons, dates, wheat, and maize (corn). A prediction model was developed using five independent variables: year, rainfall, pesticide, temperature fluctuations, and nitrogen fertilizer. Crop prediction is crucial for decision-making in agriculture and uses input variables to determine food availability for the upcoming years.

2.2 Data preprocessing and feature extraction

In order to address the various issues, which come due to incompleteness, inconsistency and missing of values against various features of the dataset, a data preprocessing technique known as normalization is introduced. After the data is preprocessed, it can generate promising results from simulations. Following feature extraction and data pre-processing, there are 12 features in the dataset, including derived features. These include longitude, latitude, altitude, and day length, quantity of precipitation, minitemp, maxitemp, ndvi, wind speed, mean temperature, standardized temperature, and yield are among the functions.

2.3 Model Selection

It is intended to develop an intelligent DSS for crop yield monitoring using RF and ANN.

2.3.1 Random Forest (RF)

We have used the RF Classifier from scikit-learn to simulate RF-based engines. With a few exceptions indicated below, the default set of settings is utilized initially:

- 'n estimators'- (n shows the trees that makeup the forest, default size is 10);
- 'Max depth' - The maximum depth of the tree (default: none). If the setting is 'None,' the documentation indicates that "tree vertices are expanded until all the Childs are pure or until all child node contain less than min samples split";
- 'Min samples split' – This parameter shows the required least number of samples to separate an internal node in the tree (it is set to 2 by default);
- 'Min samples leaf' - The very bare least number of nodes, correspondingly);
- One seven-node output layer.

Since entities are standardized real numbers, `ReLU was chosen as the activation function of optimal hidden layers. Also, since this is a multiclass classification exercise, where the output is intended to be binary ('1' for the specified class, '0' for all other classes), choosing 'softmax' makes the layer output trigger function seem appropriate. This is a multi-class classification exercise, and "categorical_crossentropy" is selected as the loss function. Adaptive performance is evaluated using "accuracy" as a metric for selection. For samples in freshly formed leaves, the default is one.

2.3.2 Artificial Neural Network (ANN)

The ANN design has a sequential structure that includes:

- 1-Input layer (102 input nodes);
- 3-Hidden layers (204, 204 and 102 nodes, respectively);
- one seven-node output layer

Figure 6 represents the adaptive performance being assessed using "accuracy" as a selection criteria. Because entities are standardized real numbers, the activation function 'relu' was used for buried layers. Also, because, this is a meticulous classification exercise with binary output ('1' for the chosen class, '0' for all other classes), choosing 'softmax' makes the layer output trigger function seem appropriate. This is a multi-class classification task with "categorical_crossentropy", as the loss function.

3. RESULTS AND DISCUSSION

In this section, various evaluation matrices used for the proposed technique are discussed.

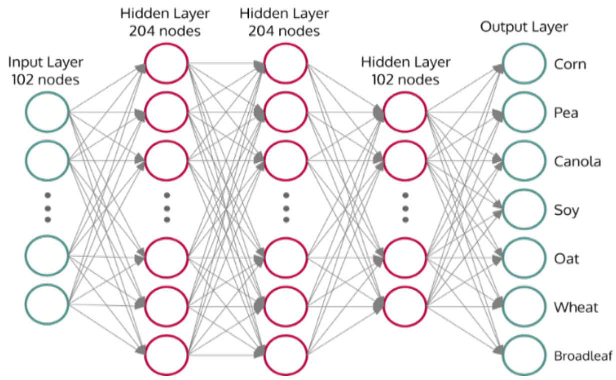

Fig. 6. The proposed Artificial Neural Network Model

3.1 Evaluation Matrices

The performance here is measured with both the evaluation and confusion matrices for both models (Random Forest and Neural Network) of the confusion matrix.

- Columns represent expected classes;
- whereas rows represent actual classes

3.1.1 Confusion Matrix

A confusion matrix is a 2x2 matrix structure that is useful for visualizing an algorithm's performance. The true positive rate (TPR) is defined as the entire number of positives that have been given that classification.

$$\mathrm{TPR} = \frac{TP}{TP+FN} \tag{1}$$

Where TP denotes the true positive and FN shows false negative.
The true negative rate is the proportion of conditions that qualify as negative.

$$\mathrm{TNR} = \frac{TP}{TN+FN} \tag{2}$$

The false positive rate is the proportion of instances that are misclassified or anticipated as being negative.

$$\mathrm{FPR} = \frac{FP}{FP+TN} \tag{3}$$

The false negative rate is the proportion of positive cases reported or anticipated as negative.

$$\mathrm{FNR} = \frac{FN}{FN+TP} \tag{4}$$

3.1.2 Accuracy

It measures the proportion of accurate predictions to all calculations.

$$\mathrm{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \tag{5}$$

3.1.3 Precision

It is the ratio between TPs combined with a number of TPs and FPs.

$$\mathrm{Precision} = \frac{TP}{TP+FP} \tag{6}$$

3.1.4 Recall

It is defined as the product of the ratio of TPs and the sum of the TP and FN numbers.

$$\mathrm{Recall} = \frac{TP}{TP+FN} \tag{7}$$

3.1.5 F1-score

Recall and accuracy are averaged mathematically, and it takes into consideration both false positive and false negative (FN) outcomes.

$$\mathrm{F1-score} = 2 * \frac{(precision*recall)}{(precision+recall)} \tag{8}$$

Figure 7 depicts the data visually, revealing a substantial variance in agricultural yields across different locations, with a focus on autumn-sown winter crops. What sticks out is that winter crops have a substantially higher amount of variance from year to year, indicating that their yields fluctuate more pronouncedly than other crop types.

Figure 8 focuses on training an algorithm utilizing multispectral satellite photos containing

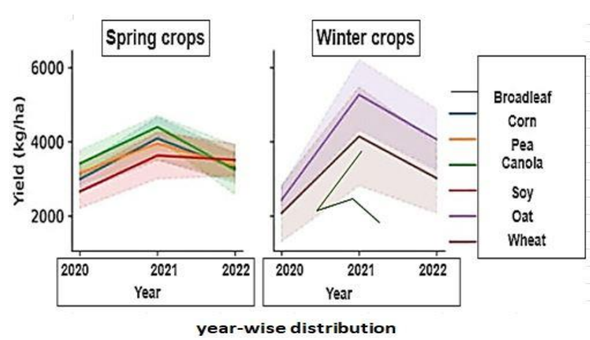

Fig. 7. Crop production statistics

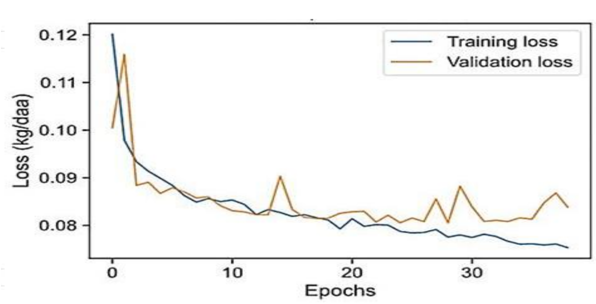

Fig. 8. Training and validation loss for the multi-temporal ANN model

information on crop kinds and their respective areas. Only these photos with crop type and area encoding were used to train the algorithm in this scenario. The training and validation losses of the multi-temporal Artificial Neural Network (ANN) model are evaluated. The best results from the single-image ANN trials served as the foundation for the training process, and subsequent photos with a pixel mask inserted as a separate channel were employed in this context. To test how well these models work, both statistics matrices and confusion matrices are used in this evaluation.

It is important to mention how crucial good data preparation is. The dataset used in present studies was originally found hampered by high feature intercorrelation, but it eventually proved to be fairly robust and representative. The neural network outperformed the random forest by a little margin. The accuracy scores for each crop variety were usually comparable, all of which were greater than 99 %. The "broadleaf" class was an outlier, having much lower accuracy values. This is to be expected that this class (with the fewest observations) is the most erroneously represented. Deep learning proved clearly superior at forecasting such harvests when focusing on the "broadleaf" class, indicating that it might be a more effective option in dealing with misrepresented classes in general.

According to Figures 9 and 10, the lowest loss was achieved at 77.53 kg/1000m^2. This figure demonstrates a 5.3% improvement over individual multi-temporal ANN findings and a 6.6% improvement over the crop classes Random Forest (RF) model. These graphs are created by mapping the distribution patterns of specific features and investigating their correlations with the dependent variable, given as "score." This study is made possible by using a custom function named "training example." This function, in particular, allows us to acquire insight into how these attributes are related with the "score." This analysis is performed on the seven attributes that have the strongest relationships with the dependent variable "score" to provide an initial comprehension of the data's behavior.

Figures 11 and 12 provide a comparison that focuses on evaluating the performance of two distinct models in predicting the same item. This evaluation entails adding farm-scale yields into predictions and then comparing these predictions to anticipated and actual crop production at the commune scale. The findings of this analysis confirm the presence of biases in numerous factors. However, it is vital to highlight that no changes will be made in advance to address these biases. The analysis is carried out with the assumption that these biases exist and will be considered throughout

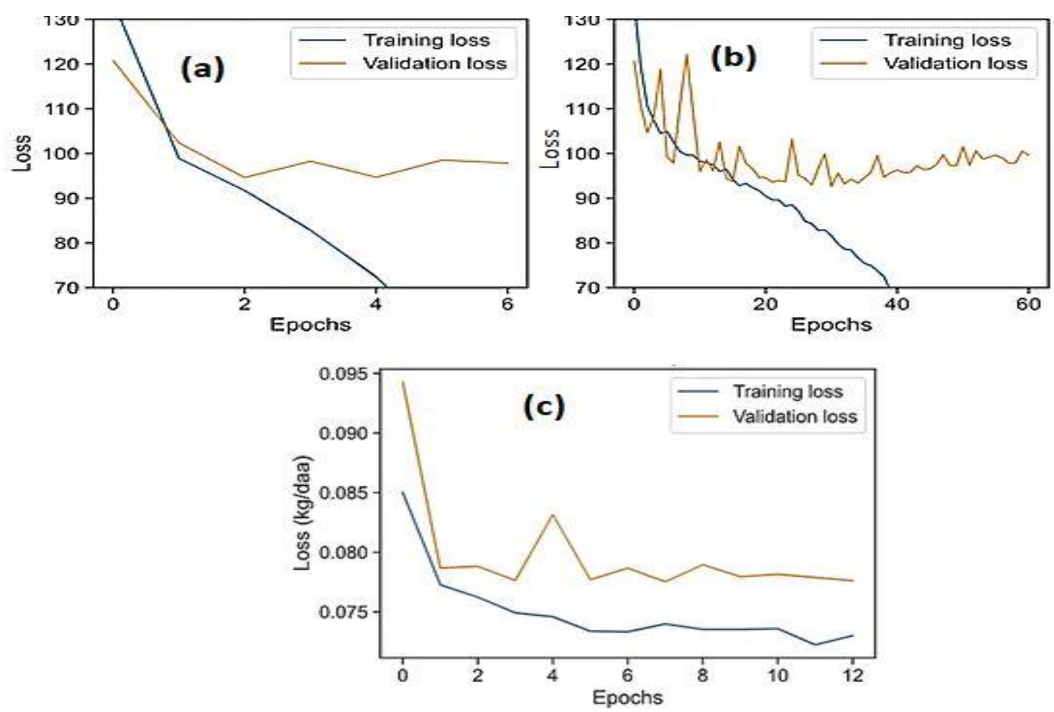

Fig. 9. Training and validation loss for the Illustration of stochastic epoch sampling (a) Simulated normal epochs, (b) With stochastic epoch sampling, and (c) Pre-trained ANN and RF.

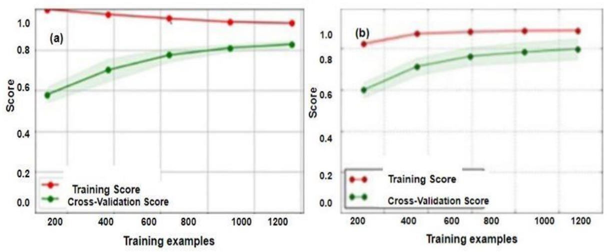

Fig. 10. Learning Curve of (a) ANN vs (b) RF

the review. It's worth noting that the sample size for each set of data points is 64 for each batch of data used. In every epoch, model has to interpretation to more than 5,000 diverse bunches. Few epochs may suffice to lead to top accuracy, minimum loss levels already in the start of the model training. Depending on how well trained classifiers performed overall, this choice might be reconsidered later. To compare deep neural networks and traditional random forests machine learning approaches, two separate sets of tensors will be produced and used in each learning experiment. After highly correlated characteristics were eliminated, this dataset now comprises 325,834 observations, which include one column for labels (integers ranging from 1 to 7); 102 columns for features. Following that, the unique features tensor and the two label tensors are divided into training and testing sets. The training set will have 80% of the observations, with the testing set holding the remaining 20%.

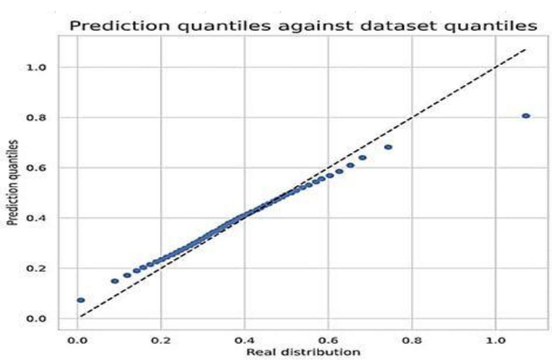

Fig. 11. ANN prediction quantiles versus real quantiles.

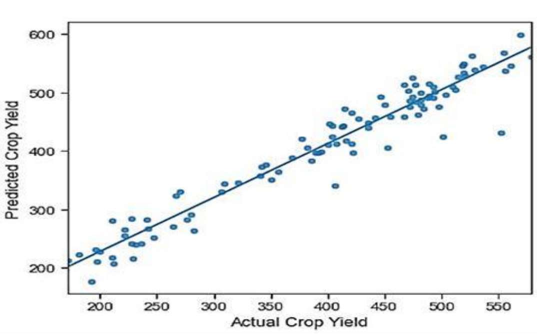

Fig.12. Comparison between actual and predicted crop yields on a commune-scale.

Table 2 represents that the grouping of optical and radar-based information produces extremely precise distant cropland mapping. It is important to note, however, that selecting the appropriate number of trees and depth parameters may have a significant influence on the results. Experiments with fewer trees and lower depth topologies revealed some deterioration, as predicted/expected. When utilized correctly, random forests are good predictors, with performance equivalent to more sophisticated, complicated algorithms. From the simulation results, it is crystal clear that NN-based model outperformed the RF-model to some extent.

Figure 13 and 14 show that the performance of the random forest classification will be preserved, with original labels (integers ranging from 1 to 7) being accommodated into one unidimensional array. The label column for neural network classification will be encoded one-hot using the "Pandas' get_dummies method" (/kaggle/input/cropland-mapping). As a result, labels will now be made up of seven binary parts, each of which refers to a different crop class, allowing for final class identification based on the array member with the greatest anticipated value.

Table 2. Comparison results

RF	Percentage	ANN	Percentage
Accuracy	99.64	Accuracy	99.82
Precision	99.29	Precision	99.79
Recall	99.29	Recall	99.38
F-Score	99.29	F-Score	99.58

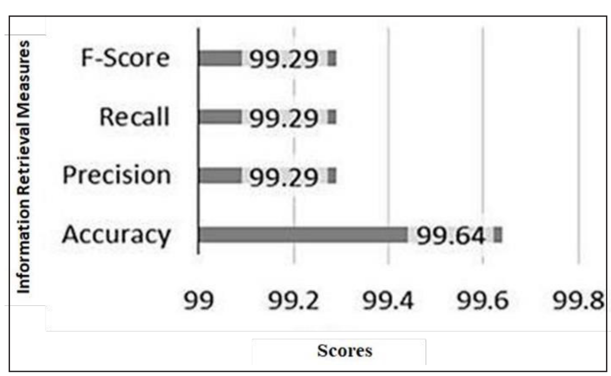

Fig. 13. RF and ANN results

Table 3: Results comparison of current and related study

Study	Year	Technique	Dataset	Precision	Recall	F1 Score	Accuracy
Zhang, et al., [60]	2022	Deep neural networks	Winter wheat (Landsat 8 imagery, Sentinel-2 imagery)	82	74	73	76
Akbar, et al., [61]	2022	CNN AlexNet, ResNet-50, VGG16	Wheat	86	77.1	76.9	85
Proposed Methodology	2023	ML + DL	Wheat	99.79	99.38	99.58	99.82

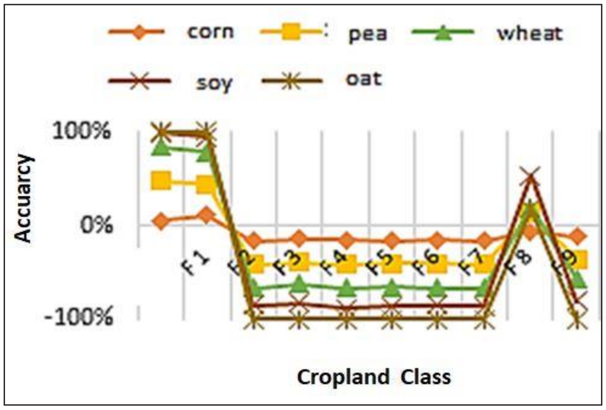

Fig. 14. The performance indicators of RF and ANN for cropland mapping

We have made comparisons with existing related studies to show that our results outperform in terms of accuracy using the different technique and the same type eof data. The empiricla evaluation of the proposed model with the existing studies, as depicted in Table 3, showed that the suggested model have achived more accuracy then the model in comparison. The suggested approach achieved the accuracy of 99.82% and super-passed all the other approaches.

4. CONCLUSIONS

AI is a technology that is emerging in the field of agriculture. It can give farmers real-time or overtime insights into their field. This allows farmers to identify areas that require irrigation, fertilization, or pesticide treatment. AI businesses are creating agricultural robots that can effortlessly do a variety of duties. These robots are programmed to harvest crops and kill plants more quickly than people. These can predict crop yield monitoring using precision farming technique that use data sensors, connected devices, remote control devices, and other technologies to allow farmers to control their fields. It is concluded from the present studies that the RF and ANN models based decision support system can be potentially used to generate cropland mapping for crop yield prediction. It is also revealed that the ANN model outperformed in crop yield prediction as compared to other models. Future research is required to look into hybrid machine learning algorithms like random forest, support vector machine, multiple regressor, logistic regressor, and deep learning algorithms like Deep convolution neural network (DCNN) and LSTM to see whether they can give rather quicker and more accurate solutions in the domain of precise agriculture. It is suggested that the DCNN and LSTM models be used in pre-foliar disease prediction to estimate crop yield, taking into account the latest large-scale data from several nations to predict fruit quality, etc. Farmers and agricultural experts may test the results.

5. ACKNOWLEDGEMENTS

The authors would like to thank Dr. Tahir Iqbal, Chairman Depatment of Farm Machinery and Precision Engineering, PMAS Arid Agriculture University for helping us in collecting and analysing the dataset and all the experimentations.

6. CONFLICT OF INTEREST

The authors declare no conflict of interest.

7. REFERENCES

1. K.S.M. Anbananthen, S. Subbiah, D. Chelliah, P. Sivakumar, V. Somasundaram, K.H. Velshankar, and M.A. Khan. An intelligent decision support system for crop yield prediction using hybrid machine learning algorithms. *F1000Research* 10: 1143 (2021).
2. K. Jhajharia, P. Mathur, S. Jain, and S. Nijhawan. Crop yield prediction using machine learning and deep learning techniques. *Procedia Computer Science* 218: 406-417 (2023).

3. S. Boonprong, C. Cao, W. Chen, X. Ni, M. Xu, and B.K. Acharya. The classification of noise-afflicted remotely sensed data using three machine-learning techniques: effect of different levels and types of noise on accuracy. *ISPRS International Journal of Geo-Information* 7: 1-21 (2018).
4. B. Devika, and B. Ananthi. Analysis of crop yield prediction using data mining technique to predict annual yield of major crops. *International Research Journal of Engineering and Technology* 5: 1460-1465 (2018).
5. S. Iniyan, V.A. Varma, and C.T. Naidu. Crop yield prediction using machine learning techniques. *Advances in Engineering Software* 175: 103326 (2023).
6. Anshul, and R. Singh. Crop Yield Prediction and Climate Change Impact Assessment Using Machine Learning Technology in Agriculture. In *Information and Communication Technology for Competitive Strategies (ICTCS 2022) Intelligent Strategies for ICT* 349-362 (2023).
7. D. Paudel, A. de Wit, H. Boogaard, D. Marcos, S. Osinga, and I.N. Athanasiadis. Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture* 206: 107663 (2023).
8. A. Nigam, S. Garg, A. Agrawal, and P. Agrawal. Crop yield prediction using machine learning algorithms. In *2019 Fifth International Conference on Image Information Processing (ICIIP),* 125-130 (2019).
9. F. Abbas, H. Afzaal, A.A. Farooque, and S. Tang. Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy* 10: 1046 (2020).
10. V. Pandith, H. Kour, S. Singh, J. Manhas, and V. Sharma. Performance evaluation of machine learning techniques for mustard crop yield prediction from soil analysis. *Journal of scientific research* 64: 394-398 (2020).
11. A. McBratney, B. Whelan, T. Ancev, and J. Bouma. Future directions of precision agriculture. *Precision agriculture* 6: 7-23 (2005).
12. E. Manjula, and S. Djodiltachoumy. A model for prediction of crop yield. *International Journal of Computational Intelligence and Informatics* 6: 298-305 (2017).
13. S. Bhanumathi, M. Vineeth, and N. Rohit. Crop yield prediction and efficient use of fertilizers. In *2019 International Conference on Communication and Signal Processing (ICCSP)*: 0769-0773 (2019).
14. C. Bojtor, S.M.N. Mousavi, A. Illés, A. Széles, J. Nagy, and C. L. Marton. Stability and adaptability of maize hybrids for precision crop production in a long-term field experiment in Hungary. *Agronomy* 11: 2167 (2021).
15. L. Meng, H.L. Liu, S. Ustin, and X. Zhang. Predicting maize yield at the plot scale of different fertilizer systems by multi-source data and machine learning methods. *Remote Sensing* 13: 3760 (2021).
16. U. Hampicke. Germany's Agriculture and UN's Sustainable Development Goal 15. *SUSTAINABLE LIFE ON LAND*, 18: 103(2021).
17. S. Khaki, L. Wang, and S. V. Archontoulis. A cnn-rnn framework for crop yield prediction. *Frontiers in Plant Science* 10: 1750 (2020).
18. S. Toomula, and S. Pelluri. An Extensive Survey of Deep learning-based Crop Yield Prediction Models for Precision Agriculture. In *Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 1* (pp. 1-12). Singapore: Springer Nature Singapore (2022).
19. L.R. Varghese, and K. Vanitha. Deep Reinforcement Learning and Model Predictive Control in Hybrid Deep Learning for Rubber Yield Forecast. *Revue d'Intelligence Artificielle* 35: 367-374 (2021).
20. E. Kalaiarasi, and A. Anbarasi. Multi-parametric multiple kernel deep neural network for crop yield prediction. *Materials Today: Proceedings* 62: 4635-4642 (2022).
21. S. Wolfert, D. Goense, and C.A.G. Sørensen. A future internet collaboration platform for safe and healthy food from farm to fork. *2014 annual SRII global conference* 266-273 (2014).
22. S. Khaki, and L. Wang. Crop yield prediction using deep neural networks. *Frontiers in plant science* 10: 621 (2019).
23. F.F. Haque, A. Abdelgawad, V.P. Yanambaka, and K. Yelamarthi. Crop Yield Prediction Using Deep Neural Network. 2020 *IEEE 6th World Forum on Internet of Things (WF-IoT)* 1–4 (2020).
24. M. Qiao, X. He, X. Cheng, P. Li, H. Luo, L. Zhang, and Z. Tian. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. *International Journal of Applied Earth Observation and Geoinformation* 102: 102436 (2021).
25. L. Gong, M. Yu, S. Jiang, V. Cutsuridis, and S. Pearson. Deep learning based prediction on greenhouse crop yield combined TCN and RNN. *Sensors* 21: 4537 (2021).
26. J. You, X. Li, M. Low, D. Lobell, and S. Ermon. Deep gaussian process for crop yield prediction based on remote sensing data. In *Proceedings of the AAAI conference on artificial intelligence* (2017).
27. K. Gavahi, P. Abbaszadeh, and H. Moradkhani. DeepYield: A combined convolutional neural network with long short-term memory for crop yield forecasting. *Expert Systems with Applications* 184: 115511 (2021).
28. M. Samaila, G. Neto, M. Fernandes, D.A. Freire, M.M, and P.R. Inácio. Challenges of securing Internet of Things devices: A survey. *Security and Privacy* 1: e20 (2018).
29. E.S. Mohamed, A.A. Belal, S.K. Abd-Elmabod, M.A. El-Shirbeny, A. Gad, and M.B. Zahran. Smart farming for improving agricultural management. *The Egyptian Journal of Remote Sensing and Space Science* 24: 971-981 (2021).
30. M.S. Farooq, S. Riaz, A. Abid, T. Umer, and Y.B. Zikria. Role of IoT technology in agriculture: A systematic literature review. *Electronics* 9: 319 (2020).
31. A. James, A. Saji, A. Nair, and D. Joseph.

CropSense–A Smart Agricultural System using IoT. *Journal of Electronic Design Engineering* 5**:** 3 (2019).

32. J.M.O. Jayamanne, R. Vandebona, and J.M.I. Karalliyadda. 2020. Optimization of Conventional Land Survey Techniques Using Modern Technology (2020).
33. K. Guan, J. Wu, J.S. Kimball, M.C. Anderson, S. Frolking, B. Li, C.R. Hain, and D.B. Lobell. The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote sensing of environment* 199: 333-349 (2017).
34. Q. Yang, L. Shi, J. Han, Y. Zha, and P. Zhu. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Research* 235: 142-153 (2019).
35. L. Wei, Y. Luo, L. Xu, Q. Zhang, Q. Cai, and M. Shen. Deep convolutional neural network for rice density prescription map at ripening stage using unmanned aerial vehicle-based remotely sensed images. *Remote Sensing* 14: 46 (2022).
36. Y. Shen, B. Mercatoris, Z. Cao, P. Kwan, L. Guo, H. Yao, and Q. Cheng. Improving Wheat Yield Prediction Accuracy Using LSTM-RF Framework Based on UAV Thermal Infrared and Multispectral Imagery. *Agriculture* 12: 892 (2022).
37. F. Wang, M. Yang, L. Ma, T. Zhang, W. Qin, W. Li, Y. Zhang, Z. Sun, Z. Wang, F. Li, and K. Yu. Estimation of above-ground biomass of winter wheat based on consumer-grade multi-spectral UAV. *Remote Sensing* 14: 1251 (2022).
38. H. Tian, T. Wang, Y. Liu, X. Qiao, and Y. Li. Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture* 7: 1-19 (2020).
39. T. Wang, B. Chen, Z. Zhang, H. Li, and M. Zhang. Applications of machine vision in agricultural robot navigation: A review. *Computers and Electronics in Agriculture* 198: 107085 (2022).
40. N. Ismail, and O.A. Malik. Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. *Information Processing in Agriculture* 9: 24-37 (2022).
41. H. Li, Z. Li, W. Dong, X. Cao, Z. Wen, R. Xiao, Y. Wei, H. Zeng, and X. Ma. An automatic approach for detecting seedlings per hill of machine-transplanted hybrid rice utilizing machine vision. *Computers and Electronics in Agriculture* 185: 106178 (2021).
42. R. Ballesteros, D.S. Intrigliolo, J.F. Ortega, J.M. Ramírez-Cuesta, I. Buesa, and M.A. Moreno. Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques. *Precision Agriculture* 21: 1242-1262 (2020).
43. K. Mochida, S. Koda, K. Inoue, T. Hirayama, S. Tanaka, R. Nishii, and F. Melgani. Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective. *GigaScience* 8: giy153 (2019).
44. M. Vázquez-Arellano, H.W. Griepentrog, D. Reiser, and D. S. Paraforos. 3-D imaging systems for agricultural applications—a review. *Sensors* 16: 618 (2016).
45. Y. Li, C.J. Randall, R. van Woesik, and E. Ribeiro. Underwater video mosaicing using topology and superpixel-based pairwise stitching. *Expert Systems with Applications* 119: 171-183 (2019).
46. L. Haalck, and B. Risse. Embedded dense camera trajectories in multi-video image mosaics by geodesic interpolation-based reintegration. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* 1849-1858 (2021).
47. E. Casoli, D. Ventura, G. Mancini, D.S. Pace, A. Belluscio, and G. Ardizzone. High spatial resolution photo mosaicking for the monitoring of coralligenous reefs. *Coral Reefs* 40: 1267-1280 (2021).
48. M. Du, N. Liu, and X. Hu. Techniques for interpretable machine learning. *Communications of the ACM* 6: 68-77 (2019).
49. B. Saravi, A.P. Nejadhashemi, P. Jha, and B. Tang. Reducing deep learning network structure through variable reduction methods in crop modeling. *Artificial Intelligence in Agriculture* 5: 196-207 (2021).
50. S. Nosratabadi, F. Imre, K. Szell, S. Ardabili, B. Beszedes, and A. Mosavi. Hybrid machine learning models for crop yield prediction. *arXiv preprint arXiv:2005.04155* (2020).
51. J. Arshad, M. Aziz, A.A. Al-Huqail, M. Husnain, A.U. Rehman, and M. Shafiq. Implementation of a LoRaWAN based smart agriculture decision support system for optimum crop yield. *Sustainability* 14: 827 (2022).
52. S.S. Dahikar, and S.V. Rode. Agricultural crop yield prediction using artificial neural network approach. *International journal of innovative research in electrical, electronics, instrumentation and control engineering* 2: 683-686 (2014).
53. F. Abbas, H. Afzaal, A.A. Farooque, and S. Tang. Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy* 10: 1046 (2020).
54. D. Elavarasan, and P.D. Vincent. Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. *IEEE access* 8: 86886-901 (2020).
55. R. Dhakar, V.K. Sehgal, D. Chakraborty, R.N. Sahoo, J. Mukherjee, A.V. Ines, S.N. Kumar, P.B. Shirsath, and S.B. Roy. Field scale spatial wheat yield forecasting system under limited field data availability by integrating crop simulation model with weather forecast and satellite remote sensing. *Agricultural Systems* 195: 103299 (2022).
56. I. Khosravi, and S.K. Alavipanah. A random forest-based framework for crop mapping using temporal, spectral, textural and polarimetric observations. *International Journal of Remote Sensing* 40: 7221-7251 (2019).
57. I.H. Sarker. Machine learning: Algorithms, real-world applications and research directions. *SN computer science* 2: 160 (2021).
58. L. Alzubaidi, J. Zhang, A.J. Humaidi, A. Al-Dujaili,

Y. Duan, O. Al-Shamma, J. Santamaría, M.A. Fadhel, M. Al-Amidie, and L. Farhan. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data* 8: 1-74 (2021).

59. H.F. Assous, H. AL-Najjar, N. Al-Rousan, and D. AL-Najjar. Developing a Sustainable Machine Learning Model to Predict Crop Yield in the Gulf Countries. *Sustainability* 12: 9392 (2023).
60. S. Zhang, H. Feng, S. Han, Z. Shi, H. Xu, Y. Liu, and J. Yue. Monitoring of Soybean Maturity Using UAV Remote Sensing and Deep Learning. *Agriculture* 13: 110 (2022).
61. S. Akbar, K.T. Ahmad, M.K. Abid, and N. Aslam. Wheat Disease Detection for Yield Management Using IoT and Deep Learning Techniques. *VFAST Transaction on Software Engineering* 10: 80-89 (2022).