International Journal of Scientific Research and Engineering Development-– Volume 7 Issue 1, Jan-Feb 2024

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Systematic Reviews of Machine Learning Applications in Corn: A Literature Review

Ronie A. Pansoy*, Erick J. Garcia**, Edwin R. Arboleda***

*(Department of Computer and Electronic Engineering,Cavite State University, Indang, Cavite, Philippines
Email:ronie.pansoy@cvsu.edu.ph)

**(Department of Computer and Electronic Engineering, Cavite State University, Indang, Cavite, Philippines
Email:erick.garcia@cvsu.edu.ph)

***(Department of Computer and Electronic Engineering, Cavite State University, Indang, Cavite, Philippines
Email:edwin.r.arboleda@cvsu.edu.ph)

--************************----------------------------------

Abstract:

In the realm of corn cultivation, Machine learning application revolutionizes conventional farming methods by imposing advanced algorithms to analyze diverse datasets. Machine learning can be applied in projecting crop yields, detecting diseases, and controlling weeds. This method eventually enhances efficiency and crop productivity. The goal of the literature review is to analyze systematically and incorporate the existing collection of research on machine learning applications in the context of corn cultivation. The review investigates the vast range of machine learning techniques employed, their impact, and their effectiveness on different aspects of corn agriculture. The review highlights the different models that are used in different applications of machine learning in corn cultivation and its impact in the economy and environment. This review gives valuable insights for researchers, practitioners, and policymakers aiming to integrate machine learning technology to optimize corn production by examining the present state of knowledge.

Keywords **—Disease Detection, Machine Learning, Weed control, Yield Prediction**

--************************----------------------------------

I. INTRODUCTION

A. *Background*

Corn or maize is considered the widely grown food crop worldwide. Maintaining a thriving maize industry plays a role, in ensuring global food security. Maize serves as a food, animal feed and raw material for industrial purposes [2]. The adoption of corn cultivation technology by farmers is influenced by factors like age, education level, farming experience and counseling frequency. Given its uses in industry, animal feed and food production corn hold importance in agriculture. Over time there have been advancements in maize production technology [3] These technological improvements have greatly enhanced farm productivity and efficiency. The utilization of techniques such as farming systems and production technologies has been found to increase maize yield while reducing labor and material costs [4]. These advancements do not boost productivity. Also contribute to cost reduction and environmental sustainability, in corn farming.

Machine learning is essential in modern agriculture because it offers numerous advantages. Precision farming also termed digital agriculture, and agri-technology are now emerging as new scientific disciplines that employ data-intensive methods to enhance agricultural output while reducing the impact on the environment [5]. Additionally, Precision farming has become an innovative tool to address the problem of the sustainability of agriculture today. The latest technology in this field is powered by machine learning (ML). The machine can learn without being explicitly programmed. Machine learning and other computing techniques are the latest developments for studying and solving different complex problems. To be able to build the models and assess the data, different analytical models have been used, including Random Forests, Support Vector Machines, Decision Trees, Artificial Neural Networks, Bayesian Networks, and so on. These techniques make it possible to examine the soil, climate, and water regime which have a huge impact on crop growth and precision farming [6].

The application of machine learning to corn production holds immense potential benefits, revolutionizing modern agriculture. It is stated by Khanal et al. [6], that in contrast to conventional techniques, machine learning algorithms offer a time- and money-efficient method for spatially predicting crop yield and parameters of the soil. There are also challenges in applying machine learning in the cultivation of corn. These include difficulties in determining who is responsible for what, a lack of transparency and explainability, problems with fairness, and worries about data ownership, privacy, and security [7].

In order for the machine learning adoption process in corn production to be effective and implementable, the review aims to identify and assess the important machine learning methods that are applied to different facets of corn farming. The evaluation of the wider effects of machine learning on key elements including yield prediction, disease detection, and weed control in the context of corn production ensues. In addition, the review seeks to critically assess the amount of literature already in existence, identifying the knowledge gaps and pointing out areas for further research into the dynamic connection of corn farming and machine learning. The review seeks to achieve these goals with a comprehensive overview of the state of the current knowledge in the subject, presenting insights into the effectiveness of machine learning applications, and directing prospects for developments in this area.

II. METHODOLOGY

A. *Selection of the Study*

The systematic literature reviews reporting the use of ML applications in corn, Searches were conducted in Scopus, Springer link, Google Scholar, IEEE, Nature, MDPI, PubMed, ResearchGate and Taylor and Francis. The following words were searched in the titles and abstracts of published studies: "ML" “machine learning” and “corn”. The research involved a systematic search and selection process to identify relevant journal papers published within the 2018-2023 timeframe. Following a framework similarto Liakos et al. [5], the initial search filtered for papers published in English within the specified period and focused on journal articles. Excluded were book chapters, reviews, dissertations, and non-English studies.

The full texts and abstracts of the discovered publications were then examined in order to evaluate their eligibility according to two main standards: they had to be published between 2018 and 2023 and had to make reference to one of the predefined categories. The co-authors worked together to guarantee that these criteria were applied consistently during the selection procedure.

The review technique is summarized in a flowchart shown in Fig. 1., which follows the PRISMA principles for transparency[8]. 38 identified journal articles that were judged relevant to the study goals were eventually obtained using this process.

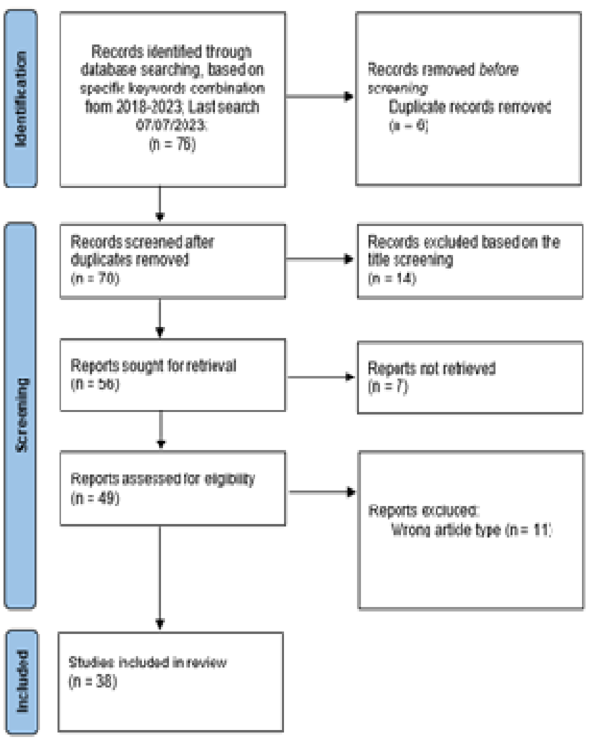

Fig. 1The flowchart of the methodology of the present systematic review following the PRISMA guidelines for exclusion criteria.

B. *Data Analysis*

The systematic literature reviews (SLRs) were examined using certain criteria pertaining to machine learning applications in corn. Basic statistics from every publication reported in each of the included SLRs were extracted individually, including the sources of data, accuracy, specificity, and sensitivity.

These performance metrics, which are defined as follows, are frequently used to evaluate how well machine learning models perform:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$sensitivity = \frac{TP}{TP + FN}$$

$$specificity = \frac{TN}{TN + FP}$$

Where: FN = false negative; FP = false positive; TN = true negative; TP = true positive

Accuracy represents the overall correctness of the model prediction; sensitivity is the fraction of correctly identified positive cases while specificity is the fraction of correctly identified negative cases.

Methods of validation and handling missing data were also extracted. Details regarding the external validation with respect to the comparison of machine learning models against traditional methods were also extracted and reviewed, not only from systematic literature reviews but also from primary studies. According to Shahhosseini et al. [9] the details of the types of machine learning techniques were also extracted, and the number of primary studies reporting the use of different machine learning algorithm typologies was determined for each SLR included.

III. MACHINE LEARNING APPLICATIONS IN CORN FARMING

A. *Yield Prediction*

Crop yield prediction is a method of calculating the yield of the crop in kilograms per hectare while taking account of several variables such as location, weather, soil characteristics, water level, and yield from the previous year [10]. The researchers used an advanced cropping system simulator available as open-source software is called Agricultural Production Systems Simulator (APSIM). Several modules, including maize, SWIM soil water, soil N, carbon, surface residue, soil temperature, and management criteria based on radiation and water use efficiency concepts, were combined to simulate biomass production using APSIM. It is suggested that increasing ML yield prediction might be accomplished by adding more soil water-related variables (either from remote sensing, simulation models, or other sources).

A study on maize yield prediction is presented with use of six algorithms and environmental variables from satellite observations, weather data, and crop progress reports found that advanced algorithms and a large composite of information are more effective. In regards to accuracy and stability, the XGBoost algorithm performs well compared to other algorithms; deep neural networks, such as CNN and LSTM, are not advantageous. Prediction is not considerably influenced by the time series variables' compositing interval. Prediction accuracy is increased by 5% when the best algorithms and inputs are combined [11].

In the paper of Nyeki et al. [12], Spatial-temporal data mining has been applied to create a new approach for maize production prediction. XY-fused Query networks, supervised Kohonen networks, neural networks with Rectangular Linear Activations, extreme gradient boosting (XGBoost), support-vector machines (SVM), and independent variables in five vegetation periods were within the models that were used. On training and test sets, the most effective technique (XGBoost) had accuracy rates of 92.1% and 95.3%, respectively. An innovative approach to treating individual units in a lattice system was presented, leading to a 97.5% increase in the Area under the curve (AUC). Extreme Gradient Boosting Trees, with 92.1% accuracy on the training set, was the top regression model. The technique also determines how site-specific soil fertility influences maize grain production.

Ma et al. [13] developed a Bayesian Neural Network (BNN)--based county-level corn yield forecast model that has been developed with publicly accessible data sources. The model surpassed six other methods over ten testing years, accurately predicting corn yield in both typical and anomalous years. Around mid-August, two months before harvest, the model performed almost at its best. Extreme heat, water stress, and crop masks raised suggested uncertainty, while sequential features decreased it.

The paper of Aghighi et al. [14] utilized advanced machine learning techniques, the study predicted silage maize yields utilizing boosted regression tree (BRT), random forest regression (RFR), support vector regression, and Gaussian process regression (GPR). Based on the data, RFR was the most stable method in 2015, while BRT outperformed the other methods every year, suggesting their capacity to estimate maize yields less sensitive to NDVI inconsistency.

Awasthi [15] proposed a framework to forecast corn yield which is based on machine learning methods. To help farmers predict the annual production of corn, the researchers incorporated data on soil, climate, weather, and crop yield. The researchers employ the ensemble bagging extreme gradient boosting (XGBoost) model, decision tree (DT), random forest (RF), and linear regression (LR) models. After comparing and analyzing every model, The researchers conclude that the Bagging XGBoost Regression model performs better than every other model, with an accuracy of 93.8% and an RMSE of 9.1.

In the paper of Kralj et al. [16], the researchers developed a linear machine learning approach that is more accurate and offers early corn yield forecast with a relative error of less than 20%. This is essential data for decisions about harvesting and storing resources. The algorithm is based on a Generalized Linear Model (GLM). Numerous climate and greenhouse gas characteristics are input parameters for our model. In the R1 (Silking) phase of the corn crop and beyond, the researchers investigated the prediction accuracy of corn yield for farmers who desired to know what the yield would be at a specific harvest date. In this test scenario, the relative error of the model was 11.63%, whereas the GLM's was 12.55%.

The study of Kayad et al. [16] assessed the spatial variability of corn grain yield by applying machine learning algorithms and vegetation indexes from Sentinel-2 pictures. Between 2016 and 2018, a 22-ha field was monitored and more than 20,000 yield observation sites were documented. For tracking within-field variability, the Green Normalized Difference Vegetation Index (GNDVI) provided the greatest R2 value of 0.48; the optimal time frame was determined to be between 105 and 135 days after planting. The machine learning technique that predicted yield variability most effectively was Random Forests.

In the research of Croci et al. [17], the researcher applied phenology, predictors, and machine learning algorithms to predict maize yield. The best was Gaussian process regression, which performed at its best late in the season. The best approaches with vegetation indices and during the tasseling phenological stage were neural networks and support vector machines using the linear basis function. The maize yield standard deviation of the performance prediction decreased and the performance of NNET improved with the use of principal component analysis (PCA).

Meng et al. [18] attempted the integration of data from numerous sources, such as soil, fertilizer, monthly climate data, satellite data (i.e., vegetation indices, or VIs), and fertilizer data, to investigate the accuracy of various inputs in yield prediction. Based on the results, yield prediction can be enhanced by combining all of the datasets and employing AB (adaptive boosting) and random forests (RF) (R2: 0.85~0.98). Furthermore, relative to other combinations (such as combinations of all data and combinations of VIs and soil data), the combination of VIs, climate data, and soil data (VCS) can predict maize yield more accurately.

B. *Disease Detection and Management*

One of the most essential food crops in the world is corn, and disease will significantly lower yields. Therefore, an essential aspect of maize production is the identification and management of corn disease [19]. The paper of Panigrahi et al. [20] focuses on supervised machine learning techniques for diagnosing maize plant diseases using plant images, including Random Forest (RF), K-Nearest Neighbor (KNN), Decision Tree (DT), Naive Bayes (NB), and Support Vector Machine (SVM). The aforementioned classification methods are investigated and compared to determine which model is most appropriate and have the highest accuracy in identifying plant diseases. Out of all the classification techniques, the RF algorithm yields the highest accuracy, at 79.23%. As a preventive precaution, the farmers will use all of the previously mentioned trained models for the early detection and classification of new image diseases.

Rajeena P. P. et al. [21] employ methods for acquiring, preprocessing, and classifying images. Preprocessing comprises steps like data augmentation, image scaling, and image reading. By changing the variables, the proposed project, which is based on EfficientNet, improves the accuracy of the corn leaf disease database. To confirm the accuracy and resilience of this method, tests are conducted on the test dataset using DenseNet and Resnet. Experimental results show that this method can obtain a recognition accuracy of 98.85%, which is much higher than other advanced procedures.

In the study of Divyanth et al. [22], Using a custom dataset of handheld photos of maize leaves afflicted with Northern Leaf Spot, Northern Leaf Blight, and Gray Leaf Spot diseases, a new two-stage semantic segmentation method was developed. The approach located, identified, and computed the area coverage for disease lesions by employing the SegNet, UNet, and DeepLabV3+ network architectures to extract leaves from diverse field backdrops. Stage one performance was highest for the UNet model, whereas mwIoU of 0.7379 and mBFScore of 0.5351 were detected as disease lesions by the DeepLabV3+ model. The development of a field-worthy disease management system was made possible by the integrated model, which accurately forecasted the severity of three diseases based on actual observations.

In the work of Haque et al. [23] a deep learning approach has been presented for identifying diseased photos of maize crops in the field. Three diseases were identified from images obtained from experimental fields: banded leaf and sheath blight, turcicum leaf blight, and maydis leaf blight. Using methods such as brightness augmentation and rotation, artificial images were produced. Using a baseline training approach, three architectures built on the 'Inception-v3' network were trained. With an average recall of 95.96% and an overall classification accuracy of 95.99%, the top-performing model was attained. The model demonstrated how the baseline training method may be applied to enhance learning and feature extraction.

The paper of Chauhan & Al [24] forecasted the early identification of agricultural disease using artificial intelligence (AI) techniques such as Random Forest (RF), K-Nearest Neighbor (KNN), Decision Tree (DT), Naive Bayes (NB), and Support Vector Machine (SVM). To determine the most effective model for identifying diseases, the researchers evaluated all available methods based on accuracy in this publication. In this instance, the Random Forest model outperforms other current models with an accuracy of 80.68%.

The objective of the paper of Padilla et al. [25] is to determine the disease by looking at the leaf in the corn. The paper explores the utilization of OpenMP and Convolutional Neural Network for corn leaf disease detection. The accuracy percentage of the Convolutional Neural Network classifier in identifying and classifying diseases was 93%, 89%, and 89% for Leaf Blight, Leaf Rust, and Leaf Spot, respectively. In leaf diseases, a high percentage of categorization was made possible by the use of OpenMP.

In the paper of Amin et al. [26], Deep learning model utilizing pre-trained convolutional neural networks (CNNs) EfficientNetB0 and DenseNet121 has been developed to differentiate between healthy and unhealthy maize plant leaves. The model enhances the diversity and quantity of images by using data augmentation techniques together with concatenation techniques to extract deep characteristics from photos. The model performs better compared to two other pre-trained CNN models with greater parameters and processing power requirements: ResNet152 and InceptionV3. With a 98.56% classification accuracy, the model outperforms ResNet152 and InceptionV3, which had respective results of 98.37% and 96.26%.

The study of Agarwal et al. [27] demonstrates the early leaf analysis technique for identifying diseases in maize crops. For analysis and testing, the researchers have used the PlantVillage dataset. The accuracy, precision, recall, storage space, model running time, and AUC-RoC are just a few of the performance indicators that have been used to assess the validity of the findings. The acquired results illustrate how effectively the suggested approach performs when compared to conventional machine learning techniques. 94% accuracy can be attained using the developed model.

Convolutional neural networks (CNN) and boosting methods have been utilized by Bhatt et al. [28] to detect pests and crop diseases in photographs of corn leaves. Along with wheat and rice, corn is one of the main food crops and is also very flexible. The objective of the research was to classify associated disease symptoms more accurately than current deep learning techniques. Employing a classifier and boosting, the researchers examined ensembles of CNN-based image features and were able to classify corn leaf images into four categories: healthy, common rust, late blight, and leaf spot. This resulted in an accuracy of 98%, which was 8% better than CNN alone.

Xu et al. [29] demonstrated an improved ResNet50-based model for detecting corn pests. The objective was to recognize diseases and pests that affect maize exactly and effectively. the researchers introduced more effective channels (environment–cognition–action) to the residual network module by using convolution and pooling processes to extract shallow-edge features and compress data. In addition to addressing the issue of network degradation, this stage creates links between channels and makes it easier to extract essential deep information. Eventually, the ResNet50 model was used to achieve 96.02% recognition accuracy through experimental validation. Numerous maize pests and diseases, such as stem borer, rust disease, gray leaf spot, maize leaf blight, and Helminthosporium maydis, have been successfully detected by this investigation. These findings provide insightful information for the intelligent management and control of diseases and pests affecting maize.

In the research of Dai et al. [30], MTDL-EPDCLD, a deep learning-based system, is proposed for efficient corn leaf disease diagnosis and detection. Two tasks constitute a component of the system: fine-grained disease classification with attention (FDCA) and rapid health status identification (RAHSI). For Task 1, a shallow CNN-4 model with a spatial attention mechanism is utilized and the accuracy is 98.73%. For Task 2, a modified version of the MobileNetV3Large-Attention model is created, yielding 94.44% accuracy and gains of 4-8% in precision, recall, and F1 score. Better crop yields, enhanced food security, and well-informed decision-making are all supported by the system.

C. *Weed Control and Management*

Researchers and farmers are becoming interested in the topic of site-specific weed management in precision agriculture [31]. In the work of Picon et al. [32], a pixel-by-pixel classification model for plant species recognition using deep learning is presented. Grass, broadleaf, and crop species were the subject of three datasets created. Model convergence is aided by the inclusion of an auxiliary classification loss and a semantic segmentation architecture. Other datasets are added to the network without requiring more manual annotation labor, leading to improved network performance. By utilizing single-species or synthetic datasets, the algorithm's performance can be doubled, and the suggested solution outperforms the state of the art.

Mota-Delfin et al. [33] created An aerial RGB image of a corn crop under weedy conditions that was utilized in a study to compare deep learning systems. There were a total of ten flight missions carried out, six of which used a ground sampling distance of 0.33 cm/pixel and four of which employed a distance of 1.00 cm/pixel. At the intersection over union thresholds of 0.25, 0.50, and 0.75, YOLOv4 detectors were compared. In all models, there was a 4.92% gain for 0.25 compared to 0.50, and a significant F1-Score penalty at 0.75. When the confidence level was higher than 0.35, YOLOv4 exhibited increased robustness in detection. When it came to plant count, YOLOv5-s obtained a mAP of 73.1%, a coefficient of determination of 0.78, and a relative mean square error of 42%.

In the study of Pathak et al. [34], the researchers proposed to classify four common weeds in a cornfield using computer vision methods. Images were collected from the soil and 21 shape features were extracted. The handcrafted simple image processing approach was successful in distinguishing lambs quarters and redroot pigweed from horseweed. However, advanced non-parametric machine learning models like k-nearest neighbor, random forest, and support vector machine (SVM) showed high accuracies, with RF outperforming others. The study recommends using the simple handcrafted simple image processing algorithm for weed identification and classification before implementing advanced ML models.

Ni et al. [35] an effective model for weed recognition in corn fields employing deep convolutional neural networks (VGG16). For common weeds, including black grass, cleavers, Charlock, common chickweed, and loose silky-bent, the model demonstrated enhanced recognition effects, with F1-values of 0.971, 0.945, 0.949, 0.959, and 0.958. The SGD optimizer with the best overall performance was identified by the researchers upon having retrained these weed datasets using various optimizers. This approach can boost corn production while saving labor expenses.

Yang et al. [36] proposed a new corn weed identification model which is SE-VGG16. Employing VGG16 as a foundation, the SE-VGG16 model integrates the SE attention mechanism to highlight relevant elements. For the extraction of features and dimensionality reduction, it reduces convolutional kernels to 1×1 and uses Leaky ReLU in place of ReLU activation. The model outperforms the original VGG16 model with an average accuracy of 99.67% in identifying maize weeds, outperforming it by a margin. The robustness, stability, and recognition rate of the model render it a viable solution for practical applications, particularly within the context of weed control in corn fields.

IV. IMPACT ASSESSMENT

A. *Economic Impact*

The integration of machine learning into the corn farming industry has the potential to have an impact on its aspects. Machine learning models have been used to predict maize crop yields in locations, including Eswatini in Africa [37]. These models, trained with opensource data have achieved levels of precision as indicated by modified R2 values ranging from 0.784 to 0.978 [38]. In the United States a new approach called Bayesian Domain Adversarial Neural Network (BDANN) was proposed for domain adaptation in predicting corn yields. Has proven to be more effective than other methods. Similarly in Ohio combining machine learning algorithms with sensed data improved the accuracy of forecasts for maize production and soil parameters compared to linear regression [6]. Another example can be seen in Africa where an explicit random forest (RF) algorithm was developed to estimate variations in maize grain yields over time and space demonstrating the potential of machine learning and remotely sensed data for mapping soil properties and corn yield [39].

In their study Khanal et al. [40] delve into the application of machine learning techniques, on resolution sensed data to predict soil properties and corn yield in a corn field. They show that these models can accurately estimate soil and crop variables providing maps of these factors across the field. The paper emphasizes that this approach is beneficial for farmers as it allows them to identify areas, with productivity, nutrient deficiency or water stress. By implementing site management practices based on these findings farmers can enhance crop performance and profitability. Additionally this method saves time and costs associated with soil and crop sampling, analysis and mapping while improving farm management decisions efficiency and effectiveness [41].

In the study conducted by Qin et al. [42] they delve into the realm of machine learning techniques to predict the nitrogen rate (EONR), for corn production. This factor plays a role in ensuring sustainable corn farming practices. The researchers evaluate four machine learning algorithms, random forest (RF) ridge regression (RR) support vector machine (SVM) and artificial neural network (ANN). They assess how well these algorithms can forecast the EONR for corn cultivation across nitrogen management scenarios. Additionally, the paper explores how incorporating soil features, from a mechanistic model, impacts the accuracy of EONR prediction. It highlights the effectiveness of the RR algorithm in predicting split application EONR. This finding is significant as it offers corn farmers the opportunity to adjust their nitrogen application rates based on crop growth stages and environmental conditions thereby improving sustainability. Additionally incorporating soil hydrological features obtained from a model has been shown to enhance the accuracy of EONR prediction. This improvement can further contribute to sustainability by optimizing nitrogen management according to soil moisture and temperature. The paper suggests that leveraging machine learning can serve as a tool, in achieving sustainability, in corn production through more precise and adaptable nitrogen management practices [42].

The use of machine learning techniques in corn farming has been shown to be cost effective, by reducing costs while simultaneously increasing productivity.It also improves the return on investment and economic sustainability resulting in increased profits and term environmental advantages [43].

B. *Environmental Impact*

The creation of models for sustainable digital agriculture depends heavily on machine learning (ML).ML techniques play a role in promoting sustainability in agriculture particularly when it comes to crop and water management. It is essential to maintain a balance between climatological factors for soil and water management [44]. By leveraging machine learning in agriculture we can encourage practices while minimizing the negative impact on our surroundings. In the realm of agriculture wireless sensor networks (WSN) have proven valuable, in monitoring soil properties and gathering data, which greatly enhances decision making processes [45]. Precision agriculture aims to optimize crop yield while minimizing waste relying on the power of machine learning and data analysis [46]. By utilizing machine learning algorithms for disease detection, soil analysis and crop production prediction we can effectively manage resources and improve food security [47]. Farmers can now conveniently. Control their crops from a distance by utilizing machine learning and Internet of Things (IoT) devices. This advanced technology aids in minimizing resource usage and enhancing the management of operations [39]. Precision farming can be made accessible by leveraging the capabilities of machine learning and Internet of Things (IoT) sensors. By utilizing machine learning, in agriculture we can achieve outcomes, reduce wastage and promote sustainable intensification. Sustainable digital agriculture benefits significantly from machine learning, which enhances soil analysis, disease detection, crop and water management, and crop production prediction. It additionally improves food security and helps with resource management issues. Including wireless sensor networks enhances the ability to make decisions. Remote crop control became possible by IoT device implementation, which lowers consumption of resources and improves operational management. With the application of this technology, sustainable intensification, waste reduction, and precision farming are encouraged.

V. CHALLENGES AND FUTURE DIRECTIONS

A. *Challenges in Implementing Machine Learning in Corn Production*

The researchers discovered four papers on Challenges in Implementing Machine Learning in Corn Farming. According to Ahmad et al, the need for a large number of disease samples to train machine learning algorithms, which is difficult to collect evenly and meticulously during production [48]. Another challenge is haploid seed selection, which is now done by trained technicians but can be automated using machine learning approaches [49]. Training credible supervised machine learning models for corn yield prediction. Necessitates representative ground truth labels, which may be limited or unavailable due to cost and labor constraints [50]. Existing deep learning models for corn yield forecasting. Frequently lack the ability to quantify prediction uncertainty and are prone to overfitting with limited training sets [13]. These difficulties highlight the need for improved data collection methods, seed selection automation, and the creation of algorithms that can manage minimal labeled data and generate uncertainty estimates.

B. *Opportunities for Future Research*

When it comes to identifying research gaps some areas require attention. Firstly, it is important to develop cost scalable methods for obtaining data from various sources, including soil samples, weather data, and satellite imagery. Additionally exploring techniques for integrating and merging datasets will be essential for extracting meaningful insights.

Secondly, designing machine learning models that provide explanations for their predictions is of importance. This will enable farmers to trust the outputs and better understand them. Moreover, there should be methods to quantify prediction uncertainty so that farmers can make well informed decisions even when faced with risks and uncertainties. Lastly addressing the challenges related to data scarcity and bias is also crucial in order to ensure results, in corn farming practices.

The researchers should explore strategies for effectively training machine learning models when we have limited labeled data, which is often the case in resource limited environments. Additionally, it is important to develop approaches that address biases in training datasets and algorithms to ensure equal outcomes for farmers, across the board.

When it comes to suggestions for advancing the use of machine learning (ML) in agriculture, two important areas to focus on are personalized farming and precision agriculture. By developing ML powered systems that provide recommendations tailored to farms and field conditions we can optimize resource allocation and maximize crop yield. Integrating real-time sensor data with ML models enables decision making allowing for adjustments in irrigation, fertilization, and pest control based on conditions.

Another crucial aspect is maintenance and disease management. ML can play a role in the detection and prevention of corn diseases, reducing crop losses and promoting practices. Additionally implementing ML based maintenance systems for farm machinery helps optimize maintenance schedules and minimize downtime.

Intelligent farming systems that utilize ML for farm management including yield prediction, resource optimization, and market analysis should also be developed. Furthermore, fostering collaboration among researchers, farmers, and industry stakeholders is essential to ensure the implementation and adoption of ML solutions in corn farming. Open-source data sharing initiatives along with programs and financial support can accelerate the development and deployment of these technologies.

By acknowledging and addressing the research gaps that have been identified as prioritizing the areas of advancement that have been recommended that can tap into the vast potential of machine learning (ML) to revolutionize corn farming. This transformation will make it more efficient, sustainable, and profitable ultimately securing food security, for generations.

VI. CONCLUSION

This review emphasizes how machine learning has influenced corn farming and how effective it is at managing diseases, yield prediction, and sustainability. Notably, the accuracy of image classifiers in weed and disease diagnosis and the precision of extreme gradient boosting (XGBoost) in agricultural production forecasts are highlighted.

Corn farming gains substantial advantages from the integration of machine learning, which provides accurate yield projections, creative domain adaptation methods, and improved forecasting accuracy through algorithmic synergy with sensory data. Machine learning minimizes expenses associated with conventional methods by providing actionable information for more effective farm management through accurate monitoring of soil parameters and corn yields.

The impact of machine learning extends beyond the prediction of nitrogen rates, enabling real-time modifications and the promotion of sustainable practices. Despite its potential, challenges such as a shortage of data, automated seed selection, and a variety of illness samples still exist. In order to overcome these challenges while enhancing operational effectiveness, the article emphasizes the importance of personalized farming and precision agriculture.

REFERENCES

[1] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, “Machine Learning Applications for Precision Agriculture: A Comprehensive Review,” *IEEE Access*, vol. 9, pp. 4843–4873, 2021, doi: 10.1109/ACCESS.2020.3048415.

[2] I. Zulfida, “FARMERS’ ADOPTION LEVEL OF CORN CULTIVATION TECHNOLOGY (Zea Mays L) IN THE SUB-DISTRICT OF LABUHANBATU DISTRICT ON THE INFLUENCE OF SOCIO-ECONOMIC ASPECTS,” *ZIRAAAH Maj. Ilm. Pertan.*, vol. 48, no. 1, Art. no. 1, Jan. 2023, doi: 10.31602/zmip.v48i1.9192.

[3] R. Kumar, A. Bhardwaj, L. P. Singh, and G. Singh, “Environmental Impact Assessment of Maize Production in Northern India,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1084, no. 1, p. 012042, Oct. 2022, doi: 10.1088/1755-1315/1084/1/012042.

[4] G. Onuwa, “Fostering Sustainable Productivity Through Maize Technology Intensification: Participant Responses,” in *Advances in Environmental Engineering and Green Technologies*, R. S. Sengar, R. Chaudhary, and H. S. Bhadauriya, Eds., IGI Global, 2022, pp. 427–436. doi: 10.4018/978-1-7998-8434-7.ch027.

[5] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine Learning in Agriculture: A Review,” *Sensors*, vol. 18, no. 8, Art. no. 8, Aug. 2018, doi: 10.3390/s18082674.

[6] S. Khanal, J. Fulton, A. Klopfenstein, N. Douridas, and S. Shearer, “Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield,” *Comput. Electron. Agric.*, vol. 153, pp. 213–225, Oct. 2018, doi: 10.1016/j.compag.2018.07.016.

[7] M. Gardezi *et al.*, “Artificial intelligence in farming: Challenges and opportunities for building trust,” *Agron. J.*, vol. n/a, no. n/a, Apr. 2023, doi: 10.1002/agj2.21353.

[8] "PRISMA." Accessed: Dec. 28, 2023. [Online]. Available: http://prisma-statement.org/PRISMAStatement/FlowDiagram?AspxAutoDetectCookieSupport=1

[9] M. Shahhosseini, R. A. Martinez-Feria, G. Hu, and S. V. Archontoulis, "Maize yield and nitrate loss prediction with machine learning algorithms," *Environ. Res. Lett.*, vol. 14, no. 12, p. 124026, Dec. 2019, doi: 10.1088/1748-9326/ab5268.

[10] S. A. Shetty, T. Padmashree, B. M. Sagar, and N. K. Cauvery, "Performance Analysis on Machine Learning Algorithms with Deep Learning Model for Crop Yield Prediction," in *Data Intelligence and Cognitive Informatics*, I. Jeena Jacob, S. Kolandapalayam Shanmugam, S. Piramuthu, and P. Falkowski-Gilski, Eds., in Algorithms for Intelligent Systems. Singapore: Springer, 2021, pp. 739–750. doi: 10.1007/978-981-15-8530-2_58.

[11] Y. Kang, M. Ozdogan, X. Zhu, Z. Ye, C. Hain, and M. Anderson, "Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest," *Environ. Res. Lett.*, vol. 15, no. 6, p. 064005, May 2020, doi: 10.1088/1748-9326/ab7df9.

[12] A. Nyéki *et al.*, "Application of spatio-temporal data in site-specific maize yield prediction with machine learning methods," *Precis. Agric.*, vol. 22, no. 5, pp. 1397–1415, Oct. 2021, doi: 10.1007/s11119-021-09833-8.

[13] Y. Ma, Z. Zhang, Y. Kang, and M. Özdoğan, "Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach," *Remote Sens. Environ.*, vol. 259, p. 112408, Jun. 2021, doi: 10.1016/j.rse.2021.112408.

[14] H. Aghighi, M. Azadbakht, D. Ashourloo, H. S. Shahrabi, and S. Radiom, "Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 12, pp. 4563–4577, Dec. 2018, doi: 10.1109/JSTARS.2018.2823361.

[15] P. Awasthi, S. Mishra, and N. Gupta, "Performance Assessment of Machine Learning Techniques for Corn Yield Prediction," in *Advanced Network Technologies and Intelligent Computing*, I. Woungang, S. K. Dhurandher, K. K. Pattanaik, A. Verma, and P. Verma, Eds., in Communications in Computer and Information Science. Cham: Springer Nature Switzerland, 2023, pp. 320–335. doi: 10.1007/978-3-031-28183-9_23.

[16] A. Kayad, M. Sozzi, S. Gatto, F. Marinello, and F. Pirotti, "Monitoring Within-Field Variability of Corn Yield using Sentinel-2 and Machine Learning Techniques," *Remote Sens.*, vol. 11, no. 23, Art. no. 23, Jan. 2019, doi: 10.3390/rs11232873.

[17] M. Croci, G. Impollonia, M. Meroni, and S. Amaducci, "Dynamic Maize Yield Predictions Using Machine Learning on Multi-Source Data," *Remote Sens.*, vol. 15, no. 1, Art. no. 1, Jan. 2023, doi: 10.3390/rs15010100.

[18] L. Meng, H. Liu, S. L. Ustin, and X. Zhang, "Predicting Maize Yield at the Plot Scale of Different Fertilizer Systems by Multi-Source Data and Machine Learning Methods," *Remote Sens.*, vol. 13, no. 18, Art. no. 18, Jan. 2021, doi: 10.3390/rs13183760.

[19] C. Guifen, Z. Shan, C. Liying, F. Siwei, and Z. Jiaxin, "Corn plant disease recognition based on migration learning and convolutional neural network," *Smart Agric.*, vol. 1, no. 2, p. 34, Apr. 2019, doi: 10.12133/j.smartag.2019.1.2.201812-SA007.

[20] K. P. Panigrahi, H. Das, A. K. Sahoo, and S. C. Moharana, "Maize Leaf Disease Detection and Classification Using Machine Learning Algorithms," in *Progress in Computing, Analytics and Networking*, H. Das, P. K. Pattnaik, S. S. Rautaray, and K.-C. Li, Eds., in Advances in Intelligent Systems and Computing. Singapore: Springer, 2020, pp. 659–669. doi: 10.1007/978-981-15-2414-1_66.

[21] F. Rajeena P. P., A. S. U., M. A. Moustafa, and M. A. S. Ali, "Detecting Plant Disease in Corn Leaf Using EfficientNet Architecture—An Analytical Approach," *Electronics*, vol. 12, no. 8, Art. no. 8, Jan. 2023, doi: 10.3390/electronics12081938.

[22] L. G. Divyanth, A. Ahmad, and D. Saraswat, "A two-stage deep-learning based segmentation model for crop disease quantification based on corn field imagery," *Smart Agric. Technol.*, vol. 3, p. 100108, Feb. 2023, doi: 10.1016/j.atech.2022.100108.

[23] M. A. Haque *et al.*, "Deep learning-based approach for identification of diseases of maize crop," *Sci. Rep.*, vol. 12, no. 1, Art. no. 1, Apr. 2022, doi: 10.1038/s41598-022-10140-z.

[24] M. D. Chauhan and E. Al, "Detection of Maize Disease Using Random Forest Classification Algorithm," *Turk. J. Comput. Math. Educ. TURCOMAT*, vol. 12, no. 9, Art. no. 9, Apr. 2021.

[25] D. A. Padilla, R. A. I. Pajes, and J. T. De Guzman, "Detection of Corn Leaf Diseases Using Convolutional Neural Network With OpenMP Implementation," in *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, Dec. 2020, pp. 1–6. doi: 10.1109/HNICEM51456.2020.9400004.

[26] H. Amin, A. Darwish, A. E. Hassanien, and M. Soliman, "End-to-End Deep Learning Model for Corn Leaf Disease Classification," *IEEE Access*, vol. 10, pp. 31103–31115, 2022, doi: 10.1109/ACCESS.2022.3159678.

[27] M. Agarwal, V. K. Bohat, Mohd. D. Ansari, A. Sinha, S. Kr. Gupta, and D. Garg, "A Convolution Neural Network based approach to detect the disease in Corn Crop," in *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, Dec. 2019, pp. 176–181. doi: 10.1109/IACC48062.2019.8971602.

[28] P. Bhatt, S. Sarangi, A. Shivhare, D. Singh, and S. Pappula, "Identification of Diseases in Corn Leaves using Convolutional Neural Networks and Boosting," presented at the 8th International Conference on Pattern Recognition Applications and Methods, Dec. 2023, pp. 894–899. Accessed: Dec. 22, 2023. [Online]. Available: https://www.scitepress.org/Link.aspx?doi=10.5220/0007687608940899

[29] W. Xu, W. Li, L. Wang, and M. F. Pompelli, "Enhancing Corn Pest and Disease Recognition through Deep Learning: A Comprehensive Analysis," *Agronomy*, vol. 13, no. 9, Art. no. 9, Sep. 2023, doi: 10.3390/agronomy13092242.

[30] D. Dai, P. Xia, Z. Zhu, and H. Che, "MTDL-EPDCLD: A Multi-Task Deep-Learning-Based System for Enhanced Precision Detection and Diagnosis of Corn Leaf Diseases," *Plants*, vol. 12, no. 13, Art. no. 13, Jan. 2023, doi: 10.3390/plants12132433.

[31] G. c Sunil, Y. Zhang, C. Koparan, M. R. Ahmed, K. Howatt, and X. Sun, "Weed and crop species classification using computer vision and deep learning technologies in greenhouse conditions," *J. Agric. Food Res.*, vol. 9, p. 100325, Sep. 2022, doi: 10.1016/j.jafr.2022.100325.

[32] A. Picon, M. G. San-Emeterio, A. Bereciartua-Perez, C. Klukas, T. Eggers, and R. Navarra-Mestre, "Deep learning-based segmentation of multiple species of weeds and corn crop using synthetic and real image datasets," *Comput. Electron. Agric.*, vol. 194, p. 106719, Mar. 2022, doi: 10.1016/j.compag.2022.106719.

[33] C. Mota-Delfin, G. de J. López-Canteñs, I. L. López-Cruz, E. Romantchik-Kriuchkova, and J. C. Olguín-Rojas, "Detection and Counting of Corn Plants in the Presence of Weeds with Convolutional Neural Networks," *Remote Sens.*, vol. 14, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/rs14194892.

[34] H. Pathak, C. Igathinathane, K. Howatt, and Z. Zhang, "Machine learning and handcrafted image processing methods for classifying common weeds in corn field," *Smart Agric. Technol.*, vol. 5, p. 100249, Oct. 2023, doi: 10.1016/j.atech.2023.100249.

[35] C. Ni, B. Tian, X. Wang, Y. Sun, and C. Fei, "A deep convolutional neural network-based method for identifying weed seedlings in maize fields," in *2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Dec. 2022, pp. 776–779. doi: 10.1109/IMCEC55388.2022.10019943.

[36] L. Yang, S. Xu, X. Yu, H. Long, H. Zhang, and Y. Zhu, "A new model based on improved VGG16 for corn weed identification," *Front. Plant Sci.*, vol. 14, 2023, Accessed: Dec. 22, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpls.2023.1205151

[37] S. Fashoto, E. Mbunge, G. Ogunleye, and J. Van Den Burg, "IMPLEMENTATION OF MACHINE LEARNING FOR PREDICTING MAIZE CROP YIELDS USING MULTIPLE LINEAR REGRESSION AND BACKWARD ELIMINATION," *Malays. J. Comput.*, vol. 6, no. 1, p. 679, Jan. 2021, doi: 10.24191/mjoc.v6i1.8822.

[38] Y. Ma and Z. Zhang, "A Bayesian Domain Adversarial Neural Network for Corn Yield Prediction," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3211444.

[39] Mayur Rajaram Salokhe, "Machine Learning: Applications in Agriculture (Crop Yield Prediction, Diease and Pest Detection)," *Int. J. Adv.*

Res. Sci. Commun. Technol., pp. 592–597, Jul. 2023, doi: 10.48175/IJARSCT-12088.

[40] S. Khanal *et al.*, "Assessing the impact of agricultural field traffic on corn grain yield using remote sensing and machine learning," *Soil Tillage Res.*, vol. 208, p. 104880, Apr. 2021, doi: 10.1016/j.still.2020.104880.

[41] S. Khanal, J. Fulton, A. Klopfenstein, N. Douridas, and S. Shearer, "Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield," *Comput. Electron. Agric.*, vol. 153, pp. 213–225, Oct. 2018, doi: 10.1016/j.compag.2018.07.016.

[42] Z. Qin *et al.*, "Application of Machine Learning Methodologies for Predicting Corn Economic Optimal Nitrogen Rate," *Dep. Agron. Hortic. Fac. Publ.*, Jan. 2018, [Online]. Available: https://digitalcommons.unl.edu/agronomyfacpub/1288

[43] V. Meshram, K. Patil, V. Meshram, D. Hanchate, and S. D. Ramkteke, "Machine learning in agriculture domain: A state-of-art survey," *Artif. Intell. Life Sci.*, vol. 1, p. 100010, Dec. 2021, doi: 10.1016/j.ailsci.2021.100010.

[44] A. Bhansali, S. Saxena, and K. C. Bandhu, "Chapter 8 Machine learning for sustainable agriculture," in *Machine Learning for Sustainable Development*, K. Kant Hiran, D. Khazanchi, A. Kumar Vyas, and S. Padmanaban, Eds., De Gruyter, 2021, pp. 129–146. doi: 10.1515/9783110702514-008.

[45] C. Salim and N. Mitton, "Machine Learning Based Data Reduction in WSN for Smart Agriculture," in *Advanced Information Networking and Applications*, L. Barolli, F. Amato, F. Moscato, T. Enokido, and M. Takizawa, Eds., in Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2020, pp. 127–138. doi: 10.1007/978-3-030-44041-1_12.

[46] A. Ingle, Priyanka Gupta, Nachiket Gaikwad, and Sanika Bhatye, "Precision Agriculture Application using Machine Learning," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 1–8, Apr. 2023, doi: 10.48175/IJARSCT-9463.

[47] M. Kagzi, "Role of Machine Learning in Promoting Sustainability," in *Machine Learning for Business Analytics*, Productivity Press, 2022.

[48] X. Ming, S. A. Ahmad, and S. Ibrahim, "Identification and Analysis of Maize Leaf Diseases and Insect Pests Based on Machine Learning," in *Proceedings of the 9th International Conference on Computational Science and Technology*, D.-K. Kang, R. Alfred, Z. I. B. A. Ismail, A. Baharum, and V. Thiruchelvam, Eds., in Lecture Notes in Electrical Engineering. Singapore: Springer Nature, 2023, pp. 327–333. doi: 10.1007/978-981-19-8406-8_24.

[49] S. FiDan and A. M. Tiryaki, "Hyperparameter Optimization in Convolutional Neural Networks for Maize Seed Classification," *Eur. J. Res. Dev.*, vol. 3, no. 1, pp. 139–149, Mar. 2023, doi: 10.56038/ejrnd.v3i1.254.

[50] Y. Ma and Z. Zhang, "A Bayesian Domain Adversarial Neural Network for Corn Yield Prediction," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3211444.

