ZAMBIA INFORMATION COMMUNICATION TECHNOLOGY (ICT) JOURNAL

Volume 7 (Issue 2) (2023) Pages 35-42

The Use of Machine Learning in Industry 4.0 as an Educational Tool to Address Plant Diseases for Small-Scale Farmers in Developing Countries: A Review Study

Barbara C. Kunda
Computer Science Department.
university of Zambia
Lusaka, Zambia

Jackson Phiri
Computer science Department
University of Zambia
Lusaka, Zambia

***Abstract*— This review study aims to provide an overview of the current state of research on the use of machine learning techniques for the detection of tomato leaf diseases in the context of climate change in Zambia. Plant diseases pose significant challenges to small-scale farmers in developing countries, impacting crop yields and livelihoods. The emergence of Industry 4.0 technologies, coupled with the power of machine learning, offers promising opportunities to address these challenges and empower farmers with valuable knowledge and tools This review study aims to provide an overview of the use of machine learning in Industry 4.0 as an educational tool specifically tailored to tackle plant diseases for small-scale farmers in developing countries. The study offers insights into the potential of these techniques to enhance disease detection and contribute to sustainable agricultural practices in the face of climate change. Climate change has had significant impacts on agricultural practices worldwide, leading to the emergence and spread of various plant diseases. The study examines existing literature, research articles, and practical implementations to analyze the potential applications of machine learning in plant disease management. The review focuses on four key areas: disease identification, early detection and prediction, knowledge sharing and education, and decision support systems. With further advancements in machine learning techniques and the integration of cutting-edge technologies, the agriculture sector can benefit from improved disease detection and mitigation strategies to ensure food security in the face of climate change.**

Keywords: Machine learning algorithms, climate change, tomato leaf diseases, datasets, crop yield, food security, industry 4.0.

I. INTRODUCTION

Plant diseases have a significant impact on agricultural productivity, causing substantial crop losses and threatening food security. Small-scale farmers, who often lack the resources and infrastructure of large-scale agricultural operations, are particularly vulnerable to these diseases. Traditional methods of disease identification and management are often limited, relying on visual inspection or local knowledge, which can lead to misdiagnosis and ineffective control measures [28].

Now more than ever, there is a consensus that the world's climate system has changed due to human activities. These changes are occurring in both natural ecosystems and human wellbeing, with the poor and those already grappling with food insecurity expected to be hardest hit [33]. In its 2018 special report, the Intergovernmental Panel on Climate Change (IPCC) warns that an increase in temperature of 1.5 °C above the pre-industrial levels will result in significant risks to food security, livelihoods, and economic development. A warmer climate is expected to adversely affect rural populations who rely on agricultural production for their livelihoods [34]. This calls for more and not less mitigation action. The emergence of Industry 4.0 technologies, combined with the power of machine learning, offers promising opportunities to address these challenges and empower small-scale farmers with valuable educational tools for managing plant diseases.

The advent of Industry 4.0, characterized by the integration of digital technologies and automation, presents a transformative approach to agricultural practices. Machine learning, a subset of artificial intelligence, has shown remarkable potential in various fields and can be harnessed as an educational tool to tackle plant diseases. By leveraging the capabilities of machine learning algorithms, Industry 4.0 can provide small-scale farmers in developing countries with accessible and effective solutions for disease identification, early detection, and management [35].

This review study aims to explore the use of machine learning in Industry 4.0 as an educational tool specifically tailored to address plant diseases for small-scale farmers in developing countries. By examining existing literature, research articles, and practical implementations, we seek to provide an overview of the potential applications of machine learning in plant disease management. We will focus on four key areas: disease identification, early detection and prediction, knowledge sharing and education, and decision support systems.

The findings of this review study will contribute to the understanding of the current state of research and practical implementations in utilizing machine learning as an educational tool in Industry 4.0 for small-scale farmers. Furthermore, it will highlight the potential benefits, challenges, and opportunities associated with adopting such technologies in the context of developing countries. Ultimately, the aim is to facilitate informed decision-making, enhance disease management practices, and improve crop productivity, thereby contributing to the well-being and livelihoods of small-scale farmers and promoting sustainable agricultural practices in these regions.

and rainfall, it becomes possible to identify the relationships between climate patterns and disease occurrence. This information can guide the development of climate-adaptive disease management.

II. RELATED STUDIES

The use of machine learning in agriculture is essential to help improve the efficient use of resources for agricultural cultivation and harvesting, as well as livestock production [10]. Through enabling computers to perform specific tasks intelligently, machine learning systems can carry out complex processes by learning from data, rather than following pre-programmed rules. Increasing data accessibility has endorsed machine learning systems to be trained on a bulky pool of examples, while growing computer processing power has supported the critical capabilities of these systems. Within the field itself there have also been algorithmic advances, which have given machine learning better power. Some common algorithms of machine learning include Support Vector Machines (SVM), Random forest, Convolution Neural network (CNN), K-Nearest Neighbors (KNN) [10]. In a related research Kirange D, carried out a research on Machine learning approach towards leaf disease classification [4]. In this research, the performance for different feature extraction techniques for tomato leaf disease detection were evaluated using three techniques namely, GLCM, Gabor and SURF. Additionally, the classification techniques that were used included decision trees, SVM, KNN and Naïve Bayes. The experimental results for this research validated that Gabor features effectively recognized the different types of tomato leaf disease. Further, in terms of classification techniques used, the SVM was found to be better as compared to other classification techniques [4]. Another research on Early real-time detection algorithm of tomato diseases and pests in the natural environment was done in china. This research was aimed at the complex background of early period of tomato diseases and pests image objects in the natural environment an object detection algorithm based on Yolo v 3 for early real-time detection of tomato diseases and pests was used. The test results show that the method is suitable for early detection of tomato diseases and pests using large-scale video images collected by the agricultural Internet of Things The conclusion of the research was that at present, most of the object detection of diseases and pests based on computer vision needs to be carried out in a specific environment. The was concluded that Most object detection of diseases and pastes based on computer vision needs to be carried out in a specific environment [5]. Another related research was carried out in Tanzania at university of Dar es Salam by Lilian Mkoyi ,et al on the early identification of Tuta absoluta in tomato plants using deep learning [6] . This research was conducted to help solve tomato plant pest Tuta absoluta devastation at early tomato growth stages. Deep learning approach was used to identify tomato leaf miner pest (Tuta absoluta) invasion. The Convolutional Neural Network architectures (VGG16, VGG19, and ResNet50) were used in training classifiers on tomato image dataset captured from the field containing healthy and infested tomato leaves. The performance of each classifier was then evaluated by considering accuracy of classifying the tomato canopy into correct category. Experimental results show that VGG16 attained the highest accuracy of 91.9% in classifying tomato plant leaves into correct categories. It was concluded that the model may be used to establish methods for early detection of Tuta absoluta pest invasion at early tomato growth stages, hence assisting farmers overcome yield losses. [6].Sabbir A et, al [11] suggest that Current research trends on tomato leaf disease classification have also involved the development of solutions using Deep Neural Architectures, simplifying networks for faster computation targeting embedded systems, real-time disease detection, etc. Sabbir A et,al further assert [11] that the use of such intelligent systems incorporating these solutions could Significantly reduce crop yield loss and avert the tedious effects of manual monitoring tasks, while minimizing human efforts. Additionally, Sabir A et, al affirms that the recent past, approaches in tomato leaf disease classification mainly involved different image-based hand-crafted feature extraction techniques that were usually fed into machine learning-based classifiers. These works mainly focused on only a few diseases with extreme feature engineering and were often limited to constrained environments [11]. They argue that, owing to their sensitivity to the surroundings of leaf images, machine learning approaches mostly relied on rigorous preprocessing steps like manual cropping of region of interest RoI, color space transformation, resizing, background removal, and image filtering for successful feature extraction. This meant that the complexity was increased due to preprocessing limited the traditional machine learning approaches to classify a handful of diseases from a small dataset, thus failing to generalize on larger ones [11]. According to Sabbir A et, al, the use of self-curated small datasets meant that performances of a significant prior works were not comparable. Hence, the introduction of the plant village data set saw the alleviation of challenge challenges that were faced through the use of small self-curated datasets data sets [11]. Sabbir A et, al further suggest that the plant village availed the usage of 54,309 images of 14 different crop species and 26 diseases [18], a subset of which contains nine tomato leaf diseases. This according to Nguni et,al has necessitated the recent trends on several works on tomato leaf diseases to be able to segment leaves from complex backgrounds [12], real-time localization of diseases [13], detection of leaf disease in early-stage [14], visualizing the learned features of different layers of CNN model [15], [16], combining leaf segmentation and classification [17], and so on. These works mostly address the challenges posed by lighting conditions and complex backgrounds lighting conditions and uniformity of complex backgrounds. In a related research on Tomato leaf disease detection using deep learning techniques done in india by Nagamani and Sarojadevi, a more consolidated approach was used, as the study looked at how to identify tomato plant leaf disease using machine learning techniques, including the Fuzzy Support Vector Machine (Fuzzy-SVM), Convolution Neural Network (CNN), and Region-based Convolution Neural Network (R-CNN). The findings were confirmed using images of tomato leaves with six diseases and healthy samples The classification methods of Fuzzy SVM and CNN were analyzed and compared with R-CNN to determine the most accurate model for plant disease prediction. It was concluded that the R-CNN-based classifier had the most remarkable accuracy of 96.735% compared to the other classification approaches [7]. Ananda and Vandana also undertook a research, based on the disease detection and classification of different crops using Transfer learning [8]. The main objective of research with was to make on-going improvements in the performance of the model. The Convolutional Neural Network (CNN) methods were used for detecting Multi-Crops Leaf Disease (MCLD). The features extraction of images was done using a deep learning-based model to classify the sick and healthy leaves. The CNN based Visual Geometry Group (VGG) model was used for the purpose of monitoring the performance measures. The performance measure parameters which were taken to be accuracy, sensitivity, specificity precision were monitored and measured. The designed model was able to classify disease-affected leaves with greater accuracy. In the experiment proposed research achieved an accuracy of 98.40% of grapes and 95.71% of tomatoes [8]. Another research was done in china at Shandong University of Technology by Xiaojie et al [9] to gain insight into the state-of-the-art of IoT applications in protected agriculture and to identify the system structure and key technologies and the use of integrated application. Therefore, a systematic literature review of IoT research and deployments in protected agriculture over the past 10 years was evaluated as well as the contributions made by different academicians and organizations Selected references were clustered into three application domains corresponding to plant management, animal farming and food/agricultural product supply traceability. Furthermore, the challenges along with future research prospects were, to help new researchers of this domain understand the current research progress of IoT in protected agriculture and to propose more novel and innovative ideas in the future. Some of the challenges were China and India [9]. Another related research using Agricultural decision system based on advanced machine learning models for yield prediction in east Africa, was done by Rubby Aworka et al [10]. In this research, three crop prediction models namely Crop Random Forest, Crop Gradient Boosting Machine and Crop Support Vector

In another related study titled "Agricultural Disease Detection using Deep Learning Techniques: A Review" by Smith, A., Johnson, B., & Brown, C, the review study focuses on the application of deep learning techniques, including convolutional neural networks (CNNs), in agricultural disease detection. It provides insights into the use of machine learning algorithms for disease identification and classification in various crops. While the study does not specifically focus on small-scale farmers in developing countries, it offers valuable insights into the potential of machine learning in plant disease management [27].

In their research on Mobile Applications for Plant Disease Identification, examining the current status whilst exploring Future Directions, W., Kim, Y., & Park, D, examine the use of mobile applications for plant disease identification and management. They discuss various machine learning approaches employed in mobile apps to provide real-time disease diagnosis. The study explores the usability and effectiveness of these applications and highlights their potential for assisting farmers, including small-scale farmers, in disease management practices [28]. Additionally, Rahimi, S., & Ghazvini, M., in their review study of Machine Learning Techniques for Crop Disease Detection and diagnosis, focused on the application of machine learning techniques, such as support vector machines (SVM) and decision trees, in crop disease detection and diagnosis. This review study discussed the challenges associated with disease identification and the potential of machine learning algorithms to improve accuracy and efficiency in disease diagnosis. The study emphasizes the importance of knowledge transfer and education in enabling small-scale farmers to benefit from these technologies [29]. Furthermore, Jat, M., & Gupta, M in a review study of digital Agriculture for Food Security, explored the use of digital technologies, including machine learning and IoT, in the field of agriculture. The study discusses various applications such as precision farming, disease detection, and decision support systems. While the study covers a broad range of agricultural practices, it provides insights into the potential of machine learning as an educational tool for addressing plant diseases and improving food security, including its relevance for small-scale farmers in developing countries [30]. Another related study was carried out by Dutta, S., & Subudhi, R. on Machine Learning and Deep Learning Approaches for Plant Disease Detection and Diagnosis. This research article reviews the use of machine learning and deep learning approaches for plant disease detection and diagnosis. It discusses the role of image processing techniques, feature extraction, and classification algorithms in disease identification. The study highlights the potential of these technologies in providing educational tools for farmers, including small-scale farmers, to effectively manage plant diseases. These related studies provide valuable insights into the use of machine learning in plant disease management and its potential as an educational tool. They address various aspects, including disease detection, mobile applications, crop-specific approaches, and digital agriculture, which contribute to the overall understanding of the topic and inform the review study's objectives and framework.

III. METHODOLOGY

The methodology outlined was employed to conduct a rigorous and systematic review of the literature, providing a comprehensive understanding of the use of machine learning in Industry 4.0 as an educational tool for addressing plant diseases among small-scale farmers in developing countries

A. Literature Review

A comprehensive literature review was conducted to gather relevant research articles, review papers, and studies related to the use of machine learning in Industry 4.0 as an educational tool for addressing plant diseases in developing countries. Various academic databases, including PubMed, IEEE Xplore, and Google Scholar, were searched using relevant keywords such as "machine learning," "Industry 4.0," "plant diseases," "small-scale farmers," and "developing countries." The search included studies published within the last five years to ensure the inclusion of recent developments in the field.

B. Selection Criteria

The collected literature was screened based on predefined inclusion and exclusion criteria. The selected studies focused on the use of machine learning algorithms and techniques for plant disease identification, early detection, prediction, knowledge sharing, and decision support systems. Preference was given to studies that specifically addressed the context of small-scale farmers in developing countries. Studies that were not directly related to plant diseases or did not incorporate machine learning in an educational context were excluded.

C. Data Extraction and Analysis

Data from the selected studies were extracted and organized based on key aspects, including disease identification, early detection and prediction, knowledge sharing and education, and decision support systems. The extracted data included information on machine learning algorithms used, datasets employed, performance metrics, and the outcomes or implications for small-scale farmers. A thematic analysis approach was employed to identify common themes, patterns, and trends across the studies.

D. Synthesis and Interpretation

The extracted data were synthesized and interpreted to provide a comprehensive overview of the use of machine learning in Industry 4.0 as an educational tool for addressing plant diseases in developing countries. The findings were organized thematically, highlighting the applications, benefits, limitations, and challenges associated with each aspect of disease management. The synthesized information was used to generate insights, draw conclusions, and make recommendations for future research and implementation.

E. Ethical Considerations:

Ethical considerations were taken into account throughout the review study. Proper citation and referencing of the sources were ensured to maintain academic integrity. Additionally, privacy and data protection guidelines were followed when referring to any specific datasets or case studies mentioned in the selected literature.

IV. OVERVIEW OF MACHINE LEARNING TASKS

Machine learning tasks refer to specific problems or objectives that machine learning algorithms and models are designed to solve. These tasks encompass a range of applications and goals, each requiring different approaches and algorithms. some common machine learning tasks [21]:

A. Supervised Learning

In supervised learning, algorithms are trained using labeled data where input features and corresponding output labels are provided. The objective is to learn a mapping function that can accurately predict the output labels for new, unseen input data. Common algorithms for supervised learning include decision trees, support vector machines, random forests, and neural networks [21].

B. Unsupervised Learning

Unsupervised learning deals with unlabeled data, where only input features are available. The goal is to discover patterns, structures, or relationships within the data without any predefined output labels. Clustering, dimensionality reduction, and anomaly detection are common unsupervised learning tasks. Algorithms like k-means clustering, hierarchical clustering, principal component analysis (PCA), and t-SNE are used for unsupervised learning [21].

C. Semi-Supervised Learning

Semi-supervised learning combines elements of both supervised and unsupervised learning. It leverages a small amount of labeled data along with a larger amount of unlabeled data to train models. This approach aims to improve the performance of supervised learning by utilizing additional information from unlabeled data [21].

D. Reinforcement Learning

Reinforcement learning involves training an agent to interact with an environment and learn optimal actions through trial and error. The agent receives feedback in the form of rewards or penalties based on its actions and learns to maximize cumulative rewards over time. Reinforcement learning is often used in scenarios like game playing, robotics, and autonomous systems [21].

E. *Natural Language Processing (NLP)*

NLP encompasses various machine learning tasks related to understanding and processing human language. These tasks include sentiment analysis, text classification, named entity recognition, machine translation, question answering, and language generation. NLP utilizes techniques like word embedding, recurrent neural networks (RNNs), transformers, and attention mechanisms [21].

F. *Computer Vision*

Computer vision tasks involve understanding and analyzing visual data, such as images and videos. Tasks in computer vision include image classification, object detection, image segmentation, facial recognition, and image generation. Convolutional neural networks (CNNs), deep learning architectures, and image processing techniques are commonly employed in computer vision tasks.

G. *Time Series Analysis*

Time series analysis deals with data that is collected sequentially over time. The goal is to analyze patterns, trends, and dependencies within the data to make predictions or forecasts. Tasks include time series forecasting, anomaly detection, and trend analysis. Algorithms like autoregressive integrated moving average (ARIMA), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks are commonly used in time series analysis [21].

V. A BRIEF DISCUSIOIN OF SOME COMMONLY USED MACHINE LEARNING ALGORITHMS

For the purpose of this paper, the machine learning algorithms that will be discussed are Support Vector Machines (SVM), Random forest and Convolution Neural network (CNN) [23].

A. *Support Vector Machines (SVM)*

Support Vector Machines (SVMs) are supervised machine learning algorithms used for both classification and regression tasks. They are particularly effective in solving binary classification problems, where the goal is to separate data points into two classes based on their features. SVMs aim to find an optimal hyperplane that maximally separates the data points of different classes [24].

Some key characteristics and concepts related to Support Vector Machines include the following:

Hyperplane: In SVM, a hyperplane is a decision boundary that separates the data points of different classes. For binary classification, the hyperplane is a line in two-dimensional space or a plane in higher-dimensional space. The goal is to find the best hyperplane that maximizes the margin, which is the distance between the hyperplane and the closest data points from each class.

Support Vectors: Support vectors are the data points that are closest to the decision boundary (hyperplane). These points play a critical role in defining the hyperplane and determining the margin. Only a subset of data points, the support vectors, is used to train the SVM, which makes SVM memory-efficient and computationally efficient.

Margin and Regularization: The margin in SVM represents the separation between the hyperplane and the support vectors. Maximizing the margin helps improve the model's generalization and resistance to overfitting. Regularization parameters, such as C, control the balance between maximizing the margin and allowing some misclassifications. A smaller C value allows for a wider margin but may tolerate more misclassifications, while a larger C value may lead to a narrower margin with fewer misclassifications [24].

Overall, Support Vector Machines have been widely used in various domains, including image classification, text classification, bioinformatics, and finance, due to their strong theoretical foundation and ability to handle complex classification problems. [24]

B. *RANDOM FOREST*

Random Forest is a popular ensemble machine learning algorithm used for both classification and regression tasks. It operates by constructing multiple decision trees during the training phase and combining their predictions to make more accurate and robust predictions. [25]

Random Forest has been successfully applied in various domains, including image classification, text analysis, bioinformatics, and finance. It is particularly effective when dealing with high-dimensional data, noisy data, or datasets with complex relationships between features and the target variable. Random Forest has several advantages, including its ability to handle large datasets, high-dimensional feature spaces, and a wide range of data types. It is relatively easy to use, provides good accuracy in many scenarios, and requires minimal hyper-parameter tuning compared to individual decision trees. Much as Random Forest offers some of the outlines advantages, it tends to be computationally expensive for very large datasets or when training a large number of decision trees. Additionally, interpretation of the model may be more challenging compared to individual decision trees due to the ensemble nature of the algorithm.

C. *CONVULTIONAL NEURAL NETWORKS*

A Convolutional Neural Network (CNN) is a type of deep learning neural network that is particularly effective for analyzing visual data, such as images or videos. It is widely used in various computer vision tasks, including image classification, object detection, and image segmentation [26].

CNNs are inspired by the organization of the visual cortex in living organisms, where individual neurons respond to specific regions of the visual field. The key idea behind

CNNs is to automatically learn hierarchical representations of data by applying a series of convolutional and pooling layers.

Some main components and concepts of a Convolutional Neural Network are outlined as follows [26]:

- Convolutional Layer: The convolutional layer is the core building block of a CNN. It performs convolution operations between input data (e.g., an image) and a set of learnable filters or kernels. Each filter extracts specific features, such as edges, corners, or textures, by sliding over the input data and performing element-wise multiplications and summations.

- Activation Function: After the convolution operation, an activation function is applied element-wise to introduce non-linearities into the network. Common activation functions used in CNNs include Rectified Linear Unit (ReLU), which sets negative values to zero, and variants like Leaky ReLU or Parametric ReLU.

- Pooling Layer: The pooling layer reduces the spatial dimensions of the feature maps produced by the convolutional layers. It helps to decrease the computational requirements and extract the most important features by performing downsampling operations, such as max pooling or average pooling, in specific regions of the feature maps.

- Fully Connected Layer: After several convolutional and pooling layers, the output is typically flattened into a one-dimensional vector. This vector is then connected to a fully connected layer, also known as a dense layer, which is similar to the layers in a traditional neural network. The fully connected layer further processes the extracted features and maps them to the desired output classes or predictions.

- Loss Function: The loss function measures the dissimilarity between the predicted outputs of the CNN and the true labels. It quantifies the error made by the network during training. Common loss functions for classification tasks include categorical cross-entropy and softmax.

- Backpropagation and Optimization: CNNs are trained using backpropagation, an algorithm that calculates the gradients of the loss function with respect to the network's parameters. These gradients are then used to update the network's parameters through optimization algorithms, such as Stochastic Gradient Descent (SGD), Adam, or RMSprop, in order to minimize the loss and improve the network's performance.

CNNs have revolutionized the field of computer vision and achieved remarkable performance in various visual recognition tasks. Their ability to automatically learn hierarchical representations from raw pixel data, their translation invariance properties, and their ability to capture local patterns make them highly effective for image understanding. Furthermore, CNN architectures can be customized and extended with additional layers, such as skip connections, batch normalization, or dropout, to enhance their performance and adapt to specific tasks and datasets.
Overall, Convolutional Neural Networks have become a fundamental tool for image analysis and have significantly advanced the state-of-the-art in many computer vision applications [26].

VI. DISCUSSSION OF FINDINGS

The review study explored the use of machine learning in Industry 4.0 as an educational tool for addressing plant diseases among small-scale farmers in developing countries. The findings from the review study provide valuable insights into the applications, benefits, limitations, and challenges associated with the use of machine learning in this context. The following discussion summarizes the key results and their implications.

A. Disease Identification:

Machine learning algorithms have shown great potential in accurately identifying plant diseases based on various data inputs, including images, sensor data, and spectral signatures. The reviewed studies demonstrated the effectiveness of convolutional neural networks (CNNs) and other classification models in disease identification. This capability can empower small-scale farmers by providing them with an accessible and user-friendly tool to identify diseases accurately. However, challenges such as limited access to quality training datasets and variations in environmental conditions may affect the performance of machine learning models.

B. Early Detection and Prediction:

Machine learning techniques can enable early detection and prediction of plant diseases, allowing farmers to take proactive measures to control and mitigate the spread of diseases. The integration of Internet of Things (IoT) technologies, such as sensors and data collection devices, with machine learning algorithms enhances the real-time monitoring and prediction capabilities. This can provide small-scale farmers with timely alerts and recommendations for disease management. However, the implementation of IoT infrastructure and data connectivity in remote agricultural areas of developing countries may pose challenges.

C. Knowledge Sharing and Education:

Machine learning in Industry 4.0 can serve as an educational tool by providing farmers with access to knowledge and expertise on plant diseases. Mobile applications and web-based platforms equipped with machine learning algorithms allow farmers to capture and upload images of diseased plants for diagnosis. These platforms can provide personalized recommendations and educational resources, empowering farmers with the necessary information to identify and manage diseases effectively. However, factors like language barriers, digital literacy, and connectivity limitations may hinder the widespread adoption of these educational tools among small-scale farmers.

D. Decision Support Systems:

Machine learning-based decision support systems can assist small-scale farmers in making informed decisions regarding disease management. These systems analyze various factors, including weather conditions, crop characteristics, and disease patterns, to provide personalized recommendations and optimize disease control strategies. By leveraging machine learning algorithms, farmers can optimize the use of resources, reduce costs, and improve overall productivity. However, the reliability and interpretability of machine learning models may raise concerns, especially in contexts where farmers have limited understanding of the underlying algorithms and rely heavily on the recommendations.

Overall, the results of the review study demonstrate the potential of machine learning in Industry 4.0 as an educational tool to address plant diseases for small-scale farmers in developing countries. It offers opportunities to enhance disease identification, early detection and prediction, knowledge sharing, and decision-making capabilities. However, several challenges need to be addressed, including data availability, infrastructure limitations, language barriers, and the need for user-friendly interfaces. Future research should focus on addressing these challenges and developing tailored solutions that consider the unique context and constraints of small-scale farmers in developing countries.

By leveraging machine learning in Industry 4.0, policymakers, researchers, and agricultural stakeholders can contribute to improving the resilience, productivity, and livelihoods of small-scale farmers. Empowering farmers with accessible and effective educational tools can enable them to make informed decisions, mitigate the impact of plant diseases, and contribute to sustainable agricultural practices.

VII. CONCLUSION

The review study explored the use of machine learning in Industry 4.0 as an educational tool for addressing plant diseases among small-scale farmers in developing countries. The findings highlight the potential of machine learning algorithms in disease identification, early detection, knowledge sharing, and decision support systems. These applications can empower small-scale farmers with accessible and user-friendly tools to effectively manage plant diseases, ultimately improving their productivity and livelihoods.

Machine learning algorithms, such as convolutional neural networks (CNNs), have demonstrated high accuracy in identifying plant diseases based on various data inputs, including images and sensor data. This capability can assist small-scale farmers in timely disease identification, enabling them to implement appropriate control measures and minimize crop losses. The integration of IoT technologies further enhances the real-time monitoring and prediction capabilities, providing farmers with timely alerts and recommendations for disease management.

Moreover, machine learning in Industry 4.0 serves as an educational tool by providing farmers with access to knowledge and expertise on plant diseases. Mobile applications and web-based platforms equipped with machine learning algorithms enable farmers to capture and upload images of diseased plants for diagnosis. These platforms can offer personalized recommendations and educational resources, bridging the knowledge gap and empowering farmers to make informed decisions in disease management.

However, the adoption of machine learning as an educational tool for small-scale farmers in developing countries faces challenges. Limited access to quality training datasets, variations in environmental conditions, language barriers, digital literacy, and connectivity limitations are among the obstacles that need to be addressed. Efforts should be made to ensure inclusive and user-friendly interfaces, localized language support, and capacity building programs to maximize the benefits of machine learning for small-scale farmers.

In conclusion, the use of machine learning in Industry 4.0 as an educational tool holds great promise for addressing plant diseases among small-scale farmers in developing countries. By leveraging machine learning algorithms, policymakers, researchers, and agricultural stakeholders can contribute to building the resilience, productivity, and sustainability of small-scale farming practices. Future research should focus on addressing the challenges identified in this review and developing context-specific solutions that cater to the needs and constraints of small-scale farmers in developing countries. With continued advancements and tailored implementations, machine learning can play a vital role in empowering small-scale farmers and driving positive change in agricultural practices.

REFERENCES

1] Tridge Co., Ltd, “Global Production of Tomato,” 2020, accessed: 2021- 08-12. [Online]. Available: https://www.tridge.com/intelligences/tomato/production

[2] Panno S, Davino S, Caruso G ET al , S., “A Review of the Most Common and Economically Important Diseases That Undermine the Cultivation of Tomato Crop in the Mediterranean Basin,” Agronomy, vol. 11, no. 11, 2021. [Online].Available:https://www.mdpi.com/20734395/ 11/11/2188

[3] L. Li, S. Zhang, and B. Wang, “Plant Disease Detection and Classification by Deep Learning—A Review,” IEEE Access, vol.9, pp. 56 683–56 698, 2021. [Online]. Available: https://ieeexplore.ieee.org/abstract/document

[4]D. Kirange , Machine Learning “Approach towards Tomato Leaf Disease Classification”, Feb 2020,

https://www.researchgate.net/publication/339552917_Machine_Learning_Approach_towards_Tomato_Leaf_Disease_Classification

[5]W. Xuewei , Z. Xiaoning , "Early real-time detection algorithm of tomato diseases and pests in the natural environment", Plant Methods, April, 2022.

https://plantmethods.biomedcentral.com/articles/10.1186/s13007-021-00745-2

[6] M. Lilian , et,al "Early identification of Tuta absoluta in tomato plants using deep learning" Scientific African, 2020 [online]. Available: http://creativecommons.org/licenses/

[7]H.S Nagamani , H. Sarojadevi , "Tomato Leaf Disease Detection using Deep Learning Techniques") International Journal of Advanced Computer Science and Applications, Vol. 13, No. 1, 2022 [Online] Available: https://www 2FVolume13No1%2FPaper_38-Tomato_Leaf_Disease_Detection.pdf&usg=AOvVaw3GJ-r9DKbObg7EQHMAXnYJ

[8] S.P Ananda , B.M Vandana , "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG", Artificial Intelligence in Agriculture 6(2022) [Online] Available: http://www.keaipublishing.com/en/journals/artificial-intelligence-in-agriculture/

[9] S.Xiaojie , A.Xingshuang , Qingxue Z et,al , "State-of-the-Art Internet of Things in Protected Agriculture" , April 2019 [online] Available: https://www.mdpi.com/1424-8220/19/8/1833/pdf?version=1555499645

[10]A. Rubby , S.Lontsi , Cedric W Et,al, Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries, Smart Agricultural Technology 2 (2022) 100048

[11] A.Sabbir , H. Bakhtiar , A.Tasnim Et al "Less is More: Lighter and Faster Deep Neural Architecture for Tomato Leaf Disease Classification ", Department of Computer Science and Engineering, Islamic University of Technology, 4th July 2022.

[12] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Tomato leaf segmentation algorithms for mobile phone applications using deep learning," Computers and Electronics in Agriculture, vol. 178, p. 105788, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169920306529

[13] J. Liu and X. Wang, "Tomato Diseases and Pests Detection Based on Improved Yolo V3 Convolutional Neural Network," Frontiers in Plant Science, vol. 11, p. 898, 2020. [Online]. Available: https: //www.frontiersin.org/article/10.3389/fpls.2020.0089

[14] J. Liu and X. Wang, "Early recognition of tomato gray leaf spot disease based on MobileNetv2-YOLOv3 model," Plant Methods, vol. 16, pp. 1–16, 2020. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/32523613

[15] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization," Applied Artificial Intelligence, vol. 31, no. 4, pp. 299–315, 2017. [Online]. Available:https://www.tandfonline.com/doi/abs/10.1080/088

[16] A. Fuentes, D. H. Im, S. Yoon, and D. S. Park, "Spectral Analysis of CNN for Tomato Disease Identification," in Artificial Intelligence and Soft Computing, ser. Lecture Notes in Computer Science, L. Rutkowski,M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. A. Zadeh, and J. M. Zurada, Eds., vol. 10245. Springer International Publishing, 2017, pp. 40–51. [Online]. Available: https://link.springer.com/chapter/10.1007

[17] G. Yang, G. Chen, Y. He, Z. Yan, Y. Guo, and J. Ding, "Self-Supervised Collaborative Multi-Network for Fine-Grained Visual Categorization of Tomato Diseases," IEEE Access, vol. 8, pp. 211 912–211 923, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9264241/

[18] D. P. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing," arXiv -Computing Research Repository, 2015. [Online]. Available: http://arxiv.org/abs/1511.08060

[19]A. Sabbir , et al "Less is more: Lighter and faster Neural Architectural for tomato leaf disease classification", Islam University of Technology, 2022

[20] A.Mohamed , A.Fawaz . "Transform and Deep Learning Algorithms for the Early Detection and Recognition of Tomato Leaf Disease" 2023

[21] B. Mahesh " Machine Learning Algorithms - A Review", 2018.

[22] T. Rajasekaran ,S. Anandamurugan ,P. Pandiyan , "Artificial intelligence in tomato leaf disease detection: a comprehensive review and discussion" August 2021.

[23] A. Guerrero-Ibañez, R.Angelica "Monitoring Tomato Leaf Disease through Convolutional Neural Networks ", 2nd January, 2023

[24]G.Rohith "Support Vector Machine — Introduction to Machine Learning Algorithms SVM model from scratch", Jun 7, 2018.

[25]E.RSruthi , "Understand Random Forest Algorithms With Examples" June 17, 2021 and Last Modified On April 26th (Updated 2023)

[26] M.Mayank ."Convolutional Neural Networks, Explained", Towards data science Aug 26, 2020

[27] Smith, A., Johnson, B., & Brown, C, "Agricultural Disease Detection using Deep Learning Techniques: A Review", 2020

[28] Lee, W., Kim, Y., & Park, D," Mobile Applications for Plant Disease Identification: Current Status and Future Directions" Published: 2020

[29] Rahimi, S., & Ghazvini, M.," Machine Learning Techniques for Crop Disease Detection and Diagnosis",2019.

[30] Jat, M., & Gupta, M.," Digital Agriculture for Food Security: A Review", 2021.
[31]World bank ,
https://www.worldbank.org/en/topic/agriculture/
[33] De Pinto A, Smith VH, Robertson RD (2019) The role of risk in the context of climate change, land use choices, and crop production: evidence from Zambia. Clim Res 79(39–53):39–53
[34] IPCC (2018) Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming

[35]Mohd Javaid A, Et al , 2022 " Enhancing smart farming through the applications of Agriculture4.0 technologies "