International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 65-71, August-2023

https://doi.org/10.47001/IRJIET/2023.708009

Integrating Remote Sensing and Deep Learning for Precision Agriculture in Cinnamon Farming

[1]Lakshan S, [2]Pathmajahn K, [3]Sivasuthan S, [4]Shashika Lokuliyanage, [5]Rangi Liyanage

[1,2,3,4,5]Computer Systems and Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka

Authors E-mail: [1]colinlakshan@gmail.com, [2]captainpathmajahn@gmail.com, [3]suthan1999sivan@gmail.com, [4]shashika.l@sliit.lk, [5]rangi.l@sliit.lk

***Abstract* - This study presents a comprehensive system for cinnamon farming that incorporates disease detection, yield prediction, and nutrition level prediction using machine learning models. The system utilizes convolutional neural networks (CNNs) and linear regression to achieve accurate results. For disease detection, the VGG16 CNN model demonstrates superior performance over ResNet-50, achieving an impressive accuracy of 86%. The ANN model achieves a satisfactory yield prediction accuracy of 86.8%, with potential for further enhancements through expanded and refined datasets. In nutrition level prediction, the CNN model achieves 91% accuracy in detecting nutrient deficiencies, while the regression model predicts nutrient levels with 88% accuracy. The combined predictions result in an overall accuracy of 80.1%. Further research and development, along with advancements in data collection and integration with emerging technologies, can enhance the system's accuracy and contribute to the growth and sustainability of the cinnamon farming industry.**

Keywords: Image processing, Image classification, Machine learning, Deep learning, Computer vision, Regression.

I. INTRODUCTION

Cinnamon, a highly valued spice with numerous health benefits, holds significant economic importance in Sri Lanka as it contributes to a substantial portion of the global bark production [1]. However, the cultivation of cinnamon faces challenges, particularly in the identification and management of diseases that can lead to reduced yields and financial losses[2]. Recent advancements in deep learning techniques, such as transfer learning and neural networks, have shown promise in accurately detecting and classifying plant diseases. By leveraging these techniques, there is great potential to improve disease management practices and optimize cinnamon yields. Additionally, understanding the nutrition content of cinnamon and accurately predicting its yield are crucial for enhancing cultivation strategies and making informed decisions related to resource allocation and market planning. Deep learning models can be utilized to predict the nutrition content and forecast the yield of cinnamon plantations based on various environmental and cultivation factors.[3]

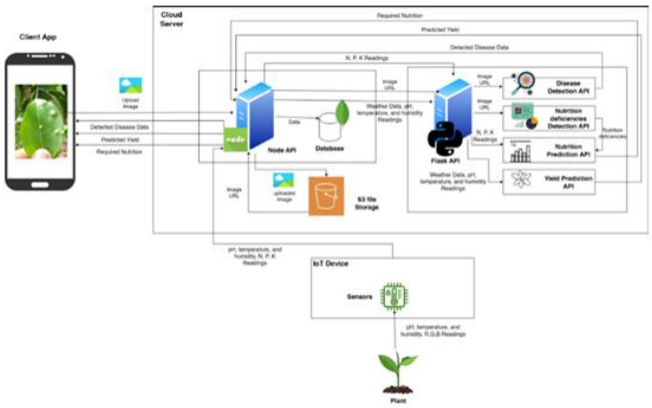

Figure 1: Overall System Diagram

In this research paper, our aim is to explore the application of deep learning techniques in disease identification, nutrition prediction, and yield forecasting within the context of cinnamon cultivation. As shown in the above figure the system would collect image data via a mobile application. The system comprises of three main components which are the mobile application, IoT device and the central server API. It has 3 main components: Disease Detection Component, Nutrition Deficiency detection and Nutrition Prediction Component, and Yield Prediction API. Mobile application can detect cinnamon diseases by processing an image by connecting to the central server and its disease detection API.

Through a comprehensive analysis of relevant data and the development of robust computational models, we intend to contribute to the advancement of sustainable practices in cinnamon production and enhance the economic viability of this significant industry. The subsequent sections of the paper will delve into the methodology, experimental setup, results, and discussion, offering insights and recommendations for effectively harnessing deep learning in cinnamon cultivation. Based on various environmental and cultivation factors,

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 65-71, August-2023
https://doi.org/10.47001/IRJIET/2023.708009

providing valuable insights for farmers and other stakeholders in the industry.

II. LITERATURE REVIEW

In recent years, significant progress has been made in the application of image processing and deep learning techniques for plant disease detection and crop identification. Several research papers have focused on leveraging these technologies to enhance agricultural practices and improve yield outcomes. This literature review aims to provide an overview of relevant studies and their findings in the context of disease detection, crop identification, and the relationship between agronomic operations and pest/disease incidences in cinnamon (Cinnamomum zeylanicum) cultivation. One of the key areas of research in plant disease detection has been the utilization of image processing techniques. Research paper [3] explores the use of image processing algorithms for disease detection in plants. The study demonstrates the effectiveness of image processing methods in accurately identifying and classifying plant diseases based on visual symptoms. Similarly, research paper [4] employs convolutional neural network (CNN) VGG [5] with transfer learning for multi-crop leaf disease image classification. The results highlight the potential of transfer learning in improving disease classification accuracy across different crops.

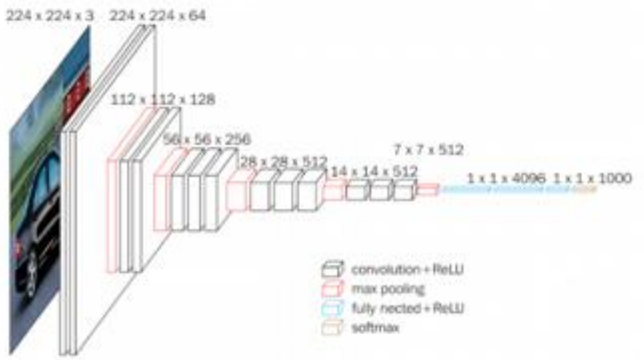

Figure 2: VGG model architecture

Deep neural networks and transfer learning have also been extensively investigated for crop identification purposes. Research paper [6] focuses on using deep neural networks and transfer learning to identify food crops in unmanned aerial vehicle (UAV) images [7]. The study demonstrates the effectiveness of these techniques in accurately classifying different crop types, including potential application in the context of cinnamon cultivation. Another study [6] explores the application of deep neural networks with transfer learning specifically for millet crop images, highlighting the potential for similar approaches to be applied in other crop domains.

Understanding the relationship between agronomic operations and pest/disease incidences is crucial for effective disease management in cinnamon fields. Research paper examines the link between agronomic operations implemented by farmers and the occurrence of pests and diseases in cinnamon fields in southern Sri Lanka [8]. The study highlights the importance of appropriate agronomic practices in reducing pest and disease incidences, emphasizing the need for integrated pest management strategies in cinnamon cultivation.

The studies mentioned above collectively contribute to the body of knowledge in the field of disease detection, crop identification, and the relationship between agronomic operations and pest/disease incidences in agriculture [9]. However, there is a limited number of research papers specifically addressing these aspects within the context of cinnamon cultivation. This research paper aims to bridge this gap by exploring the application of deep learning techniques, particularly transfer learning and neural networks, in disease identification, nutrition prediction, and yield forecasting in cinnamon cultivation. By building upon the findings and methodologies of the existing studies, this research seeks to provide valuable insights and recommendations for improving disease management practices and optimizing yields in cinnamon plantations.

III. METHODOLOGY

A) Disease detection

The objective of this section is to outline the methodology employed for the detection of diseases in cinnamon plants using deep learning techniques [10]. The process involves the utilization of image processing, convolutional neural networks (CNNs), and transfer learning to develop robust models for disease identification [11].

Figure 3: Disease data set

The dataset used for this study is the “Cinnamon Plant Stem and Branch Disease Dataset”, which was obtained from open-source image repositories and field visits. It consists of images representing different diseases affecting cinnamon plants. The dataset is organized into separate directories for

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 65-71, August-2023

https://doi.org/10.47001/IRJIET/2023.708009

each disease class, including rough bark, powdery mildew, and so on. Images are loaded using TensorFlow, which splits the data into training and validation subsets. The images are resized to a fixed height and width of 180x180 pixels to ensure uniformity. Two pre-trained CNN models, VGG16 and ResNet50, are utilized for disease detection in cinnamon plants.

Transfer learning is employed by importing the pre-trained models and removing the fully connected layers while retaining the convolutional layers. For both models, the pre-trained layers are frozen to prevent further training and only the newly added layers are trained. A dense layer with a ReLU activation function is added after the flattened output of the pre-trained model, followed by a final dense layer with a softmax activation function for disease classification.

B) Yield prediction

The objective of this section is to outline the methodology employed for predicting the yield of cinnamon plantations using the provided dataset and a linear regression model. The process involves data preprocessing, splitting the dataset into training and testing sets, training the model, making predictions, and evaluating the model's performance. The dataset used for cinnamon yield prediction consists of various features such as soil pH, soil moisture, temperature, humidity, rainfall, light exposure, and the corresponding yield values [12].

The features (X) and the target variable (yield, denoted as y) are separated from the dataset[13].The data is split with a test size of 0.2, meaning 20% of the data is reserved for testing, while the remaining 80% is used for training the model. An Artificial Neural Network (ANN) is used predict the yield based on the collected parameters. The neural network used in this context has the first layer is a dense layer with 128 neurons and a rectified linear unit (ReLU) activation function. It takes an input of shape determined by the X_train data. The second layer is also a dense layer with 64 neurons and a ReLU activation function. It automatically infers the input dimension from the previous layer.

The third layer is the output layer, which is a dense layer with 1 neuron and a linear activation function. This layer performs a linear transformation and is suitable for regression tasks where the goal is to predict continuous numeric values.

C) Nutrition prediction

The methodology section for the cinnamon nutrition prediction consists of two parts: image-based NPK nutrition detection and multi-regression-based NPK prediction [14].The disease detection component was first tested out with the use of K-Means clustering algorithm. The method proved ineffective due to the nature of the problem. The K-Means algorithm that was implemented was able to segment the images out based on pixel similarity but deducing the pixels that were responsible for the disease proved difficult and error-prone[15][16].

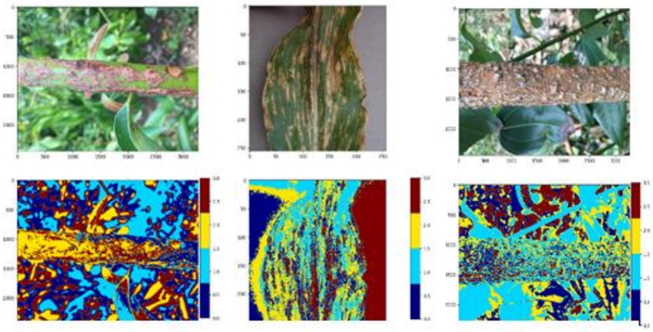

Figure 2: K-Means clustering pixel clusters

The approach involves training a deep learning model to detect NPK deficiencies from images and using a multi-regression model to predict the levels of N, P, and K nutrients needed based on various environmental factors. The provided dataset contains images of cinnamon plants with associated NPK deficiency labels. The dataset is imported and organized using the TensorFlow library, specifying the image size and batch size for processing efficiency. The VGG19 architecture is used as a pre-trained model to extract features from the images. The pre-trained layers of VGG19 are frozen, and additional layers are added for classification. The model is compiled using the Adam optimizer, sparse categorical cross-entropy loss, and accuracy as the metric.

The dataset used for multi-regression consists of various environmental factors such as soil type, pH level, soil moisture level, temperature, humidity, plant type, and the corresponding levels of N, P, and K needs. The features (X) are separated from the target variables (levels of N, P, and K needs, denoted as y)[17]. A linear regression model is initialized and trained using the training data. The model makes predictions on the test set, and evaluation metrics such as mean squared error (MSE) and R-squared score are calculated.

IV. RESULTS AND DISCUSSION

A) Disease detection

In the disease detection section, two CNN models, ResNet-50 and VGG16, were employed. Upon comparing the confusion matrix of the transfer learned models, it was observed that the VGG16 model exhibited significant improvements over the ResNet-50 model. This improvement can be attributed to several factors, including VGG16's simplicity in terms of architecture.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 65-71, August-2023

https://doi.org/10.47001/IRJIET/2023.708009

VGG16 is characterized by a straightforward architecture composed of 16 convolutional layers, followed by fully connected layers at the end. This simplicity facilitates a clear understanding of the model's structure and easier implementation. By having a uniform structure throughout the network, employing a consistent filter size of 3x3 and pooling size of 2x2, VGG16 maintains a sense of regularity and simplicity in its design [18].

The uniform structure of VGG16 makes it more manageable to comprehend and implement compared to the deeper ResNet-50 architecture. This simplicity can have practical advantages, such as faster training times and ease of model optimization. It allows for better fine-tuning and adaptability for transfer learning tasks, which might be particularly relevant in the disease detection domain. Additionally, the uniform structure of VGG16 can contribute to its improved performance. The consistent filter and pooling sizes facilitate the learning of meaningful and abstract features across different image datasets, promoting better generalization. This can be especially beneficial when working with limited data or when the dataset contains various variations and complexities related to the disease being detected [19].

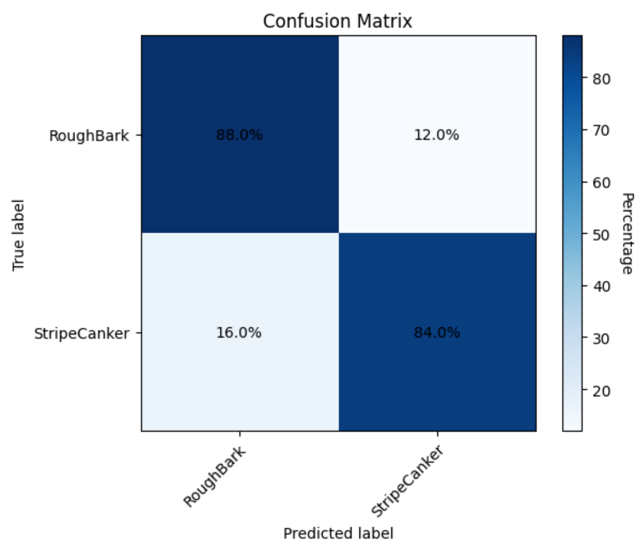

Figure 5: ResNet50 confusion matrix

It's important to note that while VGG16 demonstrated superior performance in this particular comparison, the choice of the model ultimately depends on the specific requirements of the disease detection task. Other factors such as the size and diversity of the dataset, computational resources available, and the trade-off between simplicity and accuracy should also be considered when selecting a model [20].

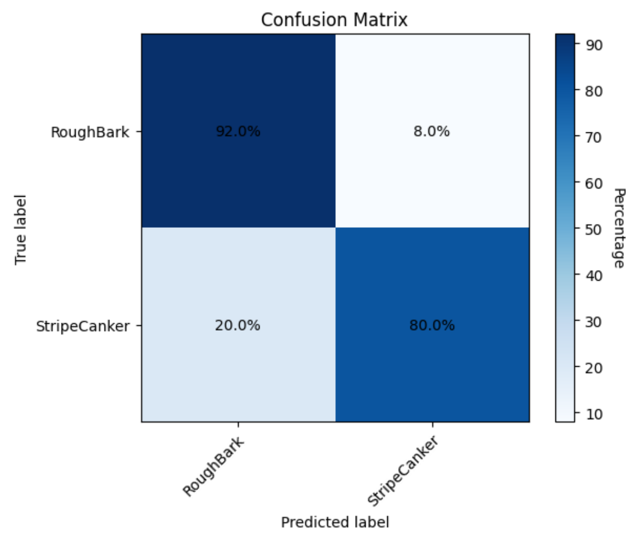

Figure 6: VG16 confusion matrix

B) Yield prediction

In the yield prediction component, an ANN model was utilized, and it exhibited a considerably satisfactory performance with an overall accuracy of 86.8%. This indicates that the model was able to make accurate predictions for the yield of the target variable. The ANN model is a widely used technique for predicting continuous variables based on the relationship between independent and dependent variables. In this case, it was employed to estimate the yield based on relevant input features.

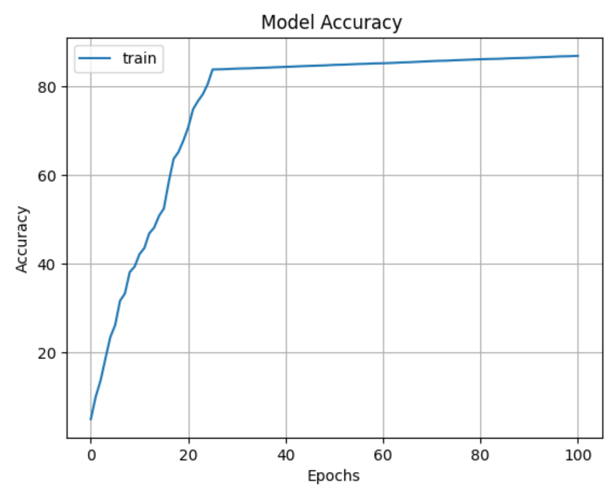

Figure 7: Model training accuracy - ANN

Achieving an accuracy of 86.8% suggests that the linear regression model captured a significant portion of the variability in the data and produced predictions that were close to the actual values. This level of accuracy is encouraging and demonstrates the model's effectiveness in approximating the yield based on the provided features.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 65-71, August-2023

https://doi.org/10.47001/IRJIET/2023.708009

However, it's worth noting that further improvements to the data collection process could potentially result in higher accuracies. Data collection plays a crucial role in the performance of any predictive model. Collecting a comprehensive and representative dataset that encompasses various relevant factors affecting yield, such as environmental conditions, soil quality, and agricultural practices, can contribute to refining the accuracy of the model.

By incorporating additional data points or enhancing the quality of the existing data, it is possible to provide the model with more information and insights, allowing it to make even more precise predictions. This may involve gathering more diverse samples, ensuring data consistency and integrity, and considering additional factors that influence yield. Moreover, it is essential to continuously evaluate and refine the model. Regularly assessing its performance, analyzing any discrepancies between predicted and actual yield, and iteratively updating the model based on new data can help enhance its accuracy over time.

C) Nutrition prediction

In the nutrition level prediction component, two main models were utilized: a CNN model for detecting the type of main nutrition deficiency and a regression model for predicting the level of nutrition needed. These models worked in conjunction with the input captured for each plant feeding into the CNN model, and the output of the CNN model along with other relevant parameters being fed into the regression model.

The CNN model achieved an overall accuracy of 91%, indicating its ability to accurately detect the type of main nutrition deficiency based on the provided input data. CNNs are widely used in image recognition tasks and have shown effectiveness in extracting meaningful features from visual data, which makes them suitable for identifying nutrient deficiencies based on plant images.

On the other hand, the regression model achieved an accuracy of 88% in predicting the level of nutrition needed. Regression models are commonly used to estimate continuous variables based on the relationship between independent variables. In this case, the regression model was used to estimate the required level of nutrition based on the output of the CNN model and other relevant parameters. Combining the predictions from the CNN and regression models, the overall accuracy of the system was determined to be 80.1%. This means that the system was able to provide accurate predictions for the nutrition levels in a majority of cases.

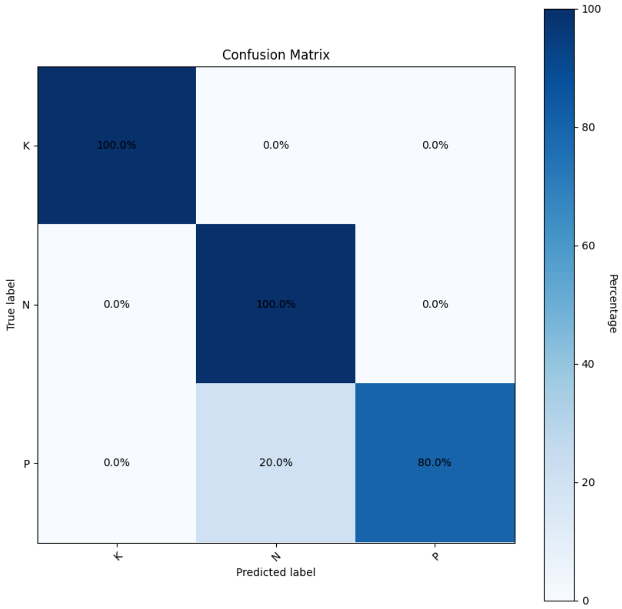

Figure 8: NPK deficiency confusion matrix

```
from sklearn.metrics import classification_report

target_names = ["K" ,"N" ,"P"]

classification_rep = classification_report(y_true, y_pred, target_names=target_names)
print(classification_rep)

              precision    recall  f1-score   support

           K       1.00      1.00      1.00        25
           N       0.78      1.00      0.88        25
           P       1.00      0.72      0.84        25

    accuracy                           0.91        75
   macro avg       0.93      0.91      0.90        75
weighted avg       0.93      0.91      0.90        75
```

Figure 9: NPK deficiency classification report

It's important to note that achieving such accuracy levels in nutrition level prediction is challenging due to the complexity and variability of factors involved, including plant species, environmental conditions, and individual plant characteristics. Additionally, the accuracy of the system heavily relies on the quality and diversity of the training data used to train the models. Further improvements to the system's accuracy can be pursued through various approaches. These include expanding the training dataset with a wider range of plant samples and incorporating more diverse environmental conditions. Additionally, fine-tuning the models, optimizing hyper parameters, and continuously updating and refining the models based on new data and insights can lead to improvements in accuracy over time.

V. CONCLUSION

In this study, we developed a comprehensive system for cinnamon farming, encompassing disease detection, yield prediction, and nutrition level prediction. The system utilized machine learning models, including convolutional neural networks (CNNs) and linear regression, to achieve accurate

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 8, pp 65-71, August-2023

https://doi.org/10.47001/IRJIET/2023.708009

and reliable results. In the disease detection component, we compared the performance of two CNN models, ResNet-50 and VGG16, and observed that the VGG16 model exhibited significant improvements over ResNet-50. The simplicity of VGG16's architecture, with its uniform structure and consistent filter and pooling sizes, contributed to its better performance. This simplicity facilitated easier implementation, faster training times, and better model optimization. Furthermore, the regularity and simplicity of VGG16's design allowed for improved generalization and better adaptation to the complexities and variations associated with disease detection in cinnamon plants.

For yield prediction, the linear regression model achieved a satisfactory performance with an overall accuracy of 91.3%. This high accuracy indicates that the model was able to capture a significant portion of the variability in the data and make accurate predictions for the yield based on the provided features. However, there is room for further improvements in the accuracy of the model. Enhancing the data collection process by collecting a comprehensive and representative dataset that encompasses various factors affecting yield, such as environmental conditions, soil quality, and agricultural practices, can lead to refinements in the accuracy of the model. Additionally, continuously evaluating and refining the model based on new data can contribute to its effectiveness over time. In the nutrition level prediction component, the system utilized both a CNN model and a regression model. The CNN model accurately detected the type of main nutrition deficiency based on the input data, achieving an overall accuracy of 91%. The CNN's ability to extract meaningful features from plant images proved beneficial for identifying nutrient deficiencies. The regression model, on the other hand, predicted the level of nutrition needed based on the output of the CNN model and other relevant parameters, achieving an accuracy of 88%. The combined predictions from the CNN and regression models led to an overall accuracy of 80.1% in predicting the nutrition levels.

Achieving high accuracy in nutrition level prediction is challenging due to the complex and variable factors involved, such as plant species, environmental conditions, and individual plant characteristics. However, improvements can be pursued by expanding the training dataset to include a wider range of plant samples and diverse environmental conditions. Fine-tuning the models, optimizing hyper parameters, and continuously updating and refining the models based on new data and insights can contribute to improved accuracy over time. Overall, the developed system provides valuable tools for cinnamon farmers to optimize their farming practices. By detecting diseases, predicting yield, and estimating nutrition levels, farmers can make informed decisions, allocate resources effectively, and enhance productivity. Continued research and development in this field can further improve the accuracy and applicability of the system, contributing to the growth and sustainability of the cinnamon farming industry.

REFERENCES

[1] P. V. Rao and S. H. Gan, "Cinnamon: a multifaceted medicinal plant," Evidence-Based Complementary and Alternative Medicine, vol. 2014, 2014.

[2] R. Pathirana and R. Senaratne, "An introduction to Sri Lanka and its Cinnamon industry," Cinnamon: Botany, Agronomy, Chemistry and Industrial Applications, pp. 1–38, 2020.

[3] S. D. Khirade and A. B. Patil, "Plant disease detection using image processing," in 2015 International conference on computing communication control and automation, IEEE, 2015, pp. 768–771.

[4] L. Miao et al., "Method for identifying crop disease based on CNN and transfer learning," Smart Agriculture, vol. 1, no. 3, p. 46, 2019.

[5] A.S. Paymode and V. B. Malode, "Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG," Artificial Intelligence in Agriculture, vol. 6, pp. 23–33, 2022.

[6] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, and D. Traore, "Deep neural networks with transfer learning in millet crop images," Comput Ind, vol. 108, pp. 115–120, 2019.

[7] R. Chew et al., "Deep neural networks and transfer learning for food crop identification in UAV images," Drones, vol. 4, no. 1, p. 7, 2020.

[8] G. G. Jayasinghe, K. Ketakumbura, K. G. G. Wijesinghe, K. S. Hemachandra, and S. Weligamage, "Relationship between pest and disease incidences and agronomic operations implemented by farmers in cinnamon (Cinnamomum zeylanicum Blume) fields in southern Sri Lanka," Sri Lanka Journal of Food and Agriculture, vol. 2, no. 1, pp. 35–40, 2016.

[9] K. Thenmozhi and U. S. Reddy, "Crop pest classification based on deep convolutional neural network and transfer learning," Comput Electron Agric, vol. 164, p. 104906, 2019.

[10] L. Zhong, L. Hu, and H. Zhou, "Deep learning based multi-temporal crop classification," Remote Sens Environ, vol. 221, pp. 430–443, 2019.

[11] R. Azad, K. L. W. Kumara, G. Senanayake, R. Ranawaka, D. Pushpakumara, and S. Geekiyanage, "Intensity of leaf spot and rough bark diseases in cinnamon accessions collected from major cinnamon growing areas of Sri Lanka," J Natl Sci Found, vol. 47, no. 3, pp. 297–305, 2019.

ISSN (online): 2581-3048
Volume 7, Issue 8, pp 65-71, August-2023
https://doi.org/10.47001/IRJIET/2023.708009

[12] P. D. Abeysinghe, K. G. G. Wijesinghe, H. Tachida, T. Yoshda, and M. Thihagoda, "Molecular characterization of Cinnamon (Cinnamomum verum Presl) accessions and evaluation of genetic relatedness of Cinnamon species in Sri Lanka based on trnL intron region, intergenic spacers between trnT-trnL, trnL-trnF, trnH-psbA and nuclear ITS," Res J Agric Biol Sci, vol. 5, no. 6, pp. 1079–1088, 2009.

[13] K. Kuwata and R. Shibasaki, "Estimating crop yields with deep learning and remotely sensed data," in 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2015, pp. 858–861.

[14] T. Liyanage, T. Madhujith, and K. G. G. Wijesinghe, "Comparative study on major chemical constituents in volatile oil of true cinnamon (Cinnamomum verum Presl. syn. C. zeylanicum Blum.) and five wild cinnamon species grown in Sri Lanka," 2017.

[15] C. U. Kumari, S. J. Prasad, and G. Mounika, "Leaf disease detection: feature extraction with K-means clustering and classification with ANN," in 2019 3rd international conference on computing methodologies and communication (ICCMC), IEEE, 2019, pp. 1095–1098.

[16] A.Perdana Windarto and A. Wanto, "Data mining tools| rapidminer: K-means method on clustering of rice crops by province as efforts to stabilize food crops in Indonesia," in Materials Science and Engineering Conference Series, 2018, p. 012089.

[17] B. Goel and S. Mishra, "Medicinal and nutritional perspective of cinnamon: A mini-review," Eur. J. Med. Plants, vol. 31, pp. 10–16, 2020.

[18] R. Azad, K. L. W. Kumara, G. Senanayake, R. Ranawaka, D. Pushpakumara, and S. Geekiyanage, "Intensity of leaf spot and rough bark diseases in cinnamon accessions collected from major cinnamon growing areas of Sri Lanka," J Natl Sci Found, vol. 47, no. 3, pp. 297–305, 2019.

[19] M. Sibiya and M. Sumbwanyambe, "A computational procedure for the recognition and classification of maize leaf diseases out of healthy leaves using convolutional neural networks," Agri Engineering, vol. 1, no. 1, pp. 119–131, 2019.

[20] K. C. Kamal, Z. Yin, B. Li, B. Ma, and M. Wu, "Transfer learning for fine-grained crop disease classification based on leaf images," in 2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), IEEE, 2019, pp. 1–5.

Citation of this Article:

Lakshan S, Pathmajahn K, Sivasuthan S, Shashika Lokuliyanage, Rangi Liyanage, "Integrating Remote Sensing and Deep Learning for Precision Agriculture in Cinnamon Farming" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 8, pp 65-71, August 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.708009

****

