



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Plant Disease Identification Using Transfer Learning-Machine Learning

¹Prof.Chetan Padolest, ²Prof. Rohan Kokatend, ³Simran Charderd

¹Assistant Professor , IT and CSE Department JDCOEMst, ²HOD MCA Department JDCOEMnd, ³MCA departmentrd

¹ Assistant Professor , IT and CSE Department JDCOEMst,
¹JDCOEM , Nagpur, India

Abstract: Plant diseases pose a significant threat to global agricultural productivity, leading to major economic losses and food insecurity. Early detection of plant diseases is crucial for effective management and prevention, reducing the need for pesticides and minimizing crop damage. This research explores the application of transfer learning in machine learning for the identification and classification of plant diseases. Transfer learning is a technique that leverages pre-trained models on large-scale datasets and fine-tunes them on smaller, domain-specific datasets, thus overcoming the challenges of limited labelled data in plant disease classification. In this study, we employ deep learning models, specifically Convolutional Neural Networks (CNNs), which have demonstrated high accuracy in image recognition tasks. By using transfer learning with pre-trained models like Visual Geometry Group (VGGNET), we evaluate the models based on classification accuracy, sensitivity, and specificity, aiming to achieve high accuracy in detecting multiple diseases across different plant species. The study's findings indicate that our suggested approach in contrast to conventional feature extraction methods. A reliable and effective method for early disease detection in agriculture is offered by the combination of VGGNet, which significantly increases crop disease prediction precision, responsiveness, and discrimination. The MobileNet Random Forest Classifier (RFC) model's experimental results showed the maximum accuracy of 92.39%. By establishing the foundation for automated crop disease diagnosis, this study advances precision agriculture and eventually helps farmers make prompt, well-informed decisions to reduce crop losses. This dataset includes 13 classes of tomato, potato, and bell pepper diseases, ranging from common fungal infections to nutrient deficiencies.

I. INTRODUCTION

A reliable and effective method for early disease detection in agriculture is offered by the combination of VGGNet, which significantly increases crop disease prediction precision, responsiveness, and discrimination. The VGGNet Random Forest Classifier (RFC) model's experimental results showed the maximum accuracy of 92.39%. By establishing the foundation for automated crop disease diagnosis, this study advances precision agriculture and eventually helps farmers make prompt, well-informed decisions to reduce crop losses.

A predictive model is trained using the retrieved characteristics in the following stages of model training and testing. Testing data is used to evaluate the performance and generalisation abilities of the trained model. Lastly, the model is used to forecast future datasets, allowing for proactive approaches to disease control in agriculture. Deep learning researchers have started to emphasise transfer learning, which uses pre-trained models to solve problems even in the absence of training data. Two well-known CNN designs, VGGNet, are renowned for their effectiveness and efficiency in image classification. While VGGNet provides outstanding performance in collecting complex features through its inception modules, VGGNet excels in situations with constrained resources due to its lightweight design.

The potential of using VGGNet architectures as feature extractors to enhance crop disease prediction is examined in this work. By virtue of our objective in fine-tuning these pre-trained models is to apply the knowledge representations to enhance the characteristics of the classification model ability to discriminate among a variety of crop diseases pictures. The application of VGGNet's transfer learning and allows us to get over the obstacles of restricted computational resources and annotated data, thus enabling precise and effective disease diagnosis. Here, we offer a comprehensive evaluation of the recommended approach and contrast its effectiveness with baseline and traditional techniques for feature extraction. We assess the effectiveness of VGGNet in extracting relevant characteristics from images of crop diseases and examine how they affect the sensitivity of predictions, accuracy as well as specificity

II. LITERATURE REVIEW

Crop diseases disrupt the crop's normal state and have an impact on its physiological composition [1]. These deficiencies are caused by plant diseases that affect and sicken the host plants. These illnesses can be caused by parasitic nematodes, bacteria, viruses, or fungi and can affect any part of a plant, both above and below ground [2]. These diseases have the potential to change agricultural crops' physical, chemical, and biological composition, which can change how the afflicted plant component functions [3]. Consequently, the altered physiology of the farm plants reduces crop yield [4]. Seasonal fluctuations, environmental conditions, the presence of particular pathogens, and crop variety characteristics are some of the factors that influence disease incidences and their spread in farm crops [5-7]. Because of this, it is challenging to anticipate and recognise any disease attacks on agricultural crops. An agricultural crop known as cassava (*Manihot esculenta* Crantz) is highly vulnerable to a wide range of diseases [8], [9]. Cassava is one of the most important staple crops grown in Africa and a vital raw material for Asian and Latin American businesses [9]. Cassava is vegetatively grown by stem cuttings, which offers numerous advantages but also puts the crop's capacity to resist illness at risk because it can spread diseases easily and rapidly from one generation to the next [9]. Therefore, anticipating, recognising, and successfully preventing the various diseases that impact agricultural crops is essential. The most efficient method of diagnosing cassava infections has been found to be physically recognising the symptoms of the various conditions [9]–[11]. It is possible to recognise each disease by its own collection of symptoms. But for a large crop field, physical identification makes early detection very challenging. In order to encourage early disease detection and control, electronic technologies for illness identification are thus required. An established method for the computerised identification and classification of images is the convolutional neural network.

The most popular deep learning method for image analysis and classification is convolutional neural networks (CNNs), a type of artificial neural networks [12]. It is now the preferred method for computer vision tasks. CNN (or ConvNets) automatically learns and adapts planar hierarchical characteristics of images using backpropagation by utilising a large number of building pieces called fully linked convolution layers. Additional enhancements, such the pooling layers, are also present in these layers [13]. The abilities acquired from solving one difficulty with CNNs can be used to tackle other difficulties. It's called transfer learning.

Transfer learning techniques, which seek to adapt the knowledge acquired from data processing in foundational domains, enable predictive modelling, which is made up of different data patterns in another domain [14]. Among the many advantages of transfer learning are the potential to enhance neural network performance, time savings during network training, and a reduction in the quantity of training data needed [15]. The aim of this study is to improve the early diagnosis of cassava infections by using CNN transfer learning. Significant progress has been made in the autonomous identification of plant diseases by the use of deep learning techniques, specifically CNNs, in image recognition and classification [16].

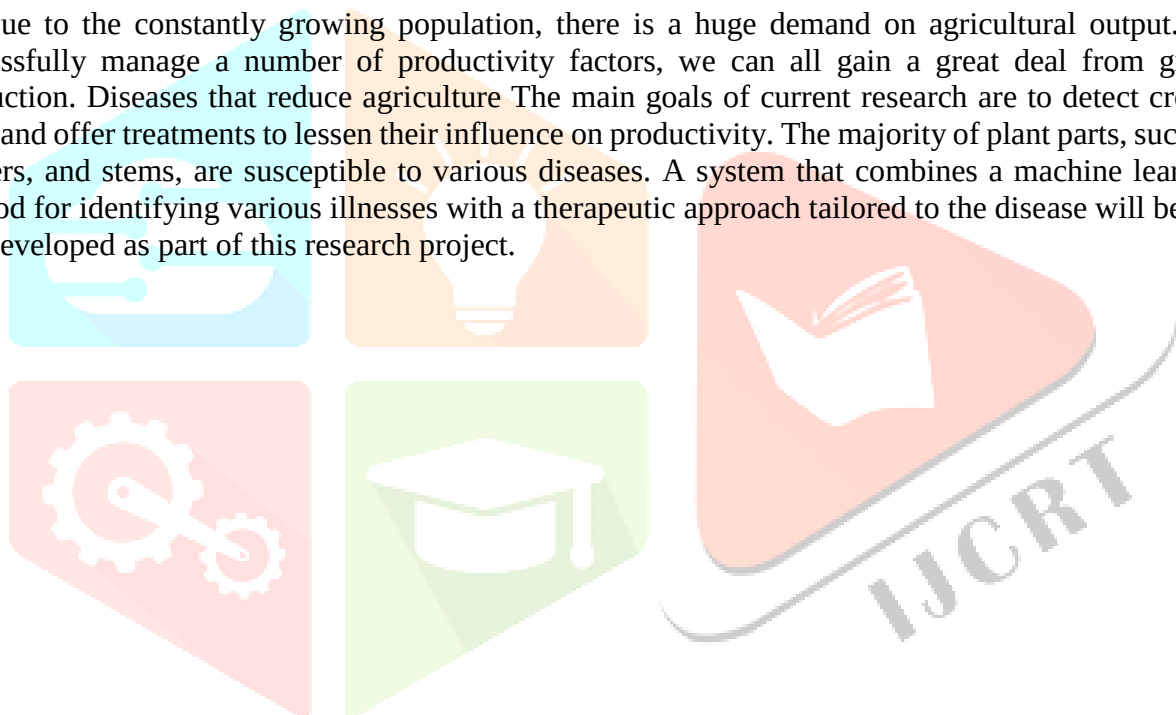
Several well-known CNN frameworks, including AlexNet, DenseNet, ResNet, Visual Geometry Group (VGG), MobileNet, Inception, and exception, have been used for picture identification and, more precisely, disease diagnosis [17]. In situations when processing resources were restricted and image resolution had to be reduced, deep convolutional neural networks were demonstrated to outperform Xception [18]. As part of the use of CNN for agricultural disease detection, a deep learning model for diagnosing cassava infections was developed on a mobile device [19]. However, there were discrepancies in the performance between the real-world videos and photos that the created model and application recorded.

A model for the identification and categorisation of cassava illnesses from an unbalanced dataset was also developed using convolutional neural networks [20]. However, the created model favoured some classes over others due to the data's extreme unbalance. Oversampling was used to balance the data, however there is a risk that the model will be overfit to the under-represented data class. Again, to enhance smart farming, a deep learning algorithm for detecting plant diseases was developed [21]. The region proposal network (RPN) was employed to overcome the leaves' backdrop as a research barrier, while the Chan-Vese (CV) approach was employed for feature detection.

The model performed admirably, although the feature extraction CV approach requires time-consuming, repetitive repeated calculations. Additionally, transfer learning has been used to detect crop diseases. The technique to identify and categorise cassava diseases from photos used CNN architectures (inception), support vector machines (SVM), and k closest neighbours (KNN) [22]. The CNN architecture yielded the best results out of the three machine learning techniques utilised with Inception, but all three achieved very high model performance.

III. PROPOSED MODEL

Due to the constantly growing population, there is a huge demand on agricultural output. If we can successfully manage a number of productivity factors, we can all gain a great deal from greater food production. Diseases that reduce agriculture The main goals of current research are to detect crop diseases early and offer treatments to lessen their influence on productivity. The majority of plant parts, such as leaves, flowers, and stems, are susceptible to various diseases. A system that combines a machine learning-based method for identifying various illnesses with a therapeutic approach tailored to the disease will be conceived and developed as part of this research project.



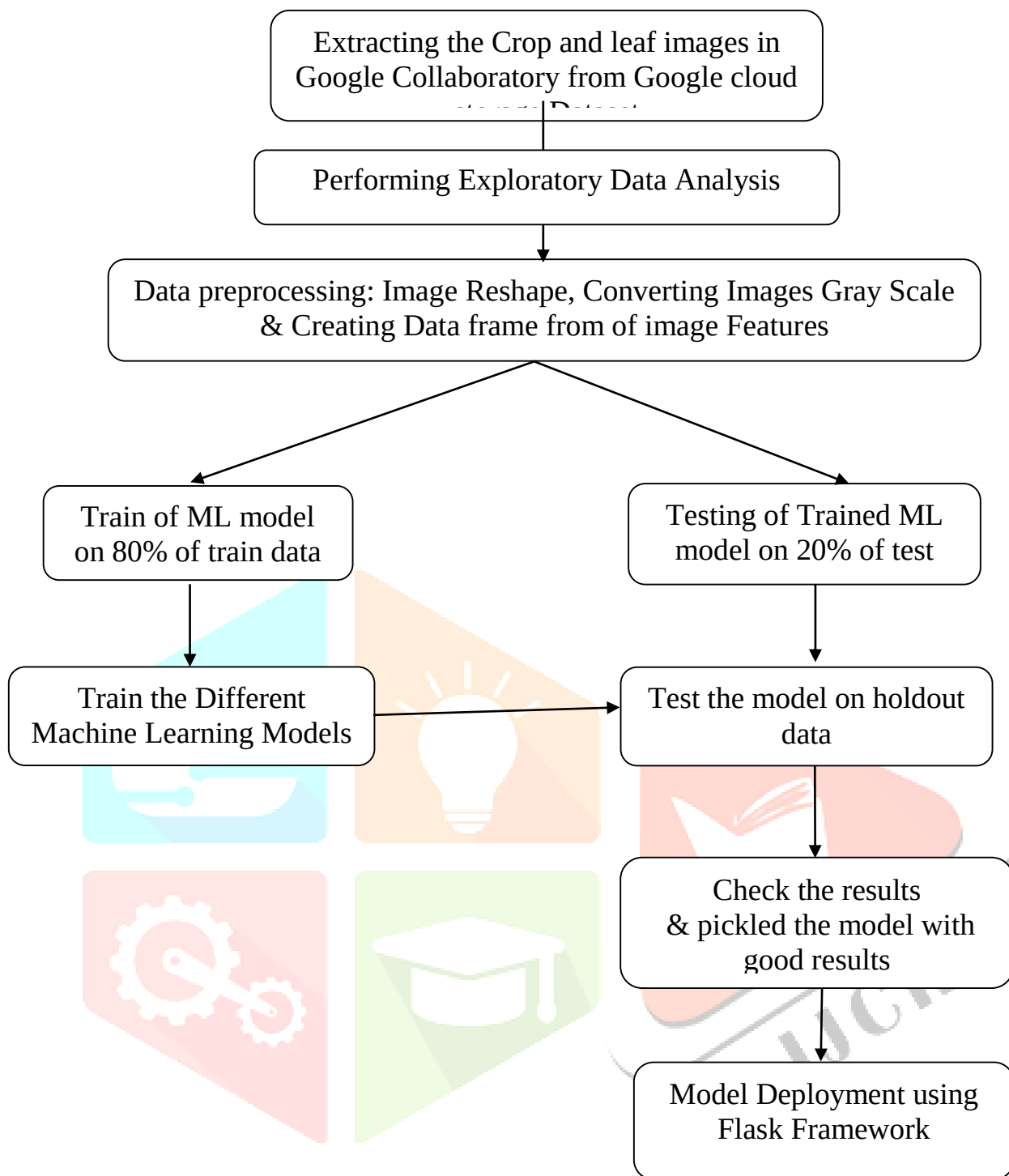


Figure 1: Proposed Methodology

As seen in Figure 2, the suggested approach predicts crop diseases using an organised manner. First, a dataset of images is taken out of cloud storage, then exploratory data analysis to learn more about the features of the dataset. methods for preparing data, such like resizing images, converting them to greyscale, and using the features of the image to create a data frame, are then used to get the data ready for additional examination. Important information is extracted from the pre-processed images. features that make use of VGG Net, which are extractors of features. Next, a testing set and a training set are produced using the dataset. VVG Net features are utilised to train and assess the Random Forest Classifier, Logistic Regression, Support Vector Machine, and KNN, respectively. This strategy provides a methodical approach to employing machine learning models and deep learning architectures for agricultural disease prediction, which will increase agricultural productivity and enable better informed decision-making.

3.1 Feature Extraction using VGG Net

VGGNet is a convolutional neural network (CNN) architecture known for its depth and performance in image classification tasks. It consists of several convolutional layers followed by fully connected layers, making it suitable for extracting rich feature representations from images. Images need to be preprocessed to match the input requirements of VGGNet, which includes resizing and normalization. Extract features from a specific layer in the VGGNet architecture. Common choices are the last convolutional layer or before the fully connected layers. After extracting the features, the extracted features for various tasks can be used as,

- **Classification:** Feed them into a different classifier.
- **Clustering:** Use them for clustering similar images.
- **Retrieval:** Compare features for image retrieval tasks.

This feature extraction technique is useful for tasks like transfer learning, where you can fine-tune the model on a new dataset or for image retrieval tasks where you compare image features. Using VGGNet for feature extraction provides a powerful way to leverage pretrained models for various computer vision tasks. By extracting high-level features, you can significantly enhance the performance of your applications while reducing the need for extensive labelled datasets.

IV.RESULT

Table 1: Algorithms and its accuracy with evaluation metrics

	Algorithm	Accuracy	Error rate	Precision	Recall	F1 score
0	K-nearest neighbors	75.88	24.11	77.24	75.88	75.16
1	Logistic Regression	59.18	40.81	60.39	59.18	58.11
2	Support Vector Machine	59.18	40.81	60.39	59.18	58.11
3	Random Forest Classifier	92.39	7.60	94.04	92.39	92.63

In the table above, you can see the results of several algorithms' accuracy tests, accuracy, precision, recall, F1 score, and error rate

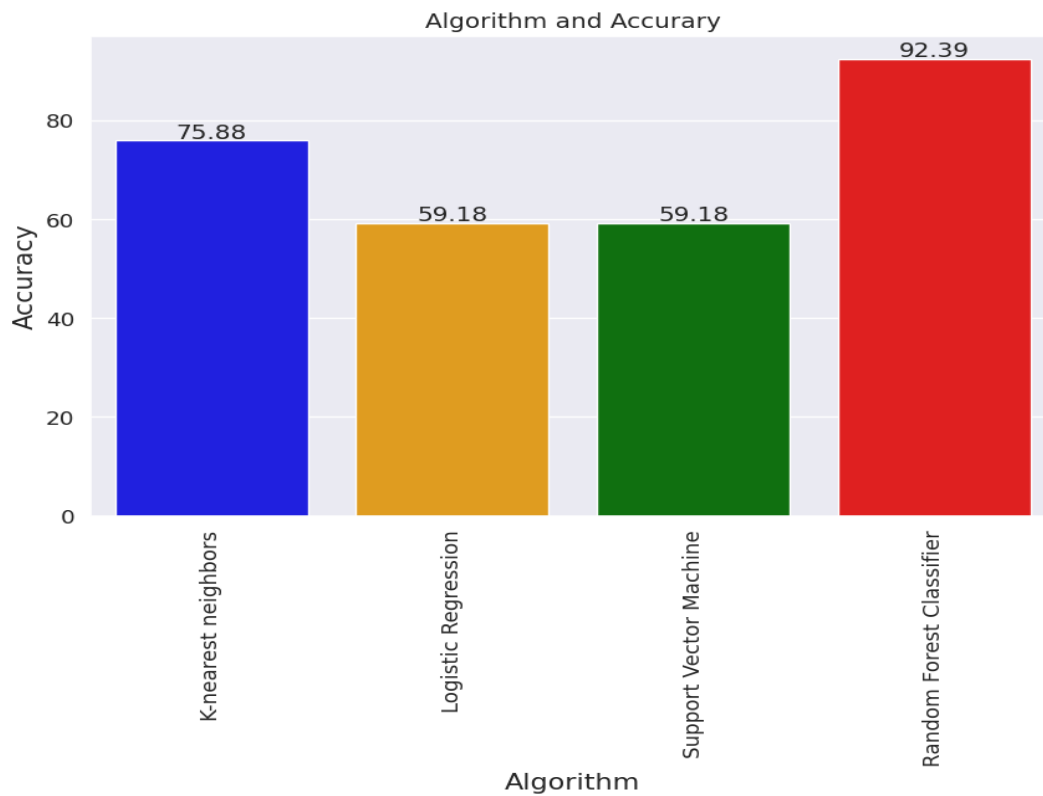


Figure 2: Graphical representation of algorithms and its accuracy

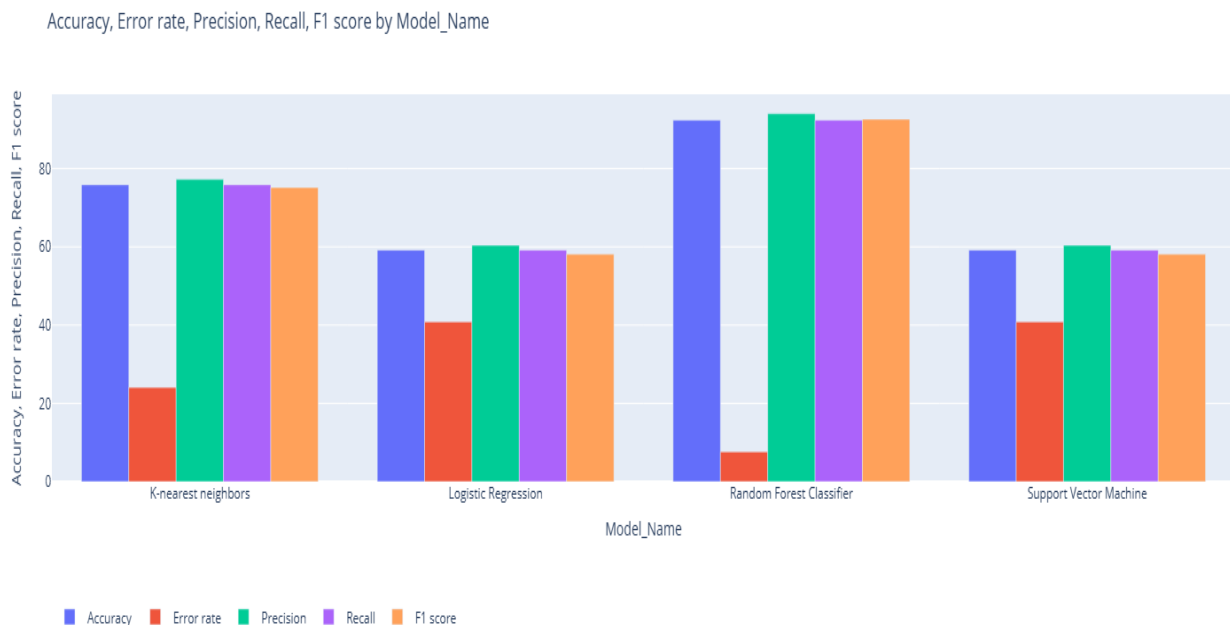


Figure 3: Graphical representation of Algorithms and its accuracy with evaluation metrics

Figure 2 and 3 displays a bar plot that contrasts how well different algorithms perform when combined with VGGNet feature extraction techniques. The bar plot above compares a number of performance metrics for various algorithms that use the VGGNet feature extraction methods, including accuracy, precision, recall, and F1 score. The y-axis shows the metrics that correspond to each algorithm, which is represented by a different bar. Out of all the methods tested, the VGGNet Random Forest Classifier (RFC) model had the best accuracy, 92.39%, as seen in the plot. The relative effectiveness of several algorithms and feature extraction techniques in crop disease prediction is made easier to understand.

V. CONCLUSION

Our research concludes by showing how well VGGNet structures work as feature extractors to enhance crop disease prediction. Transfer learning allowed us to refine these previously models on a collection of photos of crop diseases, utilising the improved categorisation using learnt representations performance. The recommended approach produces superior Predicting disease with precision, sensitivity, and specificity activities compared to baseline methods and traditional features extraction methods, based on the results of our experiments. The Random Forest Classifier (RFC) model developed by VGGNet obtained 92.39% accuracy, the highest of all algorithms assessed. Farmers may reduce losses, maximise resource allocation, and increase crop yield thanks to VGGNet's enhanced forecast accuracy, which enables prompt intervention and well-informed decision-making. In order to improve the scalability and usefulness of automated crop disease prediction in agricultural settings, future research can investigate the use of cutting-edge deep learning algorithms and integration with real-time monitoring systems.

VI. REFERENCES

1. M. C. Shurtleff, M. J. Pelczar, A. Kelman, and R. M. Pelczar, "Plant disease," Encyclopedia Britannica, 2020. [Online]. Available: <https://www.britannica.com/science/plant-disease>. (Accessed: Mar. 04, 2021).
2. L. K. T. Al-Ani and E. L. Furtado, "The effect of incompatible plant pathogens on the host plant," in *Molecular Aspects of Plant Beneficial Microbes in Agriculture*, Elsevier, 2020, pp. 47-57.
3. M. Donatelli, R. D. Magarey, S. Bregaglio, L. Willocquet, J. P. M. Whish, and S. Savary, "Modelling the impacts of pests and diseases on agricultural systems," *Agricultural Systems*, vol. 155, pp. 213-224, 2017, doi: 10.1016/j.agsy.2017.01.019.
4. R. Cerda, J. Avelino, C. Gary, P. Tixier, E. Lechevallier, and C. Allinne, "Primary and secondary yield losses caused by pests and diseases: assessment and modeling in coffee," *PLoS One*, vol. 12, no. 1, p. e0169133, 2017, doi: 10.1371/journal.pone.0169133.
5. A. C. Velásquez, C. D. M. Castroverde, and S. Y. He, "Plant-pathogen warfare under changing climate conditions," *Current Biology*, vol. 28, no. 10, pp. R619-R634, 2018, doi: 10.1016/j.cub.2018.03.054.
6. Y. Elad and I. Pertot, "Climate change impacts on plant pathogens and plant diseases," *Journal of Crop Improvement*, vol. 28, no. 1, pp. 99-139, 2014, doi: 10.1080/15427528.2014.865412.
7. J. F. Sundström et al., "Future threats to agricultural food production posed by environmental degradation, climate change, and animal and plant diseases - a risk analysis in three economic and climate settings," *Food Security*, vol. 6, no. 2, pp. 201-215, 2014, doi: 10.1007/s12571-014-0331-y.
8. G. Nakabonge, C. Samukoya, and Y. Baguma, "Local varieties of cassava: conservation, cultivation and use in Uganda," *Environment, Development and Sustainability*, vol. 20, no. 6, pp. 2427-2445, Dec. 2018, doi: 10.1007/s10668-017-9997-6.
9. J. P. Legg et al., "Cassava virus diseases: Biology, epidemiology, and management," in *Advances in Virus Research*, vol. 91, no. 1, Academic Press Inc., 2015, pp. 85-142.
10. V. Verdier, S. Ojeda, and G. Mosquera, "Methods for detecting the cassava bacterial blight pathogen: a practical approach for managing the disease," *Euphytica*, vol. 120, no. 1, pp. 103-107, 2001, doi: 10.1023/A:1017516007945.
11. J. Legg and S. Winter, "Cassava Mosaic Viruses (Geminiviridae)," in *Encyclopedia of Virology*, Elsevier, 2021, pp. 301-312.
12. P. Gopika et al., "Single-layer convolution neural network for cardiac disease classification using electrocardiogram signals," in *Deep Learning for Data Analytics*, Elsevier, 2020, pp. 21-35.
13. R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging*, vol. 9, no. 4, pp. 611-629, 2018, doi: 10.1007/s13244-018-0639-9.
14. J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang, "Transfer learning using computational intelligence: a survey," *Knowledge-Based Systems*, vol. 80, pp. 14-23, 2015, doi: 10.1016/j.knosys.2015.01.010.

15. N. Best, J. Ott, and E. J. Linstead, "Exploring the efficacy of transfer learning in mining image-based software artifacts," *Journal of Big Data*, vol. 7, no. 1, pp. 1-10, 2020, doi: 10.1186/s40537-020-00335-4.
16. J. Boulent, S. Foucher, J. Théau, and P. L. St-Charles, "Convolutional neural networks for the automatic identification of plant diseases," *Frontiers in Plant Science*, vol. 10, p. 941, 2019, doi: 10.3389/fpls.2019.00941.
17. M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant disease classification: a comparative evaluation of convolutional neural networks and deep learning optimizers," *Plants*, vol. 9, no. 10, pp. 1–17, 2020, doi: 10.3390/plants9101319.
18. H. F. Pardede et al., "Plant diseases detection with low resolution data using nested skip connections," *Journal of Big Data*, vol. 7, no. 1, pp. 1-21, 2020, doi: 10.1186/s40537-020-00332-7.
19. A. Ramcharan et al., "A mobile-based deep learning model for cassava disease diagnosis," *Frontiers in Plant Science*, vol. 10, p. 272, 2019, doi: 10.3389/fpls.2019.00272.
20. G. Sambasivam and G. D. Opiyo, "A predictive machine learning application in agriculture: cassava disease detection and classification with imbalanced dataset using convolutional neural networks," *Egyptian Informatics Journal*, vol. 22, no. 1, pp. 27- 34, 2020, doi: 10.1016/j.eij.2020.02.007.
21. Y. Guo et al., "Plant disease identification based on deep learning algorithm in smart farming," *Discrete Dynamics in Nature and Society*, vol. 2020, p. 2479172, 2020, doi: 10.1155/2020/2479172.
22. A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Front. Plant Sci.*, vol. 8, pp. 1852, 2017, doi: 10.3389/fpls.2017.01852.

