[1] Kunal Tyagi *
[2] Saksham Vats
[3] Dr Vasudha Vashisht

Implementing Inception v3, VGG-16 and VGG-19 Architectures of CNN for Medicinal Plant leaves Identification and Disease Detection

Journal of Electrical Systems

Abstract: - Medicinal plants have been a valuable source of healing and healthcare for centuries, serving as a crucial component of traditional medicine systems worldwide. The identification of these plants is a fundamental step in harnessing their therapeutic potential and ensuring their sustainable use. Accurate identification of these plants and the early detection of diseases affecting their leaves are crucial for ensuring a consistent supply of high-quality medicinal resources. In this paper, three types of CNN architectures of deep learning are used to develop robust classification models for distinguishing between healthy and diseased medicinal plant leaves. Inception v3 is known for its versatility in handling various input image sizes without the need for extensive pre-processing, while VGG-19 exhibits high accuracy, robust feature extraction, and suitability for complex image patterns. VGG-16, known for its simplicity and ease of training, provided competitive results, particularly where computational resources were limited. These networks have been pre-trained on large-scale image datasets and fine-tuned using them. These different approaches provide a varied solution to this problem by comparing their accuracy levels, feasibility and to select which one suits the best in this study.

Keywords: Artificial Intelligence, Convolutional Neural Network, Deep Learning, Machine Learning.

I. INTRODUCTION

In recent years, the fusion of technology and agriculture has opened new frontiers in the quest for efficient and sustainable farming practices. The application of deep learning, particularly *Convolutional Neural Networks* (CNNs), has revolutionized the way we approach challenges in plant biology and agriculture. Among the diverse range of CNN architectures available, Inception v3, VGG-16, and VGG-19 have emerged as powerful tools for image analysis and classification tasks. This study delves into the implementation and comparative evaluation of these three cutting-edge CNN architectures in the domain of medicinal plant leaves identification and disease detection, with a focus on contributing novel insights to the field.

Medicinal plants have played an essential role in human civilization for centuries, serving as sources of valuable compounds with medicinal properties. The sustainable cultivation and preservation of these plants are of paramount importance to both traditional and modern medicine. However, they are susceptible to a multitude of diseases and environmental stressors that can significantly impact their yield and quality.

Accurate and timely identification of medicinal plant species and the detection of diseases affecting their leaves are critical tasks in agricultural and botanical research. Traditional methods of plant identification and disease diagnosis are often labor-intensive, time-consuming, and prone to errors. In contrast, deep learning-based approaches, especially CNNs, offer a promising avenue for automating these tasks, thereby enhancing the precision and efficiency of medicinal plant management.

In traditional methodologies, the process of feature extraction is a manual endeavor, requiring human intervention to identify and define relevant features. However, in contrary, deep learning automatically extract features using kernels.

specifically tailored for the automatic extraction of features using convolutional filters.

Within the CNN framework, the process of feature extraction unfolds across multiple layers, with each layer serving a distinct role. In the lower layers, the network focuses on capturing low-level features, encompassing aspects such as [1] gradients, color variations, and key points within the input data. These low-level features are subsequently transformed into higher-level features, including edges, corners, and more intricate patterns, as the data progresses through the higher layers of the network.

In image classification tasks, the heart of the feature extraction process lies in the convolutional layers of the network, which employ convolved filters of varying sizes to discern and highlight relevant patterns and details within the input data. Additionally, pooling layers play a vital role in the process by facilitating dimensionality reduction, enhancing the computational efficiency of the network. Depending on the specific requirements of the task, pooling operations can take the form of average pooling or max pooling.

[1, 2, 3] Amity School of Engineering and Technology, Noida, Uttar Pradesh. [2] saksham.vats@s.amity.edu, [3] vvashisht@amity.edu
* Corresponding Author Email: kunal.tyagi1@s.amity.edu

In the final classification layer of the CNN, [3] the *softmax activation function* is commonly employed to assign probabilities to each potential class, ultimately leading to the classification of the input data. In essence, the transition from traditional feature extraction methods to deep learning, particularly through the utilization of CNNs, represents a significant advancement, as it automates the complex process of feature identification and extraction, streamlining tasks such as image classification while yielding accurate and efficient results.

II. Literature Review

In this section, we provide a brief overview of recent advancements in plant disease detection through the application of deep learning techniques and look into their work. One notable study by S. H. Lee et al. [4] successfully identified three types of apple diseases with a classification accuracy of 93% by employing an enhanced support vector machine (SVM). In this work, the authors also utilized the K-means clustering method to segment lesion regions. Furthermore, a combination of techniques, including the global color histogram (GCH), color coherence vector (CCV), local binary pattern (LBP), and completed local binary pattern (CLBP), was employed to extract both color and texture features from apple spots.

Previous research, as documented in references, Tan, L et al. [5] and Hassan et al. [6], investigated the utilization of isolated lesions and spots instead of considering the entire leaf as a whole. This approach is motivated by the distinct characteristics exhibited by different disease regions on the leaf. Embracing this method offers several advantages, including the capability to detect multiple diseases simultaneously on a single leaf and the potential for data enrichment by dividing the leaf image into various sub-images.

In a prior study by Ahmad W et al. [14], the methodology involved the utilization of Directional Local Quinary Patterns (DLQP) to determine keypoints within the input image during the initial stage. Subsequently, plant disease classification results were achieved by training a Support Vector Machine (SVM) classifier using the computed keypoints. In a separate study focused on the detection of tomato leaf diseases by Zhang et al. [15], various deep learning models, specifically AlexNet, GoogleNet, and ResNet, were employed. These models were subjected to experimentation with both the Stochastic Gradient Descent (SGD) and Adam optimizers. Remarkably, the ResNet model exhibited the highest level of accuracy, achieving a remarkable 97.28% accuracy rate.

A method was pioneered by researchers D. Chad et al. [7] to autonomously identify plant diseases in images of maize plants taken in real-field conditions. In the domain of rice disease identification, Lu et al. [9] introduced an approach centered on deep convolutional neural networks (CNNs). Additionally, Zhang et al. employed deep learning methods to construct a network capable of recognizing images of agricultural machinery [10]. In 2015, researchers, Yusuke Kawasaki et al. employed a 3-layer CNN-based learning system [25], which underwent training on a dataset comprising 800 cucumber leaf images. Their efforts yielded an impressive average accuracy rate of 94.9% following 40 epochs of training. Subsequently, in 2016, another study by Sharada P. Mohanty et al. [24] conducted an empirical investigation focusing on the AlexNet and GoogLeNet deep neural network (DNN) architectures. Their research resulted in a remarkable classification accuracy of 99.35% when applied to the extensive PlantVillage dataset, containing over 50,000 sample images, and this achievement was realized after 30 epochs of training.

These previous research outcomes provide a solid foundation for the current study, which seeks to extend this knowledge by implementing and comparing the Inception v3, VGG-16, and VGG-19 architectures in the specific context of medicinal plant leaves analysis. The insights gained from these prior studies underscore the potential of CNNs in automating and enhancing the accuracy of plant-related tasks, which is particularly relevant in the context of medicinal plant management and disease prevention.

III. Types Of Cnn Models

3.1 Inception V3

In contemporary computer vision, we have a highly efficient approach known as Transfer Learning. This method leverages pre-existing neural networks, particularly one developed by Google AI, to classify a broad range of visual objects with remarkable accuracy.

[19] Transfer learning is a machine learning technique that capitalizes on the use of a pre-trained neural network. In this context, the neural network model we are utilizing is Inception-v3. This model represents the integration of various concepts contributed by multiple researchers over several years, rooted in the seminal paper titled "Rethinking the Inception Architecture for Computer Vision" authored by Szegedy, et. al. Inception v2, also known as GoogLeNet, represents a significant milestone in this journey of advancement. It is the successor to Inception v1, which laid the foundation for the concept of "Inception modules." [13] These modules, characterized

by the use of multiple filter sizes (1x1, 3x3, 5x5), allowed the network to capture features at multiple scales within a single layer. [20] Inception v2, with its incorporation of batch normalization, increased depth, and efficient convolutions, represents a notable advancement in the realm of deep convolutional neural networks.

3.2 Vgg-16

VGG-16 stands as a prominent convolutional neural network (CNN) architecture, widely recognized for its excellence in the realm of computer vision. Distinguished by its distinctive characteristics, VGG-16 deviates from the conventional approach of inflating hyper-parameters, opting instead for a more streamlined design.

Central to VGG-16's architecture is its utilization of 3x3 convolution layers with a stride of 1, [12] complemented by same padding, in addition to 2x2 max-pooling layers with a stride of 2. This architectural choice remains consistent throughout the network, forming a coherent framework for feature extraction and representation. Towards the culmination of the architecture, VGG-16 incorporates two fully connected layers, culminating in a softmax activation function to yield the final output. The nomenclature "16" in VGG-16 signifies the presence of 16 layers within the network, each laden with trainable weights.

Pretrained VGG-16 models and implementations are readily available in popular deep learning frameworks like TensorFlow and PyTorch. This makes it easy for researchers and developers to use and experiment with VGG-16. This model is less prone to overfitting compared to some more complex architectures. This can be advantageous when dealing with smaller datasets or when you want a model that generalizes well.

It is not the most computationally efficient architecture, especially for deployment on resource-constrained devices. Newer architectures like ResNet, Inception etc have emerged to address these efficiency concerns. The choice of architecture depends on the specific requirements of your project and the available computational resources.

3.3 Vgg-19

VGG-19, short for the 19-layer VGG (Visual Geometry Group) convolutional neural network, is a deep learning architecture designed for image classification tasks. [16] Its primary purpose was to participate in the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) competition.

VGG-19 takes image inputs with a fixed size of 224x224 pixels. [5] It is characterized by its extensive use of 3x3 convolutional layers, each followed by a Rectified Linear Unit (ReLU) activation function. ReLU is a non-linear activation function that outputs the input if it's positive; otherwise, it produces zero. The use of 3x3 receptive fields allows VGG-19 to capture fine-grained features in images. One notable aspect of VGG-19 is that it maintains the spatial resolution of the input image by using a fixed stride of 1 pixel during convolution. This ensures that detailed information in the image is preserved. In addition to its convolutional layers, VGG-19 includes three fully-connected layers. [22] The first two fully connected layers each consist of 4096 nodes, and the final layer has 1000 nodes. The last layer aligns with the number of classes in the ImageNet dataset, which is 1000.

While it is computationally intensive and may not be as efficient as more recent architectures like ResNet or Inception, it serves as a valuable benchmark in the field of deep learning and computer vision.

IV. METHODOLOGY

In our study we have aimed at identifying leaf diseases from input images, the process involves several sequential steps.

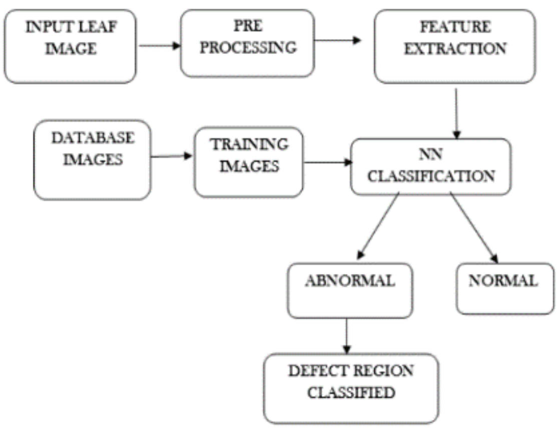

Fig. 1. CNN Methodology

4.1 Database Details

The image dataset utilized in this study comprises 13 distinct plant species, namely "Mango," "Potato," "Arjun," "Tomato," "Guava," "Basil," "Jamun," "Chinar," "Karanj," "Lemon," "Pepper Bell," "Jatropha," and "Pomegranate." Each of these plants holds significant economic and environmental value. The dataset is organized into two primary classes: "Healthy" and "Disease-Type," further subdivided into 27 subclasses. In total, the database encompasses 1,350 images, with 650 representing healthy leaf images and 700 depicting diseased leaf images. The image collection process involved images from Kaggle including other reputed sources. This dataset was reduced in size exponentially for the purpose of this study from the original dataset of 12,340 images.

4.2 Splitting The Dataset

Training Data: This is the portion of the dataset used to train the machine learning or deep learning model. The model learns patterns, features, and relationships in the training data.

Testing Data: This is a separate subset of the dataset that the model has never seen during training. It is used to evaluate the model's performance and its ability to generalize to new, unseen data A common split ratio is done, 70-30, where 70% of the data is used for training and the remaining 30% is used for testing. Randomization helps prevent any bias in the data order.

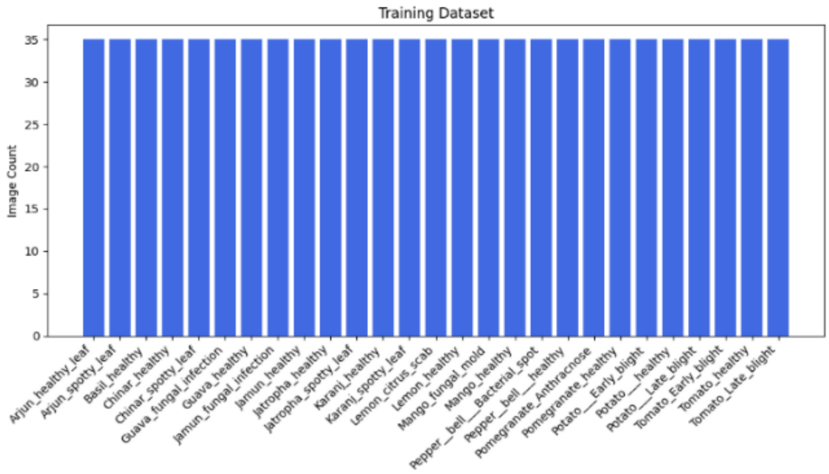

Fig. 2. Bar Graph – Training Dataset

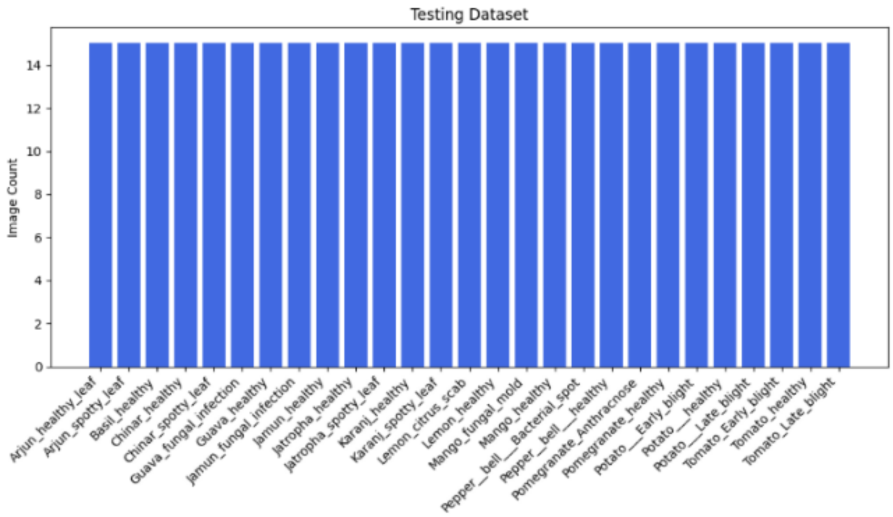

Fig. 3. Bar Graph – Testing Dataset

4.3 Image Data Generator

ImageDataGenerator is employed for data augmentation, expanding the training set and diversifying images to mitigate overfitting. It is a utility class provided by popular deep learning libraries such as TensorFlow's Keras and Keras standalone. It is commonly used for real-time data augmentation during the training of deep convolutional neural networks (CNNs) for computer vision tasks, such as image classification, object detection, and image segmentation. Dataset is also scaled down into size like 255 px in this study.

```
train_datagen=ImageDataGenerator(
    rescale=1./255,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True
    )
test_datagen=ImageDataGenerator(rescale=1./255)
```

Fig. 4. Working of ImageDataGenerator

V. Result And Discussion

Based on the provided accuracy metrics alone, Inception V3 is the most accurate choice for our dataset study. However, it's important to always thoroughly evaluate these models considering the factors mentioned above and potentially conduct further testing to make an informed decision.

Table I: Performance comparison between different CNN models

Model	**Epochs**	**Accuracy**	**Loss**
Inception v3	10	70.12 %	3.36
VGG-16	10	63.70 %	1.14
VGG-19	10	57.28 %	1.54

Inception v3 achieved the highest accuracy among the three models, indicating its superior ability to recognize and classify medicinal plant leaves. This may be attributed to its deeper architecture and the specific design of inception modules, which allow it to capture more complex features.

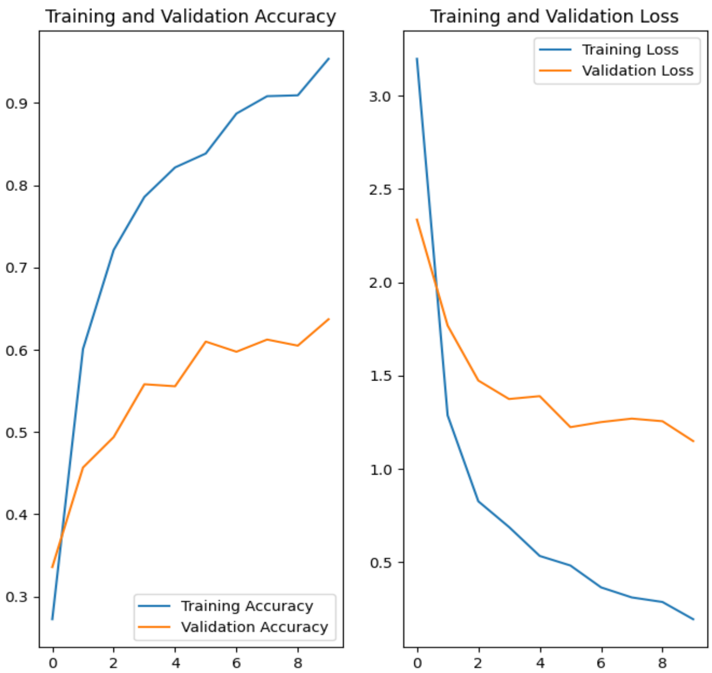

Fig. 5. Inception v3 Graph - Training and Validation Accuracy(left), Training and Validation Loss(right)

VGG-16, while showing a reasonable accuracy of 63.7%, falls slightly behind Inception V3. This suggests that the VGG-16 model might not be as effective at capturing the intricate details and variations present in medicinal plant leaves. It could still be useful for some applications, but it may not be the optimal choice for this specific task.

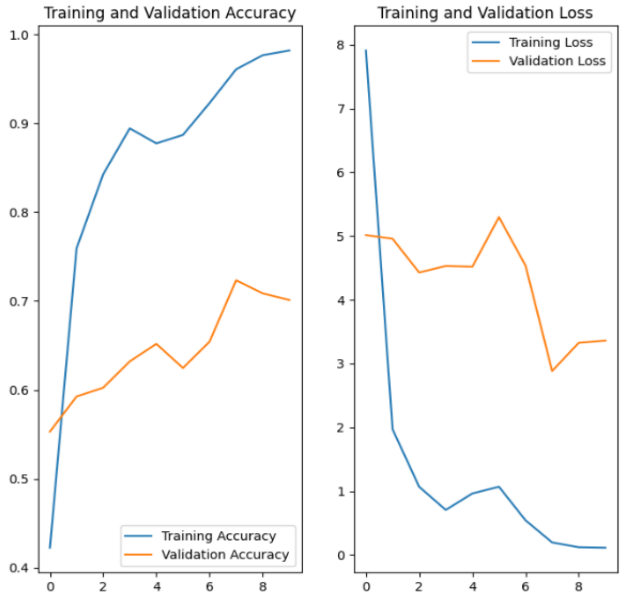

Fig. 6. VGG-16 Graph -Training and Validation Accuracy(left), Training and Validation Loss(right)

VGG-19 achieved an accuracy of 57.28%, which is the lowest among the three models. This outcome indicates that the additional layers in VGG-19 may not have contributed positively to the recognition of medicinal plant leaves. This lower accuracy might be due to over-parameterization, making it harder to generalize.

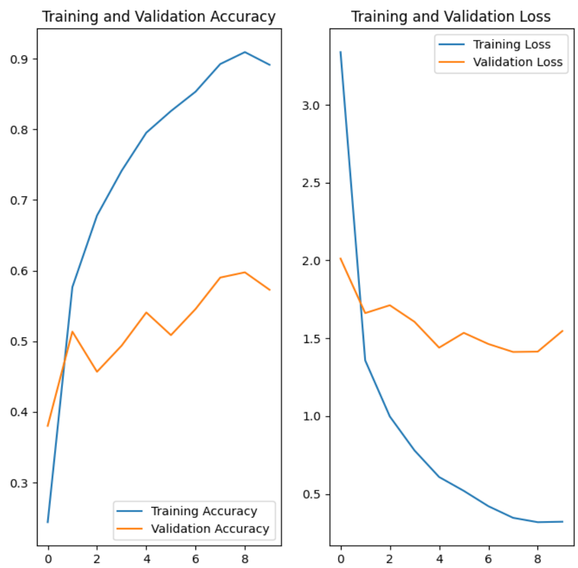

Fig. 7. VGG-19 Graph -Training and Validation Accuracy(left), Training and Validation Loss(right)

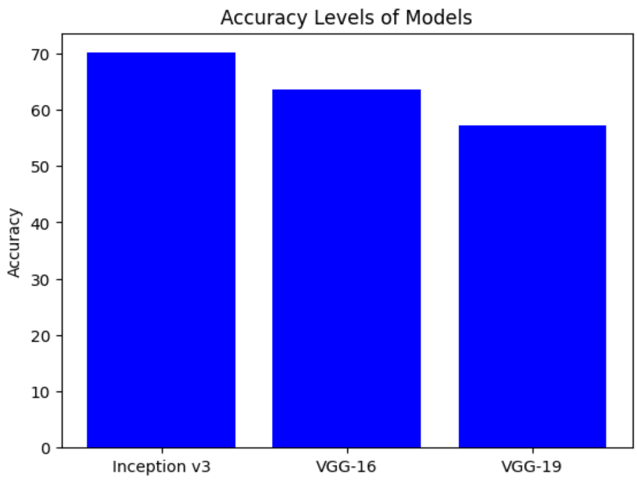

Fig. 8. Bar Graph – Value Accuracy of Models

Table II: Parameters comparison between different CNN models

Model	Total params	Trainable params
Inception v3	23185211 (88.44 MB)	1382427 (5.27 MB)
VGG-16	21144411 (80.66 MB)	6429723 (24.53 MB)
VGG-19	26454107 (100.91 MB)	6429723 (24.53 MB)

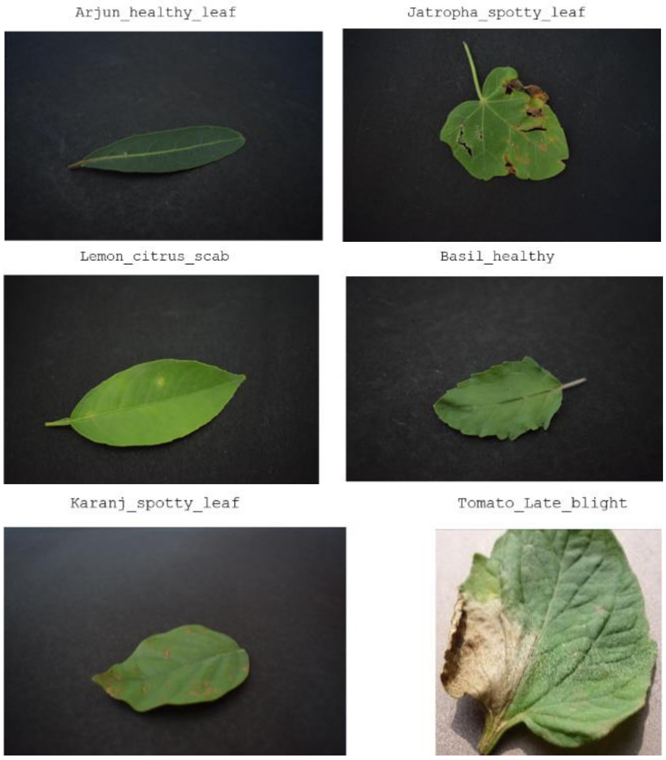

Fig. 9. Results collected on different leaves by VGG-16, VGG-19 and Inception v3 row wise

VI. Future Work

The implementation of Inception v3, VGG-16, and VGG-19, architectures for medicinal plant leaves identification and disease detection represents a significant step towards improving agricultural practices and plant conservation. However, there are several avenues for future research and development that can enhance the effectiveness and impact of this work: subjects and topics of current interest. Technical papers submitted for publication must advance the state of knowledge and must cite relevant prior work.

1) Expand the Dataset: The current dataset used in this study can be expanded to include a more extensive collection of medicinal plant species and a more diverse range of diseases. A larger and more varied dataset can further improve the model's ability to generalize to different plant species and disease types and improve accuracy exponentially. Future research can focus on fine-tuning the hyperparameters of the CNN architectures and exploring different optimization techniques. No of epochs can be increased to get better results, fine-tuning can lead to better model performance and faster convergence.
2) Ensemble methods, such as combining predictions from multiple models, can be explored to boost classification accuracy and robustness. Combining Inception v3, VGG-16, and VGG-19 models, or even different variants of these architectures, could be beneficial.
3) Transfer Learning from Pre-trained Models: Continue to explore the benefits of transfer learning by leveraging pre-trained models on even larger datasets. Fine-tuning these models for medicinal plant identification and disease detection can yield improvements in both accuracy and convergence speed.
4) Multi-Modal Data Fusion: Incorporate additional data modalities such as hyperspectral imaging, thermal imaging, or data from IoT sensors to provide a more comprehensive view of plant health. Fusion of multi-modal data can enhance disease detection accuracy and early warning systems. Develop real-time disease monitoring systems using CNN architectures. Integrating CNN models with drones or robots equipped with cameras can enable continuous monitoring of plantations, allowing for early disease detection and intervention.

5) Explainable AI (XAI): Investigate techniques for making CNN models more interpretable and explainable. This can help farmers and researchers understand why a particular diagnosis was made, increasing trust and adoption of the technology.

VII. CONCLUSION

The implementation and comparative evaluation of Inception v3, VGG-16, and VGG-19 architectures for medicinal plant leaves identification and disease detection mark a significant advancement in the domain of agriculture, botany, and computer vision. This study has illuminated the potential of cutting-edge deep learning techniques to revolutionize the way we manage medicinal plant resources and combat diseases that threaten crop yields and biodiversity. The findings of this research highlight the strengths and capabilities of each CNN architecture. Inception v3, with its efficient use of parameters and multi-scale feature extraction, demonstrates its prowess in capturing intricate details of medicinal plant leaves and their diseases. VGG-16, with its simplicity and depth, remains a formidable choice for feature learning and classification tasks. VGG-19 known for its simplicity and strong performance in image classification tasks. VGG-19's straightforward design and pre-trained models make it valuable for various computer vision applications. Moreover, the study underscores the importance of data preparation, with a meticulously curated dataset containing a diverse array of plant species and diseases. The utilization of data augmentation techniques and transfer learning from pre-trained models amplifies the models' performance and generalization capabilities.

Ultimately, the implementation of Inception v3, VGG-16, and VGG-19 for medicinal plant leaves identification and disease detection is a testament to the transformative potential of deep learning in preserving botanical diversity, enhancing agricultural sustainability, and safeguarding the invaluable medicinal resources that contribute to human health and well-being. As we venture forward, the fusion of technology and botany promises to yield greener, healthier, and more prosperous tomorrows for our planet and its inhabitants.

REFERENCES

[1] Amara, J., Bouaziz, B., Algergawy, A., et al.: A deep learning-based approach for banana leaf diseases classification. In: BTW (Workshops), pp. 79–88 (2017).

[2] Ferentinos, K.P.: Deep learning models for plant disease detection and diagnosis. Comput. Electr. Agri. 145, 311–318 (2018)

[3] J. G. A. Barbedo, "Plant disease identification from individual lesions and spots using deep learning", Biosyst. Eng., vol. 180, pp. 96-107, Apr. 2019.

[4] S. H. Lee, H. Goëau, P. Bonnet and A. Joly, "New perspectives on plant disease characterization based on deep learning", Comput. Electron. Agricult., vol. 170, Mar. 2020.

[5] Lu, J.; Tan, L.; Jiang, H. Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification. Agriculture 2021, 11, 707.

[6] Hassan, S.M.; Maji, A.K.; Jasiński, M.; Leonowicz, Z.; Jasińska, E. Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach. Electronics 2021

[7] D. Chad, W. H. Tyr, S. Chen et al., "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," Phytopathology, vol. 107, pp. 1426–1432, 2017.

[8] C. Ni, D. Wang, R. Vinson, M. Holmes, and Y. Tao, "Automatic inspection machine for maize kernels based on deep convolutional neural networks," Biosystems Engineering, vol. 178, pp. 131–144, 2019.

[9] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," Neurocomputing, vol. 267, pp. 378–384, 2017.

[10] Z. Zhang, H. Liu, Z. Meng, and J. Chen, "Deep learning-based automatic recognition network of agricultural machinery images," Computers and Electronics in Agriculture, vol. 166, p. 104978, 2019.

[11] Jiang, P., Chen, Y., Liu, B., He, D., Liang, C.: Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access 7, 59069–59080 (2019)

[12] Geetharamani, G., Pandian, A.: Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Comput. Electr. Eng. 76, 323–338 (2019)

[13] J. Hu, Z. Chen, M. Yang, R. Zhang and Y. Cui, "A Multiscale Fusion Convolutional Neural Network for Plant Leaf Recognition," in IEEE Signal Processing Letters, vol. 25, no. 6, pp. 853-857, June 2018,

[14] Ahmad W, Shah S, Irtaza A (2020) Plants disease phenotyping using quinary patterns as texture descriptor. KSII Trans Internet Inf Syst 14(8):3312–3322.

[15] Zhang, Keke & Wu, Qiufeng & Liu, Anwang & Meng, Xiangyan. (2018). Can Deep Learning Identify Tomato Leaf Disease?. Advances in Multimedia. 2018. 1-10. 10.1155/2018/6710865.

[16] Gajjar, R., Gajjar, N., Thakor, V.J. et al. Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform. Vis Comput 38, 2923–2938 (2022).

[17] I. Ahmad, M. Hamid and S. Yousaf, "Optimizing Pretrained Convolutional Neural Networks for Tomato Leaf Disease Detection,"

[18] H. F. Pardede, E. Suryawati, D. Krisnandi, R. S. Yuwana, and V. "Machine learning based plant diseases detection: A review," in 2020 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET),

[19] Andrew L. Maas, Awni Y. Hannun, and Andrew Y. Ng. 2013. Rectifier Nonlinearities Improve Neural Network Acoustic Models. ICML (2013).

[20] A. A. Joshi and B. D. Jadhav, “Monitoring and Controlling Rice Diseases Using Image Processing Techniques,” 2016

[21] K. Renugambal and B. Senthilraja, “Application of image processing technique in plant disease recognition,” International journal of engineering research and technology, vol. 4, no. 03, Mar. 2015.

[22] Sun, Yu, Yuan Liu, Guan Wang, and Haiyan Zhang, “Deep learning for plant identification in natural environment,”Computational intelligence and neuroscience, vol. 2017,

[23] Aston Zhang, Zachary C. Lipton, Mu Li, and Alexander J. Smola. 2020. Dive into Deep Learning. Retrieved August 15, 2020 from https://d2l.ai/d2l-en.pdf

[24] Sharada P. Mohanty, David P. Hughes, and Marcel Salathé. 2016. Using Deep Learning for Image-Based Plant Disease Detection. Front Plant Sci 7, (2016), 1419.

[25] Yusuke Kawasaki, Hiroyuki Uga, Satoshi Kagiwada, and Hitoshi Iyatomi. 2015. Advances in Visual Computing, 11th International Symposium, ISVC 2015, Las Vegas, NV, USA, December 14-16, 2015, Proceedings, Part II. Lect Notes Comput Sc (2015), 638–64