Deep Learning-Based Potato Leaf Disease Classification

Smrithi G S
Department of Computer Science
Vellore Institute Of Technology
Chennai, India
smrithy.gs@vit.ac.in

Dogiparthi Aasrith
Department of Computer Science
Vellore Institute Of Technology
Chennai, India
dogiparthi.aasrith2021@vitstudent.ac.in

Yarraguntla Kalyan Chakravarthy
Department of Computer Science
Vellore Institute Of Technology
Chennai, India
kalyan.chakravarthy2021@vitstudent.ac.in

Hanuma Vihari G
Department of Computer Science
Vellore Institute Of Technology
Chennai, India
hanumavihari.g2021@vitstudent.ac.in

***Abstract*— This project explores the automated classification of potato diseases using deep learning and convolutional neural network (CNN) models. The study employs a diverse dataset comprising images of healthy potatoes and those afflicted with common diseases such as late blight and early blight. Through meticulous preprocessing and fine-tuning of pre-trained CNN models, including Inception V3, ResNet-50, VGG-19, and MobileNet, the research aims to discern the most effective model for accurate disease classification. Remarkably, after extensive evaluation, ResNet-50 emerges as the top performer, achieving an outstanding accuracy rate of 99.84%. This exceptional accuracy underscores ResNet-50's ability to distinguish subtle differences between healthy and diseased potato plants, highlighting its potential for precise disease diagnosis in agriculture. The findings hold promise for revolutionizing disease management practices, offering farmers efficient tools to mitigate yield losses and ensure food security. By leveraging deep learning techniques, this research contributes to the development of real-time systems for on-field disease detection, enabling timely interventions and targeted agricultural practices. Moreover, the study sets a precedent for employing advanced technology in agricultural sectors, emphasizing the importance of machine learning in addressing pressing challenges faced by farmers worldwide. Ultimately, the successful implementation of deep learning models in potato disease classification signifies a significant step forward in enhancing agricultural productivity and sustainability, paving the way for a more resilient and efficient food production system in the face of evolving environmental and economic pressures.**

Keywords— Deep learning, Convolutional neural networks, Mobile Net, Inception V3, ResNet50, VGG19

I. Introduction

Potato is one of the most important food crops globally, serving as a staple food for millions and a crucial source of nutrition and income for farmers. However, potato cultivation faces formidable challenges, chief among them being the prevalence of various diseases that can significantly impact yield and quality. Among these diseases, late blight and early blight stand out as major threats, capable of causing substantial economic losses if not effectively managed. Additionally, accurately distinguishing between healthy potato plants and those afflicted with diseases is essential for implementing timely and targeted interventions to minimize yield losses and ensure food security.

Traditional methods of disease identification in potatoes rely heavily on visual inspection by experts, which can be time-consuming, labor-intensive, and prone to human error. Moreover, the efficacy of such methods may vary depending on the expertise of the observer and the stage of disease progression. In recent years, the advent of artificial intelligence (AI) and deep learning (DL) techniques, particularly deep learning and convolutional neural networks (CNNs), has offered promising avenues for automating disease diagnosis and classification in agriculture. By leveraging vast amounts of data and computational power, these techniques have demonstrated remarkable capabilities in accurately recognizing patterns and discerning subtle differences in complex datasets, including images of plant diseases.

In this context, this research aims to harness the power of deep learning and CNN models to develop an automated system for the classification of potato diseases, with a specific focus on late blight, early blight, and healthy plants. By analyzing high-resolution images of diseased and healthy potato plants, we seek to train and evaluate multiple CNN architectures, including Inception V3, ResNet-50, VGG-19, and MobileNet, to determine the most effective model for accurate disease classification. The ultimate goal of this project is to provide farmers and agricultural stakeholders with a reliable and efficient tool for early detection and management of potato diseases, thereby contributing to enhanced productivity, sustainability, and food security in potato cultivation systems globally.

II. LITERATURE REVIEW

Potato blight diseases, including early blight and late blight, pose significant challenges to potato cultivation worldwide. Early detection and accurate classification of these diseases are crucial for minimizing crop losses and ensuring food security. Various approaches have been proposed to address this issue, with recent advancements focusing on the application of deep learning and convolutional neural networks (CNNs) for automated disease classification [1,2].

One notable study proposes an end-to-end approach using deep learning to classify potato blight diseases, aiming to develop a user-friendly web-based application and mobile app for farmers. The system utilizes CNNs to process uploaded images of potato leaves, achieving high accuracy in distinguishing between early blight, late blight, and healthy plants. The proposed method integrates data collection, model building, MLOps, backend server development, and deployment on Google Cloud Platform (GCP), demonstrating the feasibility of using deep learning for practical agricultural solutions [3,4].

In a similar vein, another study proposes a multi-level deep learning model for recognizing potato leaf diseases, achieving impressive accuracy of 99.75% in distinguishing between healthy and diseased leaves. This model utilizes a unique CNN architecture coupled with ResNet50 image segmentation for accurate disease detection. The research outperforms existing models in terms of accuracy and computational cost, highlighting the effectiveness of deep learning in potato disease classification [5,6].

Furthermore, several studies emphasize the importance of deep learning-based methods, such as CNNs, for accurately identifying and categorizing potato diseases. By training on extensive datasets of potato plant photos, deep learning models can precisely classify various diseases, aiding in early disease detection, prevention, and crop management. Notably, some studies report accuracy rates exceeding 99% in potato disease classification tasks, showcasing the potential of deep learning algorithms in revolutionizing agricultural practices [7,8].

Moreover, the literature underscores the significance of utilizing deep learning techniques to address the challenges associated with potato diseases. By leveraging CNNs and transfer learning, researchers have achieved remarkable accuracy in classifying diseases based on leaf conditions, surpassing traditional classification methods. Some experiments report average accuracies of 91% to 95.36% using deep neural network approaches, affirming the effectiveness of these methods in potato disease detection and classification [9,10].

Additionally, studies highlight the need for early disease detection in potato cultivation to mitigate the adverse effects of diseases on crop yield and quality. Deep learning-based approaches offer a promising solution by enabling prompt responses and supporting sustainable agriculture. By accurately identifying and categorizing diseases, deep learning models contribute to increased agricultural productivity and plant health, addressing the challenges faced by farmers worldwide [11,12].

In conclusion, the literature survey underscores the significant role of deep learning and CNNs in revolutionizing potato disease classification and detection. With advancements in technology and methodology, researchers have achieved impressive accuracy rates exceeding 99% in distinguishing between healthy and diseased potato plants. These findings highlight the potential of deep learning-based approaches in improving agricultural practices and ensuring food security in potato cultivation [13,14,15].

III. PROPOSED METHODOLOGY

A. *Dataset Description*

Potato leaf disease detection in an early stage is challenging because of variations in crop species, crop diseases symptoms and environmental factors. These factors make it difficult to detect potato leaf diseases in the early stage. Various machine learning techniques have been developed to detect potato leaf diseases. However, the existing methods cannot detect crop species and crop diseases in general because these models are trained and tested on images of plant leaves of a specific region. In this research, a multi-level deep learning model for potato leaf disease recognition has developed. At the first level, it extracts the potato leaves from the potato plant image using the YOLOv5 image segmentation technique. At the second level, a novel deep learning technique has been developed using a convolutional neural network to detect the early blight and late blight potato diseases from potato leaf images. The proposed potato leaf disease detection model was trained and tested on a potato leaf disease dataset. The potato leaf disease dataset contains 4062 images collected from the Central Punjab region of Pakistan.

Split	**Leaf condition**	**No.of images**
Train	Early blight	1302
	Late blight	1139
	Healthy	816
Test	Early blight	326
	Late blight	285
	Healthy	204
Total		4,062

B. *Data Preprocessing*

Gather a diverse dataset of potato leaf images representing different stages of health and various disease conditions. This dataset should include high-quality images captured under different lighting conditions and angles. Remove any irrelevant or noisy images from the dataset. Ensure that the images are properly labeled with the corresponding disease categories (e.g., late blight, early blight, healthy). Resize all the images to a uniform size to ensure consistency across the dataset. Common sizes for input images in deep learning models include 224x224 or 299x299 pixels. Additionally, standardize the pixel values of the images to a common scale. Augment the dataset to increase its size and diversity. Common augmentation techniques include random rotations, flips, shifts, and changes in brightness and contrast. Data augmentation helps improve the model's ability to generalize to unseen data and reduces overfitting. Normalize the pixel values of the images to improve convergence during training. This typically involves scaling the pixel values to have zero mean and unit variance across the dataset. Divide the dataset into training, validation, and testing sets. The training set is used to train the model, the validation set is used to tune hyperparameters and monitor the model's performance during training, and the testing set is used to evaluate the final performance of the trained model. Implement a data loader to efficiently load and batch the preprocessed images during training. This helps optimize memory usage and training speed, especially when working with large datasets.

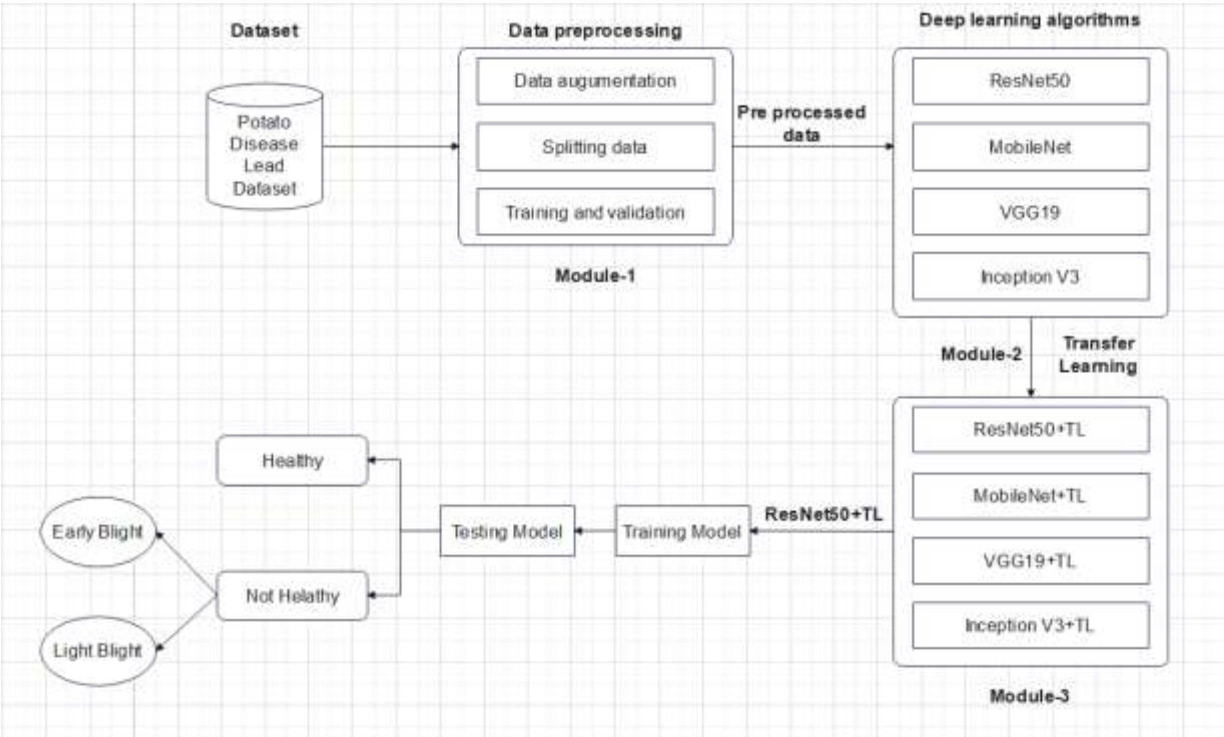

Figure.1 Complete Architecture Diagram

C. Deep Learning Algorithms

VGG19:

VGG19 is convolution neural community architecture renowned for its simplicity and effectiveness in image class duties. It turned into developed via Visual Geometry Group on the oxford college. VGG 19 has total of 19 layers that consists of sixteen convolution layers and 3 fully connected layers as mentioned in Fugure-2. Each convolutional layer is prepared with a 3x3 clear out, with a stride of one and padding to preserve spatial decision. VGG19 network architecture follows an uncomplicated pattern. In this the convolutional layers are stacked with rectified linear unit (ReLu) activations, accompanied with the aid of max pooling layers to lessen spatial dimensions. VGG19's deep architecture helps hierarchical characteristic extraction, permitting it to parent complicated patterns inside pics. VGG19 uniform structure make contributions massive adoption and serve as a benchmark for comparing the overall performance of extra architectures.

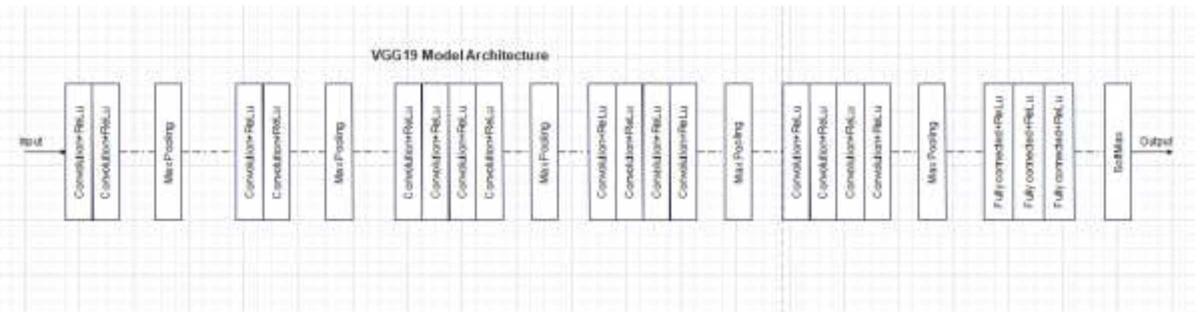

Figure.2 VGG19 Model Architecture

ResNet50:

ResNet50, short for Residual Network with 50 layers, represents a ground breaking architecture in deep studying, especially within the realm of photograph recognition and laptop imaginative and prescient responsibilities. Developed by using Microsoft Research, ResNet50 sticks out for its inventive use of residual connections, which deal with the vanishing gradient hassle encountered in education very deep neural networks. The architecture consists of 50 layers, along with convolutional layers, pooling layers, and absolutely connected layers as seen in below architecture diagram Figure-3. Central to its layout are residual blocks, where in shortcut connections allow gradients to go with the flow more directly during schooling. This architecture enables the development of deeper networks while retaining practicable complexity and avoiding degradation in accuracy. ResNet50 has validated especially powerful in diverse applications, consisting of picture type, item detection, and photo segmentation. The capacity to extract tricky functions from photographs have made ResNet50 a staple in the deep studying network.

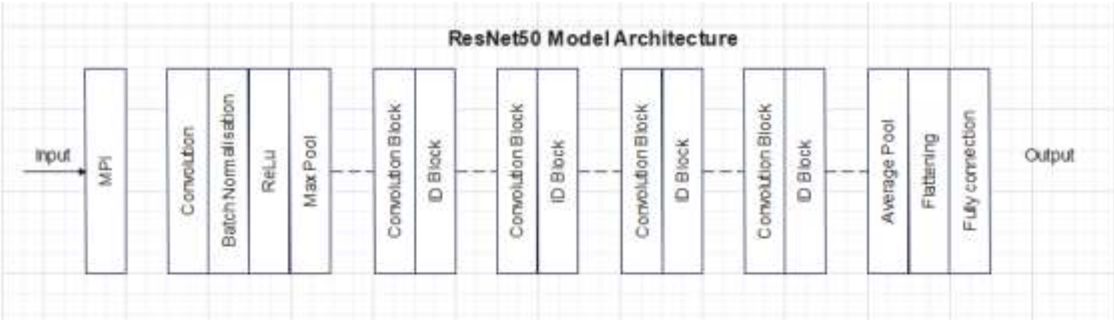

Figure.3 ResNet50 Model Architecture

Inception V3:

InceptionV3 represents an enormous development in convolutional neural network (CNN) architectures, especially designed for image classification and reputation tasks. Developed via Google researchers, InceptionV3 is renowned for its innovative use of inception modules, which permit the network to seize and technique features at more than one spatial scale. The structure features a deep network with meticulously crafted modules that include various kernel sizes, permitting the model to extract each nice and coarse-grained feature from input pics as mentioned in the below Figure-4. Furthermore, the structure employs auxiliary classifiers at some stage in training to mitigate the vanishing gradient problem. InceptionV3 has tested top notch overall performance in photo category competitions and actual-global packages, showcasing its capacity to address numerous datasets and complicated visible tasks efficaciously. Its versatility, performance, and superior overall performance have made InceptionV3 as a great one within the field of deep gaining knowledge of for computer vision.

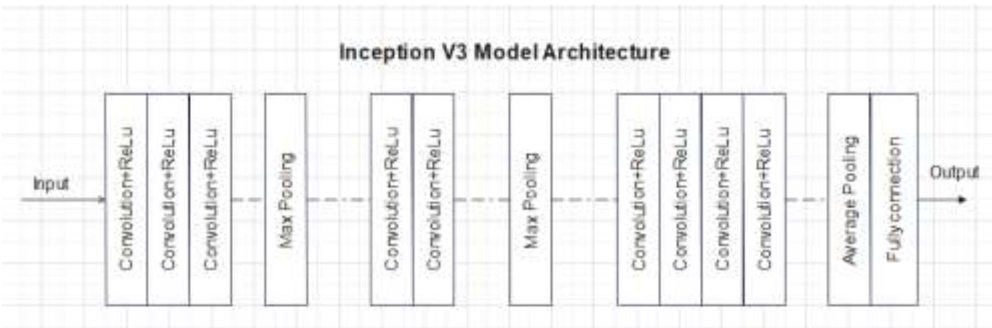

Figure.4 InceptionV3 Model Architecture

MobileNet:

MobileNet is a pioneering deep learning model designed specifically for mobile and embedded devices, catering to the increasing demand for efficient and lightweight neural networks. Developed by Google Research, MobileNet employs depth wise separable convolutions, which significantly reduce computational complexity while preserving the model's accuracy. This architectural innovation allows MobileNet to achieve high performance on tasks such as image classification, object detection, and semantic segmentation, all while maintaining a small memory footprint and fast inference speed. MobileNet's versatility and efficiency make it a popular choice for a wide range of real-world applications, including mobile apps, edge computing devices, and IoT devices, where resource constraints are prevalent. Its ability to strike a balance between accuracy and efficiency has cemented MobileNet's position as a cornerstone in the field of deep learning for mobile and embedded systems.The layers about the architecture is mentioned and drawn in the Figure-5.

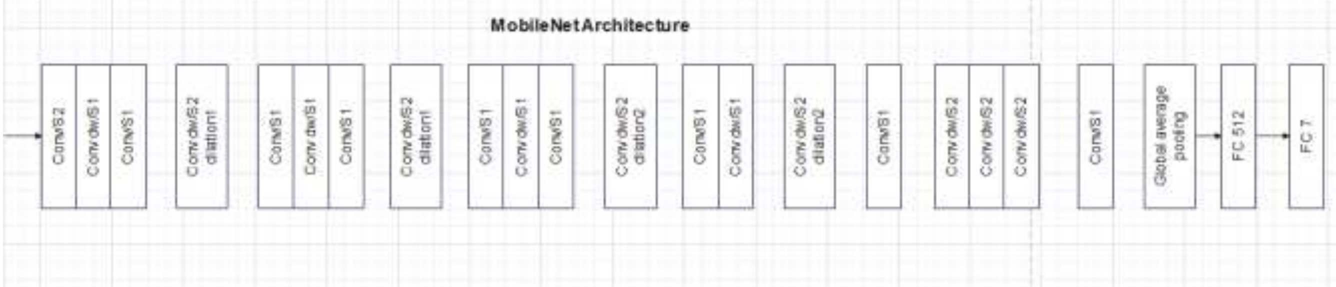

Figure.5 MobileNet Model Architecture

D. Transfer Learning

VGG19:

In this changed model of the VGG19 structure, the top layer has been eliminated, and new layers inclusive of a dense layer, batch normalization, and flatten layer have been integrated as we drawn in Figure-6. By removing the top layer, the structure turns into greater adaptable for numerous obligations which include characteristic extraction or switch getting to know. Batch normalization layers assist stabilize and boost up the training manner through normalizing the center of each layer to have 0 suggest and unit variance. The addition of a dense layer enables the network to learn complex relationships between functions extracted from the previous layers, enhancing its ability to categories or predict results based on the enter data. The inclusion of a flatten layer reshapes the output from the previous layers right into a one-dimensional array, preparing it for enter into the dense layer. These adjustments increase the flexibility and performance of the VGG19 structure, making it greater appropriate for a broader range of applications and enhancing its basic effectiveness in deep gaining knowledge of responsibilities.

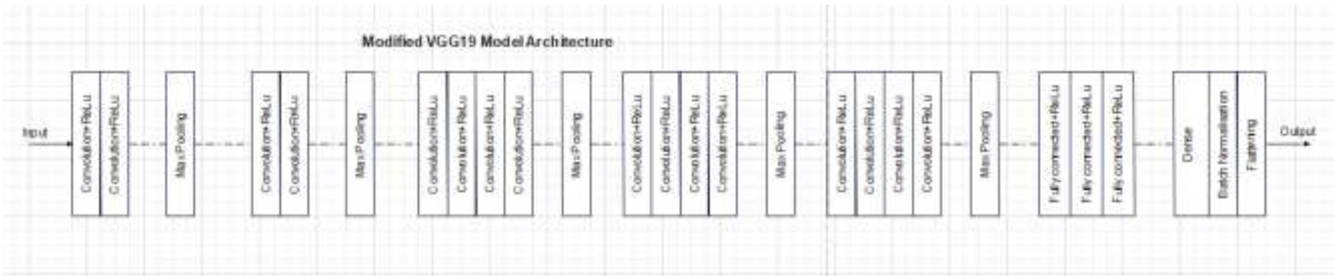

Figure.6 Modified VGG19 Model Architecture

ResNet50:

In this modified version of the ResNet50 architecture, the top layer has been eliminated and replaced with additional layers to decorate its talents. Two layers of batch normalization had been added, serving to normalize the activations of the community, which aids in stabilizing and accelerating the education manner. Following the batch normalization layers, a Dense-ReLU layer mixture has been added, in which the Dense layer helps the mastering of complex relationships in the facts, at the same time as the ReLU activation feature introduces non-linearity, allowing the community to version greater intricate styles. Finally, a Dense-SoftMax layer is appended to provide the very last output probabilities across multiple classes, leveraging the SoftMax activation characteristic to convert uncooked rankings into possibility distributions as mentioned in the below Figure-7. These modifications augment the ResNet50 structure, enhancing its adaptability and performance in various responsibilities consisting of photo type, item detection, and picture segmentation.

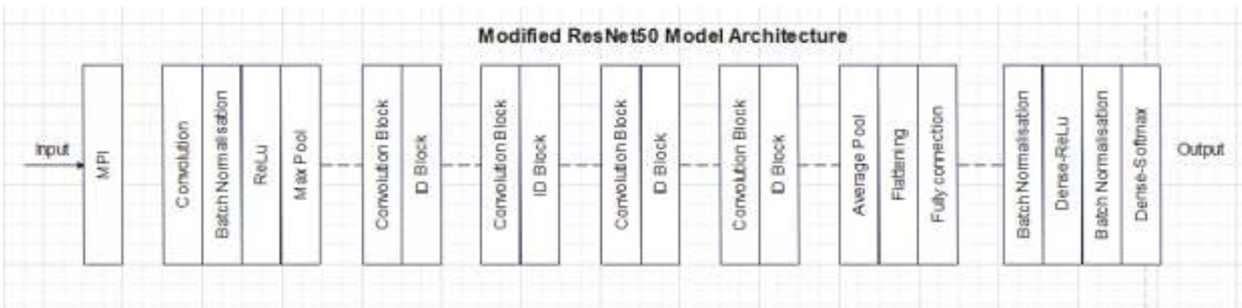

Figure.7 Modified ResNet50 Model Architecture

Inception V3:

In this custom designed variation of the InceptionV3 architecture, the top layer has been removed to facilitate adjustments that beautify its functionality and adaptableness for specific obligations. The inclusion of a flatten layer serves to reshape the output of the previous layers right into a one-dimensional array, preparing it for enter into next layers. Additionally, the mixing of batch normalization layers allows stabilize and accelerate the training system by using normalizing the center of every layer, which aids in mitigating troubles which includes inner covariate shift and gradient vanishing or exploding. Furthermore, a dense layer has been introduced to permit the community to learn difficult relationships and styles inside the function space, thereby improving its capability for classification or regression obligations. All the layers have been mentioned in the below architecture diagram Figure-8. This showcases the flexibility and scalability of deep learning architectures like InceptionV3.

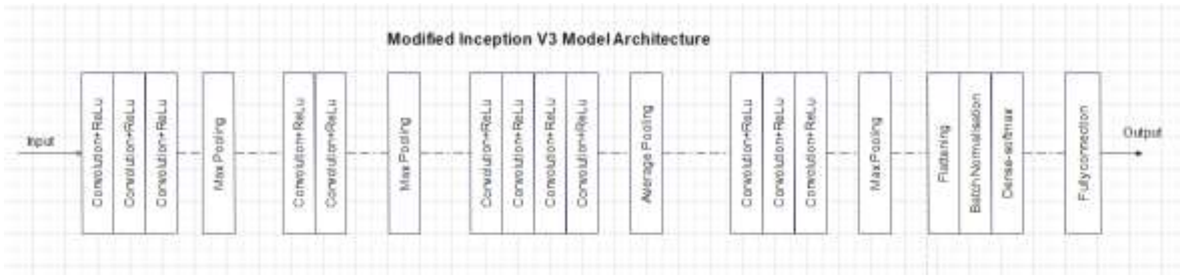

Figure.8 Modified InceptionV3 Model Architecture

MobileNet:

In the task of enhancing mobile networks through modification, a common approach involves augmenting a pretrained model by removing its top layer, effectively nullifying its previous output, and then appending additional layers to tailor it to specific objectives as mentioned in the below diagram Figure-9. In this particular scenario, after discarding the top layer, a sequence of layers is added to the model. The first added layer is a flattening layer, which reshapes the output of the preceding layer into a one-dimensional array, facilitating compatibility with subsequent layers. Following this, two batch normalization layers, strategically positioned to normalize the inputs to the following dense layers, are incorporated. Batch normalization aids in stabilizing and accelerating the training process by mitigating internal covariate shift. Finally, two dense layers are appended, promoting the model's capacity to capture intricate patterns and relationships within the data. This augmented architecture not only streamlines the computational efficiency of the mobile network but also fortifies its capacity for nuanced feature extraction and predictive accuracy.

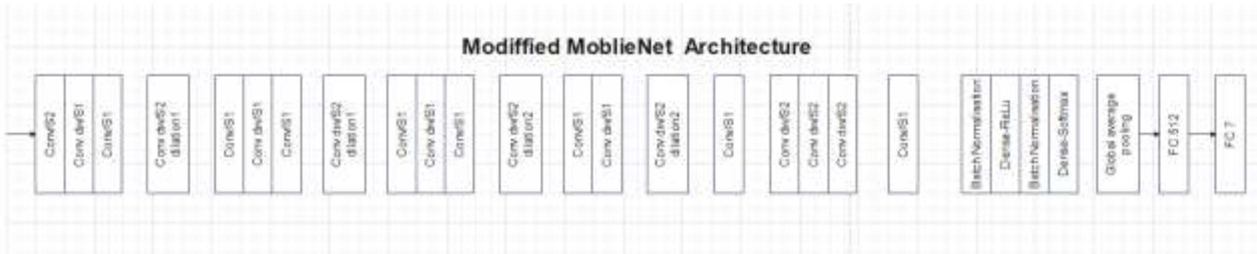

Figure.9 Modified MobileNet Model Architecture

IV. RESULTS AND DISCUSSION

Model	Loss	Accuracy	Val_loss	Val_acc
VGG19	0.1290	0.9535	0.5529	0.8025
ResNet50	0.0100	0.9971	0.2215	0.9590
InceptioV3	0.9734	0.5757	0.9508	0.5704
MobileNet	0.0037	0.9933	0.3149	0.9426

The modified models have achieved the results shown in above table. It provides a detailed breakdown of training loss, training accuracy, validation loss, and validation accuracy for each model. InceptionV3 records lowest accuracy of 57% with highest loss of 97%, and ResNet50 has highest accuracy of 99% with lowest loss of 1%. So, in hope of increasing the accuracy we tried transfer learning approach to have modified models.

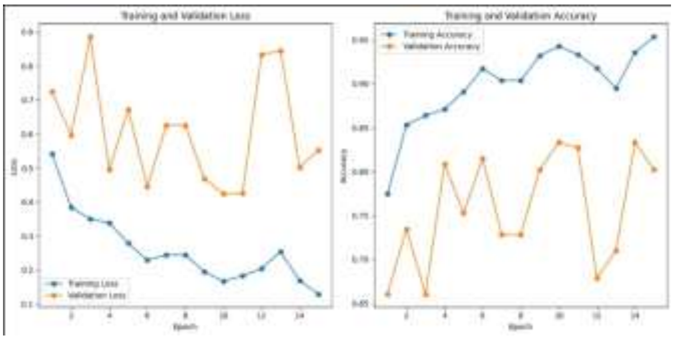

Figure.10 VGG19 training and validation performance graph for loss and accuracy metrics across different epoch values.

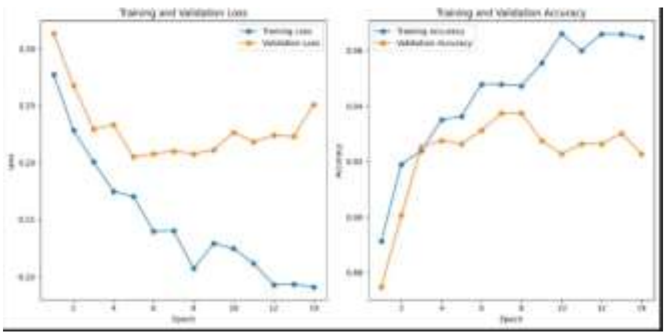

Figure.12 InceptionV3 training and validation performance graph for loss and accuracy metrics across different epoch values.

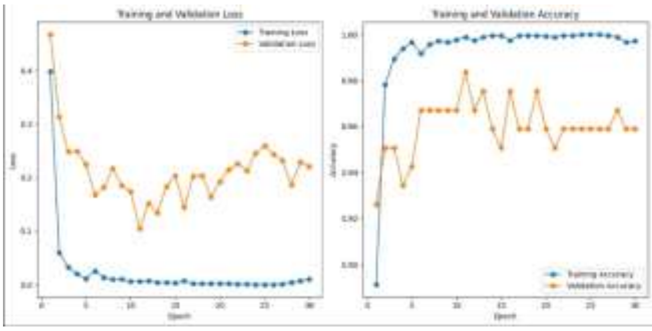

Figure.13 ResNet50 training and validation performance graph for loss and accuracy metrics across different epoch values.

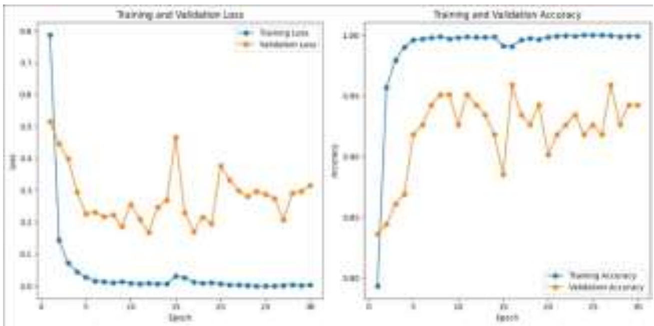

Figure.14 MobileNet training and validation performance graph for loss and accuracy metrics across different epoch values.

REFERENCES

[1] Singha, A., Moon, M. S. H., & Dipta, S. R. (2023, December). An End-to-End Deep Learning Method for Potato Blight Disease Classification Using CNN. In *2023 International Conference on Computational Intelligence, Networks and Security (ICCINS)* (pp. 1-6). IEEE.

[2] Singh, I., Jaiswal, A., & Sachdeva, N. (2024, January). Comparative Analysis of Deep Learning Models for Potato Leaf Disease Detection. In 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 421-425). IEEE.

[3] Durai, S., Sujithra, T., & Iqbal, M. M. (2023, June). Image Classification for Potato Plant Leaf Disease Detection using Deep Learning. In 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS) (pp. 154-158). IEEE.

[4] Gupta, A., & Tyagi, L. K. (2023, December). Potato Disease Detection Using Deep Learning. In 2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM) (pp. 01-08). IEEE.

[5] Wasalwar, Y. P., Bagga, K. S., Joshi, V. K., & Joshi, A. (2023, April). Potato Leaf Disease Classification using Convolutional Neural Networks. In 2023 11th International Conference on Emerging Trends in Engineering & Technology-Signal and Information Processing (ICETET-SIP) (pp. 1-5). IEEE.

[6] Wasalwar, Y. P., Bagga, K. S., Joshi, V. K., & Joshi, A. (2023, April). Potato Leaf Disease Classification using Convolutional Neural Networks. In 2023 11th International Conference on Emerging Trends in Engineering & Technology-Signal and Information Processing (ICETET-SIP) (pp. 1-5). IEEE.

[7] R, R., Bhargavi, M., Reshma, S., Manasa, C., & Chandana, J.H. (2023). Potato Leaf Disease Detection Using Deep Learning. 2023 9th International Conference on Smart Structures and Systems (ICSSS), 1-6.

[8] Shiffa, T. S., Suchithra, M. S., & Vijayakumar, A. (2023, December). Potato Leaf Diseases Detection Using Machine Learning And Deep Learning. In 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS) (pp. 1-6). IEEE.

[9] Sharma, O., Mohapatra, S., Mohanty, J., Dhiman, P., & Nonkra, A. (2023, May). Predicting Agriculture Leaf Diseases (Potato): An Automated Approach using Hyper-parameter Tuning and Deep Learning. In 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC) (pp. 490-493). IEEE.

[10] Charisma, R. A., & Adhinata, F. D. (2023, February). Transfer Learning With Densenet201 Architecture Model For Potato Leaf Disease Classification. In 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE) (pp. 738-743). IEEE.

[11] Kumar, A., Trivedi, N. K., & Tiwari, R. G. (2023, August). Disease Identification in Potato Leaves Using a Multi-Tier Deep Learning Model. In 2023 3rd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.

[12] Li, L. H., & Tanone, R. (2023, January). Disease Identification in Potato Leaves using Swin Transformer. In 2023 17th International Conference on Ubiquitous Information Management and Communication (IMCOM) (pp. 1-5). IEEE.

[13] Kukreja, V., Baliyan, A., Salonki, V., & Kaushal, R. K. (2021, August). Potato blight: deep learning model for binary and multi-classification. In 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 967-672). IEEE.

[14] Kasani, K., Yadla, S., Rachamalla, S., Hariharan, S., Devarajula, L., & Andraju, B. P. (2023, April). Potato Crop Disease Prediction using Deep Learning. In 2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT) (pp. 231-235). IEEE.

[15] Samant, D., Dhawan, R., Mishra, A.K., Bora, V., Diwakar, M., & Singh, P. (2023). Potato Leaf Disease Detection Using Deep Learning. 2023 IEEE World Conference on Applied Intelligence and Computing (AIC), 752-757.