International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Early Predictive Model for Detection of Plant Leaf Diseases Using MobileNetV2 Architecture

Mahesh T R[1*], Vinoth Kumar V[2*], R Sivakami[3], I Manimozhi[4], N Krishnamoorthy[5], B Swapna[6]

Submitted: 13/11/2022 **Accepted:** 15/02/2023

Abstract: In farming, it is vital to recognize diseases of plant leaves and to improve the prediction quality of diseased plant leaves. Several laboratory-based techniques such as polymerase reaction decrease in agricultural output, and pesticide application, have really been identified for recognizing different diseases of plant leaves with human sight. Agriculture yield is improving every day as a result of current technological advancements. However, they are very time-intensive and costly for farmers. Deep Learning (DL) methods may help boost crop yields by identifying recently upgraded methodologies and diverse systematic patterns. To improve the reliability of the measurements, researchers focused on new methodologies in deep learning algorithms for diagnosing leaf diseases. Every model is essential and focuses on the path of deep learning applications as well as the challenges faced by farmers. The mobilenetv2 architecture is used in this study to determine how to diagnose diseases that affect leaves. This design is built on an inverted residual structure, with narrow bottleneck layers serving as the input and output of the residual block and extended representations as the inputs. Lightweight depth-wise convolutions are used in this architecture to select leaf characteristics in the intermediate expansion stage. This network has 53 layers in the very beginning. There are 32 filters in the first completely convolution layer, followed by 19 more bottleneck layers. To retain representational strength, non-linearities were eliminated in the narrow layers. The proposed method enables the decoupling of the input/output leaves from the transformation's expressiveness, offering a framework for additional study. On Imagenet classification, COCO object detection, and VOC picture segmentation, the performance is evaluated. The proposed model provides an accuracy of 95% in identifying the plant diseases.

Keywords: *Deep Learning (DL), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Mobilenetv2 model, K-Nearest Neighbour (KNN), Convolutional Neural Network (CNN)*

1. Introduction

Producing enough food to meet societal demand is now possible thanks to developed technologies. The food and the crops safety and security, however, remained unachieved. Farmers face difficulties due to factors such as climate change, a decrease in pollinators, plant diseases, and other issues [1]. Priority needs to be given to establishing an important foundation for these elements. The utilization of analysis and detection techniques employing current technology aids farmers in solving such issues. Scientists and researchers analyse plant leaf diseases to pinpoint the major problems and difficulties. Below are a few of them: The leaf image must be of excellent quality and accessible to the general audience. Dataset requirements, noisy data affecting the leaf samples, segmentation process that may identify diseases but samples that must go through training and testing, classification is another challenge, and when it comes to detecting leaf diseases, the colour of the leaves may vary due to environmental factors, and a variety of diseases can be seen in various kinds of plants, making disease detection quite challenging.

The recognition of diseases of a plant leaf has been subjected to human gloss by vision scan, and agricultural production and cost can be considerably improved with the aid of a neural network [2]. When a crop is infected with illness, it produces a low yield and grows slowly. This disease is caused by bacterial, viral, and fungus infections, but can also be induced by changes in the environment, etc. Because the crop severely affects output and quality, plant leaf disease has remained one of the most serious

[1]Department of Computer Science and Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bangalore, India (trmahesh.1978@gmail.com)
[2]School of Information Technology and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India (drvinothkumar03@gmail.com)
[3]Department of Computer Science and Engineering, Sona College of Technology, Salem (sivakamir@sonatech.ac.in)
[4]Department of Computer science and Engineering East Point College of Engineering & Technology, Bangalore, India (drmanimozhi.i@eastpoint.ac.in)
[5]School of Information Technology and Engineering, Vellore Institute of Technology, Tamil Nadu, India (krishnamoorthy.n@vit.ac.in)
6Department of Electronics and Communication Engineering Dr MGR Educational and Research Institute, Chennai-95 (swapna.eee@drmgrdu.ac.in)

Correspondence:trmahesh.1978@gmail.com;drvinothkumar03@gmail.c

food safety hazards. However, the need to increase accuracy frequently comes at a price: current, cutting-edge networks demand significant computing resources that are beyond the scope of many embedded and mobile systems. In this research, we presented a neural network architecture designed for mobile and limited resource devices. MobilenetV2 has a novel layer module that reduces memory needed for processing. This proposed architecture reduces the number of operations, memory required and having the similar accuracy, our architecture advances the state of the art computer vision models. The proposed model can be applied on mobile devices making it simpler to identify plant leaf disease.

1.1. Deep Learning (DL)

Deep learning is based on ANN. Since neural networks are used to duplicate human minds, it is a kind of neural network that aids in brain copying. It is a popular topic nowadays since we didn't have nearly as much processing power and data in the beginning [3]. Neurons are the building blocks of deep learning. This allows attaining versatility and capability. This describes the universe as a layered network of ideas by deep learning; each approach is characterized by its connection to simple matters, and more abstract mathematical representations are calculated here.

1.2. Convolutional Neural Networks (CNN)

This is a form of ANN model that is used in pattern recognition systems. This CNN is specifically developed for photon information processing. With the help of natural language processing (NLP) and decision support systems, CNN will process the image powerfully, and artificial intelligence (AI) will employ deep learning to do both informative and creative activities. The device frequently uses clips and pictures for classification. The CNN is a system integration device that was designed to mimic the behaviour of cells in the brain. In general, neural networks are not the best choice for image processing, thus these pictures should be dimmed and in lower quality. Neurons are structured similarly to the forebrain, which in humans and other species is capable of receiving sensory images. To prevent the fragmented picture interpretation problem that plagues neural networks, the levels of cells are structured in a certain way to span the entire field of vision. CNN utilizes a multi-layered perceptron-like technology that is intended to reduce the processing time.

2. Related Work

The authors [1] presented a CNN model to assess and categorize numerous diseases that affect specific leaves, and they used methods include segmentation, extraction of features, and machine learning. The accuracy of Alex net, as per the authors of this work, is significantly small. The authors [2] concentrated on machine learning techniques and how they should be applied in agriculture. Although these approaches are nearly beneficial in vision-based farming, the finding reveals that they cannot be effectively implemented due to the camera's small focus length. Focusing on support vector machines, in this study, the authors [3] discovered a method for recognizing tomato leaf disease identification. Plant leaf image separation, scaling of the leaf picture and background elimination procedure using morphological processes is among the techniques employed.

The authors [4] concentrated on diagnosing diseases in 25 varieties of plants, reducing the need for collecting and analysing samples. It was indeed excessively similar. In this, the authors [5] accurately recognized the disease in the leaves of the plant and reported the solution. The authors [6] discovered the key common ailments in guava fruit, including anthracnose, Phomopsis fungus, and parasite-affected fruits like canker. This work [7] proposed a CNN model for categorizing various diseases affecting tomato plant leaves. A LeNet architecture has been used in the model. This methodology is designed to categorize the diseases produced by tomato plant leaves into ten separate categories. The CNN model had an accuracy rate of 94-95 percent. To train and evaluate many plants leaf pictures, the author employed a CNN model [8].

A trained AlexNet architecture is deployed to identify the morphological phase of selected species. The accuracy of this model was 87%. This model delivers positive accuracy when differentiated to other techniques. A visual computation Guava leaf disease diagnostic approach has been introduced in this study [9]. The authors used 125 Guava leaf specimens (25 in each class) to study five diseases, and discovered that SVM outperforms k-NN models in this instance. The authors [8] proposed a realistic and effective method for separating the defective regions of fruits. Some work has previously been done on tropical fruits for autonomous examination, diagnosis, and classification of various diseases, such as black spot, rot, or citrus Huanglongbing, among many others [9–13]. Some attempts at CNN-based seasonal fruit disease categorization have been made, but no earlier studies have focused on guava fruit diagnostics.

This study [14] used deep CNN and the Caffe architecture to build a plant disease identification technique based on plant leaf picture categorization. The presented method can classify thirteen different kinds of crop diseases with a degree of confidence ranging from 91 percent to 98 percent for discrete examinations, with a mean accuracy of 96 percent. The authors [15] suggested an ANN type model that was tested on 600 seasonal fruits. The best validation accuracy for normal fruit kinds is 85% and afflicted anthracnose fruit kinds are 77%. This study [16] has presented another ANN-based disease diagnosis method using a grape plant leaf model.

The paper is organized further as follows. The third section provides the proposed methodology, discusses the dataset taken for the experiment, pre-processing steps carried out and the algorithms being implemented. The fourth section provides the results of the proposed methodology and the conclusion section summarizes the overall performance achieved by the proposed methodology with respect to accuracy.

3. Methodology

There are totally five stages in the proposed methodology and the process diagram is shown in figure 1.

Step 1: **Import the given image from dataset**

The proposed work brings in data set using keras preprocessing image data create function and also generates size, rescale, range, zoom range, horizontal flip. Then bring photo dataset is brought in from the folder through the data create function.

Step 2: **Data preprocessing**

To convert photos of various sizes into a uniform setting, data pre-processing is used. All of the plant leaf photos were first scaled down to the same size of 224 pixels. The input photos are 224 by 224 by 3, in size. Data augmentation is used as a regularization technique while expanding the training dataset to improve the model's generalizability. To build batches of real-time augmented image data, Keras class Image Data Generator has been used. For the training dataset, the augmentation techniques of width shift and height shift were used with a range of 0.1.

Step 3: **Prepare training and validation dataset**

The dataset is split into training dataset and validation dataset. The training dataset is used to build the model.

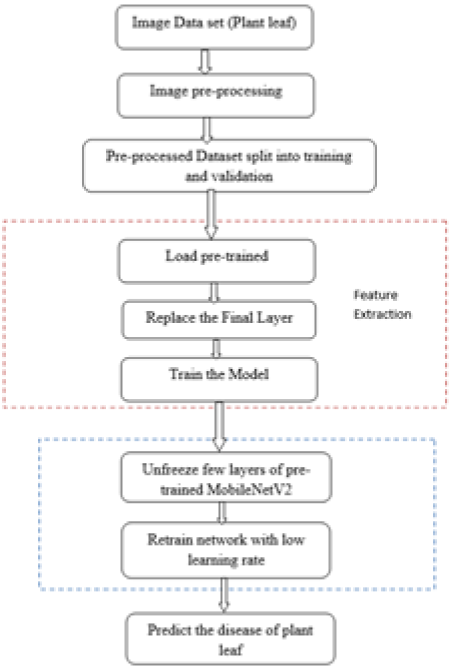


Fig. 1. Process diagram

Step 4: Using MobilenetV2 network

In feature extraction, MobilenetV2 network is used. The network uses 1 x 1 point convolution to expand input channels and extract input features and convolution integrator linearly to combine the output features while reducing the size. It replaces ReLU6 with a linear function to activate the output channel size. MobilenetV2 is depicted in figure 2.

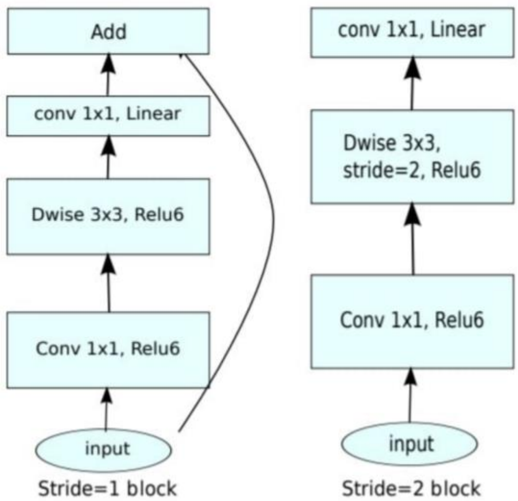


Fig. 2. MobilenetV2 Network

The MobilenetV2 uses reverse block to combine features and features when traversing convolution to gain more function. The base model was pre-trained using the ImageNet dataset. Several layers were added to top of convolutional layers of MobilenetV2 [17]. The base model uses the pre-trained weights, and layers are being frozen. Thus, in the training process only the weights of the head model were trained. Adam optimizer was used because it reaches convergence quickly. Less epochs were used and the batch size is 32. Categorical cross entropy is used as a loss function that the network needs to minimize during training. The data used in each iteration varies and the resulting model has better generalization ability [18-20].

Step 5: Plant disease identification

Input images were given using keras preprocessing package. That input image is converted into an array value using the NumPy package [21-22]. Already the leaf in the dataset is classified. It predicts leaf diseases using a predict function. It will determine the remedy for a particular plant leaf disease.

3.1 Dataset

From Kaggle the Plant Village dataset has been taken, which has 54303 healthy and unhealthy leaves separated into 38 classifications depending on plant type and condition. The features include colours, forms, and textures [23-26]. Some of the common leaf diseases are shown in figure 3.

The unprocessed directory includes the following variations:

- original RGB colour photos
- segmented: RGB photos with only the leaf segmented and colour corrected grayscale:
- gray scaled variant of the photos

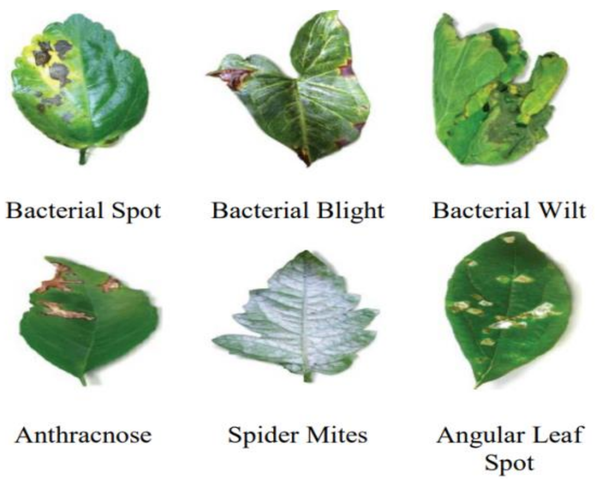


Fig.3. Different Leaf based Diseases

3.2 Noise Removal

Image transmission, capture and other processes will all produce noise. Noise suppression is an important consideration when performing classification procedures that are dependent on image quality [27, 28].

A filter is used to reduce Gaussian noise while maintaining the integrity of the other features. When the input image is designated by fp and the maximum intensity value is marked as f_{max}, a function is calculated. F_p is computed using the equation (1)

$$F_p = \begin{cases} 1, \; f_p = 0 \; or \; 255 \\ \exp\left(-\frac{(f_p - f_{\max})^2}{2*8*\sigma}\right), \; otherwise \end{cases} \qquad (1)$$

Here, σ is the standard deviation. It is applied independently for every one of the R, G, and B parts of each photo because we have colour photos. Finally, intended image is arrived by integrating multiple elements.

3.3 Removal of Over-fitting

Data Augmentation: To obtain greater generalization ability and prevent over fitting, data augmentation is used on the training dataset. Width shifting, zooming, rotating, horizontal flipping, Re-scaling, shearing, height shifting, and other techniques are employed to enhance data. This method reduces the computational complexity of a model by reducing function that determines loss, as it is hypothesized that perform better on unknown testing data.

L2 regularization: Using equation shown in (2), L2 regularization has been to reduce the square of the gap between the target and forecast classes.

$$J(\theta) = \frac{1}{2m}\left(\sum_{j=1}^{m} \left(h_\theta\left(x^i - y^{(i)}\right)^2\right) + \lambda \sum_{j=1}^{n} \theta_j^2\right)) \quad (2)$$

3.4Activation Functions

The dataset is trained using a fit generator function and classifier as well as training per epochs, validation data, total number of epochs and validation steps.

The photos are supplied into the DL through the entry layer, and the neuron performs a transformation based on the equation (3).

$$X = (weight * input) + bias \qquad (3)$$

The output of this transformation is sent to the next hidden layer using the Activation function as shown in equation (4), and the process is repeated.

$$Y = Activation((weight * input) + bias) \qquad (4)$$

The activation functions used by this model are ReLu and Softmax. The ReLu is given by the equation (5)

$$f(z) = \max(0, z) \qquad (5)$$

The Softmax is given by the equation (6).

$$\pi(y)j = \frac{e^{yj}}{\sum_{z=1}^{z} e^{yj}} \text{ for } j = 1,2,3,4 \ldots k \qquad (6)$$

3.4.1 Gray Scale Base

Gray data included inside a leaf might be seen as significant traits. Leaf characteristics including form, veins, and damaged portions of the leaf seem heavier than the adjacent leaf areas. Several new feature extraction techniques examine split leaf areas for local gray minima. To enhance the quality of local dark spots and therefore make identification simpler, the input pictures are first upgraded using counterpoint and pale morphological procedures.

Dark patches are extracted using a low-level gray-scale threshold. The approach is based on three layers. It employs tiered leaf gray scale behaviour in pyramid pictures. There are three tiers of leaf location. The higher two levels are based on photos of various resolutions, while the bottom level is based on an edge detection algorithm.

3.4.2 Edge Base

This work involved evaluating line drawings of leaves taken from pictures in order to find leaf characteristics. The photos are first upgraded with a median filter to remove noise and histogram analysis to modify contrast.

The MobilenetV2 model will be used for extraction of features. Convolution depth separators are used for the MobilenetV2 model. To begin, the mobilenetV2 network expands the input channels using 1x 1 point convolution. The deep convolutional model is then used to retrieve the incoming features, and the convolution aggregator is used to linearly aggregate the resulting characteristics while minimizing the capacity of the network. It substitutes the ReLU6 to activate the resulting channel size to match the input after lowering the size. The inverse block is also used by the MobilenetV2 system to mix features across brief networks and features while navigating convolution to increase result efficiency. The source streams and filters are divided into distinct areas using depth convolution.

3.5 Layers Used

This section provides the explanation of different layers.

3.5.1 Convolution Layer

Because it pulls information from the image, the Convo layer is also known as the Feature Extractor Layer. To begin, a piece of the picture is linked to the Convo layer, which performs the previously mentioned convolution process as well as calculating the dot product among the receptive region and the filter. The result is an integer that represents the output level. Repeat the process with a Stride filter on the subsequent receptive field of the same input picture. It will continue doing it over and over until the entire image has been processed. The output for another layer will be the outcome.

3.5.2 Pooling Layer

The pooling layer is used to lower the geographical region of the source images after convolution. It is used across two convolutions. Applying FC after the Convo layer without max pooling will be resource intensive. As a result, pooling is the sole technique to minimize the geographical region of the source images. In such a single scale layer with a Stride of 2, max pooling was applied. The 4 × 4 input appears to be reducing to 2 x 2 proportions.

3.5.3 Fully Connected Layer

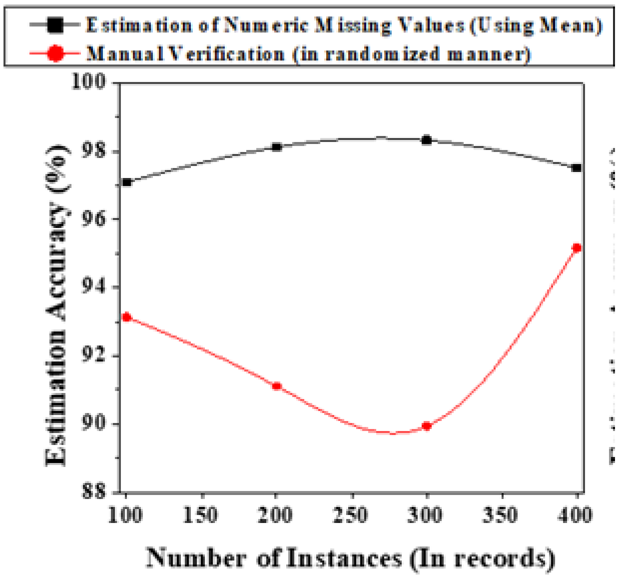
The fully collected layer appears to include weights and biases. It connects neurons in one layer with those in other. It's used to teach individuals how to categorize images into various groups.

3.5.4 Output Layer

The label, which is one-hot encoded, is contained in the resulting layer.

4. Results

Initially, the few vital steps taken in the pre-processing stages are evaluated. So the essential process is to access the estimation accuracy of mean and mode methods in identifying the missing values over 10860 considered instances. Thus, figure 4 (a) and 4 (b) depicts the resultants of the process in comparison with the randomized manual verifications. The results precisely state that the average estimation accuracy of both mean and mode methods in identifying numerical and nominal missing values are 97.77% and 96.93%, respectively.



(a)

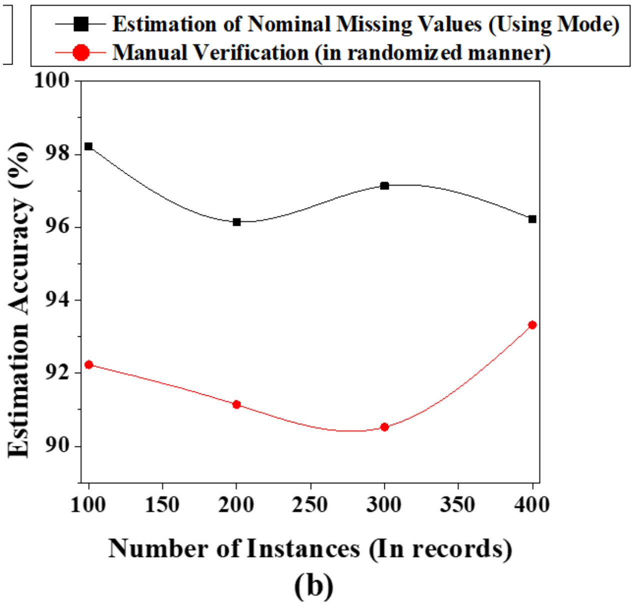


Fig. 4. Missing Value Estimation Accuracy Using Mean and Mode Method

Confusion Matrix of Healthy and Diseased Plant leaves by considering 38 classes as two classes namely Healthy and Diseased for simplified calculation of performance metrics is shown in table 1.

Table 1: Confusion Matrix of Healthy and Diseased Plant Classes on test dataset

Leaf Type	Healthy	Diseased
Healthy	10345	515
Diseased	557	10303

True Positive (TP) = 10345 False Positive (FP) =557

False Negative (FN) = 515 True Negative (TN) =10303

Let TP be t_{pos}, FP be f_{pos} and FN be f_{neg}, TN be t_{neg}. The different performance metrics are calculated as shown in the equations from (7) to (10).

$$Accuracy = \frac{(t_{pos} + t_{neg})}{(t_{pos} + t_{neg} + f_{pos} + f_{neg})} \quad (7)$$

$$Recall = \frac{t_{pos}}{(t_{pos} + f_{neg})} \quad (8)$$

$$Precision = \frac{t_{pos}}{(t_{pos} + f_{pos})} \quad (9)$$

$$F1 - Score = \frac{(2 * Precision * Recall)}{(Recall + Precision)} \quad (10)$$

The figure 5 shows the performance of the proposed model against different performance metrics.

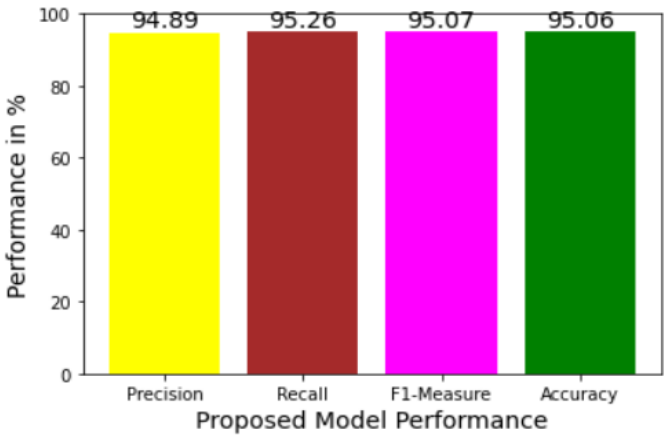
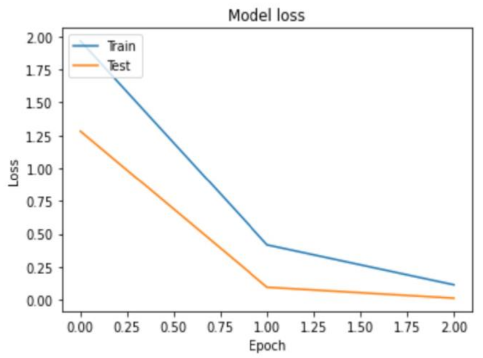
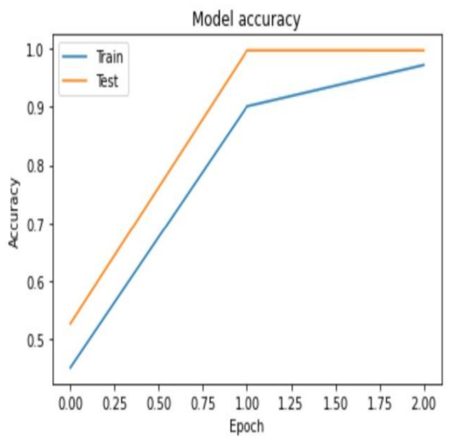


Fig.5. Performance of the proposed model

The figure 5(a) depicts the model loss per epoch and the figure 5 (b) shows the model accuracy per epoch and it was observed that the average model accuracy is 95%.



5 (a)



5(b)

Fig.5 (a) Model loss per epoch, (b) Model accuracy per epoch

In this research, we change the base model of the proposed method into InceptionV3, ResnetV2 and then measure their performance on the Plant Village Dataset. These networks are pre-trained using ImageNet dataset and added with the same head model. MobilenetV2 has the highest accuracy and most lightweight network architecture. InceptionV3 gives the highest precision value compared to other models.

Table 2 provides the comparison among different techniques against different performance parameters.

Table 2: Comparison among different techniques

Base Model	Accuracy	Precision	Recall	F1-Measure
MobilenetV2	95.06	94.89	95.26	95.07
ResNet50V2	94.23	92.30	94.25	93.26
InceptionV3	94.89	95.20	94.12	94.65

Table 3 shows the comparison of the proposed work with the related models and the proposed model accuracy is 95% and it outperforms compared to other models.

Table 3: Comparison with existing studies

Study	Plant Leaf considered	Disease detected	Method used	Accuracy (%)
Proposed Work	Apple	Apple Scrab	Mobilenet V2	95
Abirami, M.T.S [7]	Guava	Algal Leaf Spot	k-Nearest Neighbor	88
Amanda Ramcharan et al,[8]	Cassava	Red mite	InceptionV3	93
Peng Jiang et al. [9]	Apple Leaf	Mosaic Grey spot Alternaria leaf spot Brown spot Rust	INAR-SSD	78
JagadeeshDevdas Pujari et al. [15]	Mango Grape Pomegranate	Anthracnose	K-means clustering	76

5. Conclusion

It is expected to be known from any field dataset, as well as prior datasets utilized, to diagnose and forecast plant disease using the CNN model. This leads to the following disease predictions: Because this technique covers a wide range of plant leaves, a farmer can choose which plants to nurture and which plants to avoid. He can also make a list of all the potential leaves. It aids the farmer in making judgments about which crop's plant leaves to grow. The proposed method consists of a pre-trained base model and trainable head model. The proposed model provides an accuracy of 95% in identifying the plant diseases. Comparison of base model architecture shows that MobilnetV2 gives the highest accuracy than ResNet50V2, InceptionV3. It uses an average pooling layer followed by two fully connected layers, giving more accuracy while maintaining efficiency. In addition, this technology considers previous production data and assists the farmer in determining the demand for and pricing of particular plants on the market. Further improvement of the dataset is needed, to extend this research to a large number of plant species. This study can be extended with a custom optimizer that can be used in resource constrained environments.

Data availability

The dataset used in the research work will be shared by the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Madhulatha, G., and O. Ramadevi. "Recognition of Plant Diseases using Convolutional Neural Network." In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), pp. 738-743. IEEE, 2020.

[2] D. Devarajan D Stalin Alex, T R Mahesh, Rajanikanth Aluvalu, V Vinoth Kumar, Uma Maheswari, S Shitharth, "Cervical Cancer Diagnosis Using Intelligent Living Behaviour of Artificial Jellyfish Optimized with Artificial Neural Network," in IEEE Access, 2022, https://doi.org/10.1109/ACCESS.2022.3221451

[3] Rahman, Md Arifur, Md Mukitul Islam, GM ShahirMahdee, and MdWasiUlKabir. "Improved Segmentation Approach for Plant Disease Detection." In 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1-5. IEEE, 2019.

[4] Tetila, Everton Castelão, Bruno Brandoli Machado, Gabriel Kirsten Menezes, Adair da Silva Oliveira, Marco Alvarez, Willian ParaguassuAmorim, Nícolas Alessandro De Souza Belete, GercinaGonçalves da Silva, and HemersonPistori. "Automatic recognition of soybean leaf diseases using UAV images and deep convolutional neural networks." IEEE Geoscience and Remote Sensing Letters 17, no. 5 (2019): 903-907.

[5] Ramakrishna, M.T.; Venkatesan, V.K.; Izonin, I.; Havryliuk, M.; Bhat, C.R. Homogeneous Adaboost Ensemble Machine Learning Algorithms with Reduced Entropy on Balanced Data. Entropy 2023, 25, 245. https://doi.org/ 10.3390/e25020245

[6] Al Haque, ASM Farhan, Rubaiya Hafiz, MdAzizul Hakim, and GM Rasiqul Islam. "A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach." In 2019 22nd International Conference on Computer and Information Technology (ICCIT), pp. 1-6. IEEE, 2019.

[7] S. Roopashree, J. Anitha, T.R. Mahesh, V. Vinoth Kumar, Wattana Viriyasitavat, Amandeep Kaur, An IoT based authentication system for therapeutic herbs measured by local descriptors using machine learning approach, Measurement, Volume 200, 2022, 111484, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2022.111484

A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," Frontiers in plant science, vol. 8, p. 1852, 2017.

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

[9] Mahesh, T.R., Vinoth Kumar, V., Vivek, V. et al. Early predictive model for breast cancer classification using blended ensemble learning. Int J Syst Assur Eng Manag (2022). https://doi.org/10.1007/s13198-022-01696-0

[10] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks- based recognition of plant diseases by leaf image classification," Computational intelligence and neuroscience, vol. 2016, 2016.

[11] P. Chaitanya Reddy, R. M. S. Chandra, P. Vadiraj, M. Ayyappa Reddy, T. R. Mahesh and G. Sindhu Madhuri, "Detection of Plant Leaf-based Diseases Using Machine Learning Approach," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2021, pp. 1-4, doi: 10.1109/CSITSS54238.2021.9683020.

[12] Mahesh, T. R., et al. "An Efficient Ensemble Method Using K-Fold Cross Validation for the Early Detection of Benign and Malignant Breast Cancer." International Journal of Integrated Engineering 14.7 (2022): 204-216.

[13] M. Tahmooresi, A. Afshar, B. Bashari Rad, K. B. Nowshath, M. A. Bamiah, "Early detection of breast cancer using machine learning techniques," Journal of Telecommunication, Electronic and Computer Engineering, vol. 10, no. 3-2, pp. 21-27, 2018.

[14] Ebru Aydindag Bayrak, Pinar Kirci, TolgaEnsari,"Comparison of machine learning methods for breast cancer diagnosis.2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), pp. 1-3,2019.

[15] Dursun Delen; Glenn Walker; Amit Kadam (2005). "Predicting breast cancer survivability: a comparison of three data mining methods", Procedia Computer Science34(2), 113–127.

[16] Asri, Hiba; Mousannif, Hajar; Moatassime, Hassan Al; Noel, Thomas (2016). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. Procedia Computer Science, 83(), 1064–1069.

[17] Akbugday, Burak (2019). 2019 Medical Technologies Congress (TIPTEKNO) - Izmir, Turkey (2019.10.3-2019.10.5)] 2019 Medical Technologies Congress (TIPTEKNO) – "Classification of Breast Cancer Data Using Machine Learning Algorithms",1–4.

[18] T R, M., Kaladevi, A. C., J M, B., Vivek, V., Prabu, M., & Muthukumaran, V. (2022). An Efficient Ensemble Method Using K-Fold Cross Validation for the Early Detection of Benign and Malignant Breast Cancer. International Journal of Integrated Engineering, 14(7). https://doi.org/10.30880/ijie.2022.14.07.015 Retrieved from https://publisher.uthm.edu.my/ojs/index.php/ijie/article/view/10810

[19] Qinghua Huang, Yongdong Chen, Longzhong Liu, Dacheng Tao, Xuelong Li, On combining biclustering mining and adaboost for breast tumor classification, IEEE Trans. Knowl. Data Eng. 32 (4) (2020)728–738.

[20] Tamoor, Maria and Younas, Irfan. 'Automatic Segmentation of Medical Images Using a Novel Harris Hawk Optimization Method and an Active Contour Model'. 1 Jan. 2021: 721 – 739.

[21] SapnaJuneja, AbhinavJuneja, GauravDhiman, SanchitBehl, SandeepKautish, "An Approach for Thoracic Syndrome Classification with Convolutional Neural Networks", Computational and Mathematical Methods in Medicine, vol. 2021, Article ID 3900254, 10 pages, 2021. https://doi.org/10.1155/2021/3900254.

[22] K. K. Jha, R. Jha, A. K. Jha, M. A. M. Hassan, S. K. Yadav and T. Mahesh, "A Brief Comparison On Machine Learning Algorithms Based On Various Applications: A Comprehensive Survey," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2021, pp. 1-5, doi: 10.1109/CSITSS54238.2021.9683524.

[23] Siddique, N.A., Paheding, S., Elkin, C.P., &Devabhaktuni, V.K. (2021). U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access, 9*, 82031-82057.

[24] Dharahas, Reddy T., and T. R. Mahesh. "A Pragmatic Approach for Detecting Brain Tumors Using Machine Learning Algorithms." BIOSCIENCE BIOTECHNOLOGY RESEARCH COMMUNICATIONS Special Issue 14.11 (2021).

[25] Mahesh, T. R., Vinoth Kumar, V., Muthukumaran, V., Shashikala, H. K., Swapna, B., & Guluwadi, S. (2022). Performance Analysis of XGBoost Ensemble Methods for Survivability with the Classification of Breast Cancer. Journal of Sensors.

[26] R, M. T. ., Goswami, T. ., Sriramulu, S. ., Sharma, N. ., Kumari, A. ., & Khekare, G. . (2022). Cognitive Based Attention Deficit Hyperactivity Disorder Detection with Ability Assessment Using Auto Encoder Based Hidden Markov Model. International Journal of Communication Networks and Information Security (IJCNIS), 14(2), 53–65. https://doi.org/10.17762/ijcnis.v14i2.5464

[27] Ramana, Kadiyala, Madapuri Rudra Kumar, K. Sreenivasulu, Thippa Reddy Gadekallu, Surbhi Bhatia, Parul Agarwal, and Sheikh Mohammad Idrees. "Early prediction of lung cancers using deep saliency capsule and pre-trained deep learning frameworks." Frontiers in oncology 12 (2022).