Ecological Informatics 81 (2024) 102618

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

An ensemble of deep learning architectures for accurate plant disease classification

Ali Hussein Ali [a], Ayman Youssef [b], Mahmoud Abdelal [c], Muhammad Adil Raja [d,*]

[a] Basra University, Department of Computer Science, Iraq
[b] Electronics Research Institute (ERI), Department of Computers and Systems, Cairo, Egypt
[c] Benha University, artificial intelligence department and Mechatronics Engineering department, Cairo, Egypt
[d] Regulated Software Research Centre (RSRC), Dundalk Institute of Technology (DkIT), Dundalk, Ireland

ARTICLE INFO

Keywords:
Plant leaf disease
Agriculture
Ensemble deep learning models

ABSTRACT

A substantial fraction of agricultural produce loss can be attributed to plant diseases. Agricultural yield loss can have far-reaching consequences for a country's economy and contribute to global food insecurity. Early detection of plant diseases can be instrumental in maintaining global health and welfare. A pathologist's visual evaluation is typically used to make an early diagnosis of plant diseases. This technique involves experts or farmers examining plants with the naked eye and classifying the disease depending on their previous experience. This conventional approach includes drawbacks like low accuracy and the need for human expertise. This motivates researchers to investigate automated systems for the early diagnosis of plant diseases.

To achieve this goal an ensemble of different deep learning architectures (DenseNet201, efficientNetB0, inceptionresnetV2, efficientNetB3) is introduced to increase the classification accuracy of plant leaf diseases. In this work, a novel image-processing technique is proposed to increase the efficiency of deep-learning models. Also, a data balancing technique is used to solve the problem of the imbalanced dataset. Five different deep-learning models are trained and tested using the largest plant disease dataset; PlantVillage. Ten different ensembles (chosen randomly) of the deep learning models are tested and compared to find the ensemble with the highest accuracy.

The proposed ensemble model was able to achieve 99.89% accuracy on the New PlantVillage dataset. PlantVillage is a challenging dataset with 38 classes. Achieving high accuracies on such a dataset proves the ability of the system to generalize on unseen data or real-world scenarios. A comparison with the state-of-the-art is made with other available models from the literature. A section about this is added to show the superior performance of the proposed ensemble model in terms of accuracy and F1-score.

1. Introduction

Today, agriculture is considered as one of the greatest means of food production for the ever-increasing human population. Deep learning models are used in solving many problems associated with agriculture (Kamilaris and Prenafeta-Boldú, 2018). One of the problems that highly affects agriculture is leaf disease. It is estimated that 35% of the agricultural produce is wasted due to plant diseases (Oerke, 2006). This wasted produce affects the economy of the country and causes an undesirable increase in prices for food supplies. In (He et al., 2021) it is stated that plant disease may result in 13% to 22% annual yield loss or, equivalently, billion of US dollars in economic costs.

There are traditional techniques for the detection and classification of plant diseases. These techniques involve farmers examining plants with the naked eye and trying to classify the disease depending on their previous experience. However, in the current technologically advanced era, an automated system (Sankaran et al., 2010) that utilizes deep-learning models must be the fastest and most accurate approach. These deep learning models have the advantage of detecting plant diseases early even before symptoms are visible to the human eye (Mzoughi and Yahiaoui, 2023). This is a considerable stimulus for scientists and researchers to investigate methods and means for automated identification and classification of plant diseases. Previously it was a complex task to build such systems and provide them with needed datasets.

* Corresponding author.
E-mail addresses: aymanmahgoub@eri.sci.eg (A. Youssef), mahmoud.abdelaal@fci.bu.edu.eg (M. Abdelal), adil.raja@dkit.ie (M.A. Raja).

https://doi.org/10.1016/j.ecoinf.2024.102618
Received 22 September 2023; Received in revised form 22 April 2024; Accepted 24 April 2024
Available online 22 May 2024
1574-9541/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

However, with new technologies like smartphones, it has become easier to build such systems and collect needed datasets. This paper proposes single and ensemble models for plant disease classification. In the case of a single model, the model can be downloaded directly on a smartphone. In the case of large ensemble models, the models can be downloaded on a server. Images will be taken using smartphones from the field and sent to the server.

Early recognition would be crucial to stop the disease from further spreading to other areas of the plant. The advantage is that by spotting crop illnesses as soon as they manifest on the leaves, then they may be treated and kept under surveillance.

Farmers frequently cannot detect the onset of diseases. Occasionally they incorrectly identify the diseases that harm agricultural output. Given that different illnesses need different treatment modalities, incorrect disease identification leads to inappropriate management of plant diseases. As a result, despite medical intervention, a plant's health does not improve. In addition to financial and manpower costs, this leads to a considerable decline in production.

To date, many regions of the world, especially developing nations, still rely on conventional methods for plant disease classification. When a rapid outbreak of a disease occurs, researchers from universities travel to the region to advise the farmers. This conventional method needs a large number of experts to be able to travel to different regions where the disease outbreak has happened.

There is a huge loss in output, time, money, and product quality when a plant disease is incorrectly diagnosed. For successful cultivation, it is essential to determine the condition of the plant under test. Due to the advent of the Internet and advancements in computer vision, it has become possible to offer a practical solution to this issue. A much more desirable solution is to identify diseases in leaves with high accuracy as early as possible and take appropriate action. It is hoped that the solution suggested in this work, utilizing deep learning techniques, could be a crucial part of resolving this predicament.

In this work, novel contributions in the area of plant disease classification are presented. These are as follows:

- A novel image processing technique applied specifically for the problem of plant disease classification is proposed. The accuracy of any image classification deep learning system depends on the imaging technique used on the image dataset. Hence, in this work, a novel method for Image enhancement is proposed.
- A class-weighted algorithm for data balancing is proposed to solve the problem of imbalanced datasets. The imbalanced dataset can cause the deep learning model to become biased towards certain classes. Hence, a well-known simple data balancing technique is proposed in this work.
- Five different deep learning models are tuned to work with the proposed dataset. Five pre-trained models are fine-tuned for the plant disease classification problem.
- The final contribution is to build ensemble models from the trained deep learning architectures to achieve high classification accuracy. In this final step, we test some ensembles from the five proposed deep-learning models.

The rest of this article is organized as follows. In Section 2, we present the previous related work done by researchers on problems related to plant disease prediction. Section 3 introduces the methodology of the proposed ensemble models. Section 4 discusses the results of the experiments. Section 5 shows a comparison between the proposed work and similar works from the literature that comprises the state-of-the-art. Section 6 presents our concluding remarks.

2. Related work

The main steps of any image disease classification problem are the pre-processing of image data, feature extraction, and finally the development of the classification model. Many pre-processing stages can be used in image disease classification. These include image enhancement, color image transformation, image resizing, and image filtering (Iqbal et al., 2018). After pre-processing the diseased region is segmented. Features are extracted from the segmented areas for the training process using segmentation methods (Madhogaria et al., 2011) such as K-means clustering (Hartigan and Wong, 1979), edge detection (Marr and Hildreth, 1980), fuzzy logic(Zadeh, 1988), Otsu thresholding (Xu et al., 2011), and histogram matching (Shen, 2007). Finally, the features (which are vectors of numbers) are used to train a model (like Support Vector Machine (SVM), K-Nearest Neighbor (KNN), etc) with a suitable classification technique to detect the disease. Some of the well-known and popular classification techniques include Deep Neural Network (DNN), SVM (Cortes and Vapnik, 1995), Artificial Neural Network (ANN) (Jain et al., 1996), KNN (Peterson, 2009), etc.

In (Anandhakrishnan and Jaisakthi, 2022) a novel Deep Convolutional Neural Network (DCNN) with a small number of layers is proposed. The proposed DCNN model was trained on tomato leaf images collected from PlantVillage dataset and achieved 98.4% accuracy on the test data. The proposed model is compared to the work proposed in (Geetharamani and Pandian, 2019). Comparison shows that the proposed model not only achieves higher accuracy but also has a smaller number of layers. However, the proposed model is trained on tomato leaf images only.

In (Sangeetha and Rani, 2021) a transfer learning-based method using VGG16 and VGG19 is employed. The two models are compared to show which model achieves higher accuracy. It can be seen from the results that VGG16 has an accuracy of 97.79% which is higher than that of VGG19 with 94.7% accuracy. However, the proposed model is trained on tomato leaf images only.

In (Zhang et al., 2018), DCNN based model is introduced using different architectures such as AlexNet, GoogLeNet, and ResNet. This model is used to detect the disease of the maize leaves. These models are trained on nine different classes. One class represents healthy leaves, and the rest of the eight classes represent infected leaves. These diseases are Curvularia leaf spot, dwarf mosaic, gray leaf spot, northern leaf blight, brown spot, round spot, rust, and southern leaf blight. This model achieves 97.28% as the highest accuracy. The model proposed in this work is trained on nine classes only.

Another approach based on deep learning that was applied to recognize leaf diseases is proposed in (Ghazi et al., 2017). In this work, the authors tested their methodology on the LifeCLEF 2015 dataset. The dataset contains 91,758 labeled images of different plant organs (e.g. flowers, fruits). It utilizes three different models, namely, AlexNet, GoogLeNet, and VGGNet. However, the proposed work achieves small accuracies on the proposed dataset. As the best-achieved accuracy was 78.44% using VGGnet. The results also show the superior performance of AlexNet when trained from scratch due to its simpler architecture.

Several data augmentation methods are applied to the data for a Convolutional Neural Network (CNN)-based model in (Shijie et al., 2017). The data augmentation technique enhances the diversity of the data without the need to attach or add a new dataset to the training process. They utilize Generative Adversarial Networks (GANs), flipping, cropping, shifting, Principal Components Analysis (PCA), color, noise, and rotation techniques. The experiments conducted in this research prove that the performance of four augmentation techniques (cropping, flipping, GANs and rotation) is better than the other augmentation methods. It is obvious that using a combination of various augmentation methods is significantly more advantageous than using individual methods alone.

A different augmentation method is proposed in (Pandian et al., 2019). GAN and Neural Style Transfer (NST) augmentation methods were applied for plant leaf disease detection. Different transfer learning models like VGG16, ResNet, and InceptionV3 were tested using these augmentation techniques. A comparison between the performance of these methods shows that using the combination of the methods is better than using individual techniques. A comparison between the different three models show the superior performance of InceptionV3 over other models. The model achieves an accuracy higher than 91.5%.

In (Pandian et al., 2022) a new dataset was created using a number of open-source plant leaf datasets. Three different data augmentation techniques were applied to balance the dataset Basic Image Manipulation (BIM), Deep Convolutional Generative Adversarial Network (DCGAN) and NST.

The dataset consists of 147,500 images of 58 different healthy and ill plant leaf classes and one no-leaf-disease class. The coarse-to-fine technique is used to choose the best hyperparameter values for the model. The systems achieved a high accuracy of 99.9655%. The proposed work does not compare its results with other models on the same dataset. Hence, the evidence of the superior performance of the proposed models and architecture is not clear.

In (Sharma et al., 2023) a new deep lightweight CNN architecture (DLMC-Net) is proposed for multi-class classification of plant leaves. The model is tested for performing real-time leaf disease classification. The system was able to achieve 93.56% accuracy. In the proposed model, a sequence of collective blocks along with a passage layer was employed to extract deep features. The system achieves low accuracy while our proposed system achieves higher accuracies compared to this system. In (Tiwari et al., 2021) a new dense CNN model is proposed for plant disease classification. The proposed model was tested on a dataset of more than 25,000 images and 27 classes of healthy and non-healthy diseases. The proposed methodology achieved 99.58% accuracy and the processing time for a single plant leaf image is .016 s. Although the system can reach high accuracies, the system works with 27 classes only.

In (Abd Algani et al., 2023) a novel deep-learning model for plant disease classification is proposed. The proposed model is a combination of Ant Colony Optimization (ACO) and CNN. The research authors named it “ACO-NN”. In this work, ACO was used for feature extraction and CNN was used as a classifier. The proposed model is compared with other different deep learning models including CNN without any optimization. The higher accuracy (99.98%) that the proposed model achieved can be seen. However, the increase in accuracy of the proposed “ACO-CNN” over CNN is only around 0.01%. In (Vo et al., 2023) an ensemble model is built from two transfer learning models (EfficientNetB0 and MobileNetV2). The proposed Ensemble achieves 99.77% accuracy. The proposed work however introduces an ensemble of two models only. Our proposed work tests five different models and tests ten different ensembles.

In (Kaya and GÜrsoy, 2023) a novel multi-headed DenseNet-based architecture is proposed. The model relies on fusing Red-Green-Blue (RGB) and segmented images. The model achieves an average of 98.17% on 54,000 images with 38 different classes. Table 1 shows a tabular comparison among various research studies reported in this section.

Table 1
Related Work comparison.

ref	contribution	disadvantages	accuracy
(Anandhakrishnan and Jaisakthi, 2022)	a novel model for tomato leaves disease classification	model trained only on tomato leaves images from PlantVillage	98.4%
(Sangeetha and Rani, 2021)	new pre-trained models are used and compared	the proposed model is trained on tomato leaves images only	99.77%
(Zhang et al., 2018)	different architectures such as AlexNet, GoogLeNet, and ResNet are proposed for plant disease classification.	This work is trained on nine classes only	97.28%
(Ghazi et al., 2017)	utilizes three different models, namely, AlexNet, GoogLeNet, and VGGNet.	small accuracies compared to other work from literature	78.44% using VGGnet
(Pandian et al., 2019)	novel augmentation methods are proposed and tested on three deep learning models	low accuracy	91.5%
(Pandian et al., 2022)	a new dataset is developed and new data augmentation algorithms are tested	The system doesn't compare its work with related work using same dataset	99.9655%
(Sharma et al., 2023)	new deep lightweight CNN architecture227 (DLMC-Net) is proposed for multi-class classification228 of plant leaves.	The system achieves low accuracy	93%
(Tiwari et al., 2021)	a new dense CNN model is proposed for plant disease classification	the system works with 27 classes only	99.58%
(Abd Algani et al., 2023)	a novel deep-learning model for plant disease classification is proposed The proposed model is a combination of ACO and CNN. The research authors named it “ACO-NN”	However, the increase in accuracy of the proposed ACO-CNN over CNN is only around 0.01%.	99.98%
(Vo et al., 2023)	an ensemble model is built from two transfer learning models (EfficientNetB0 and MobileNetV2).	The proposed work however introduces an ensemble of two models only.	99.77%

3. Methodology

The approach proposed in this work can be explained as follows. First, the dataset is balanced using the class-weighted technique. This step is important as the deep learning models have a bias towards classes with a large number of images. This needs to be solved using a data balancing technique. Also, the data imbalanced in the proposed dataset is small hence a simple algorithm like class-weighted is proposed. Then a novel image pre-processing technique is proposed. A number of different models from the state-of-the-art deep learning models are tested. The results of the five best models are reported. Also, different ensemble models are tested and compared. The results show that ensemble models can increase the accuracy of the classification hence introducing more stable models.

3.1. Dataset description

This dataset is recreated using offline augmentation from the original dataset. The dataset that we used in this work is the new PlantVillage dataset. It is freely available online as a GitHub repository.[1] This dataset consists of about 87 K RGB images of healthy and sick crop leaves which are categorized into 38 different classes. The classes are based on different types of diseases that different plants can exhibit through their leaves. Table 2 shows the different classes of diseases that different plants can catch during their lifetime. The total dataset is divided into 80/20 ratio of training and validation set preserving the directory structure.

Fig. 1 shows some example images from the dataset along with their corresponding diseases. In this work a new model that can classify 38 type of plant disease is proposed.

[1] Web link: https://github.com/spMohanty/PlantVillage-Dataset

Fig. 1. Grid of images contains plants and their corresponding diseases.

3.2. Image pre-processing

This step is mainly about enhancing the input images for processing. This pre-processing helps to improve the quality of the images by removing undesirable distortion including spores, dust, etc. It could be used to adjust the images' colors too. The objective of this step is to provide clear images for further analysis. Basically, in this step, the input image is converted from the RGB image to L*a*b* color space (L* = Luminosity layer, a* & b* chromacity layers).

For image enhancement, different techniques for plant images, Histogram equalization, contrast limited adaptive histogram equalization (CLAHE), and gamma correction have been tested with Adaptive Median Filter (AMF), Median Filter (MF), Total Variation Filter (TVF) and Gaussian Denoising Filters (GDF). The experiments proved that CLAHE with AMF has the best performance in the image pre-processing step. Fig. 2 shows the image enhancement technique combining CLAHE with AMF. In the proposed technique the input image is converted from RGB to Lab color space at this step. Then we apply CLAHE algorithm on the L-channel. After that, the L-a-b channels are combined again. Eventually, the image is converted back to the RGB plane and different filtering techniques are applied to it (see Fig. 3).

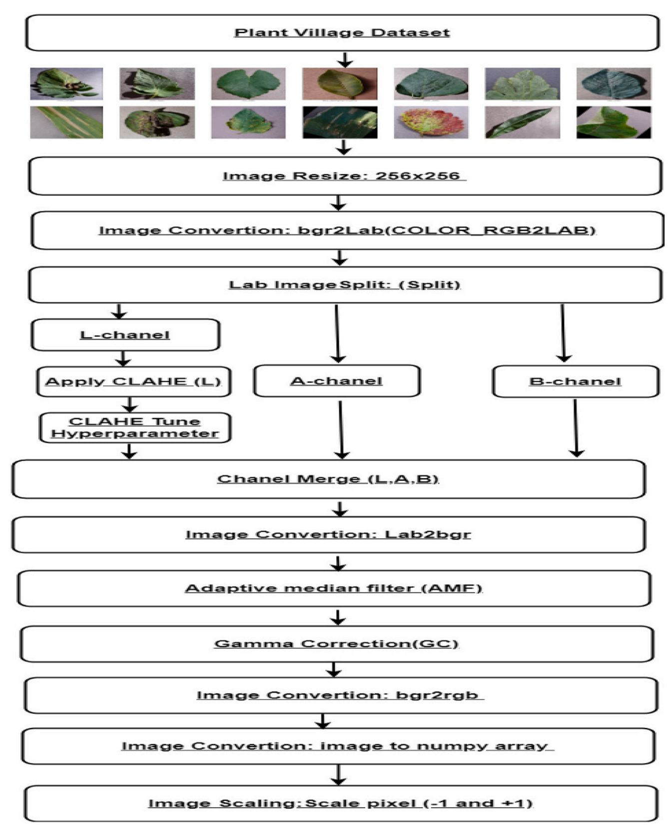

Fig. 2. Preprocessing step.

3.3. Data balancing

The PlantVillage dataset is an unbalanced dataset. This means that the classification system built on this dataset will have a bias towards some classes. However, the PlantVillage dataset is lightly imbalanced so the proposed data balancing technique is a class-weighted technique. This technique helps balance the training of the ML model with a small computational cost which is suitable for lightly imbalanced datasets.In the class-weighing technique, a higher weight is assigned to the minority classes during training by the ML algorithm. On the other hand, low weight is assigned to the majority classes. This makes the ML algorithm put more emphasis on learning the features of the minority class. This helps to improve model performance by giving more attention to

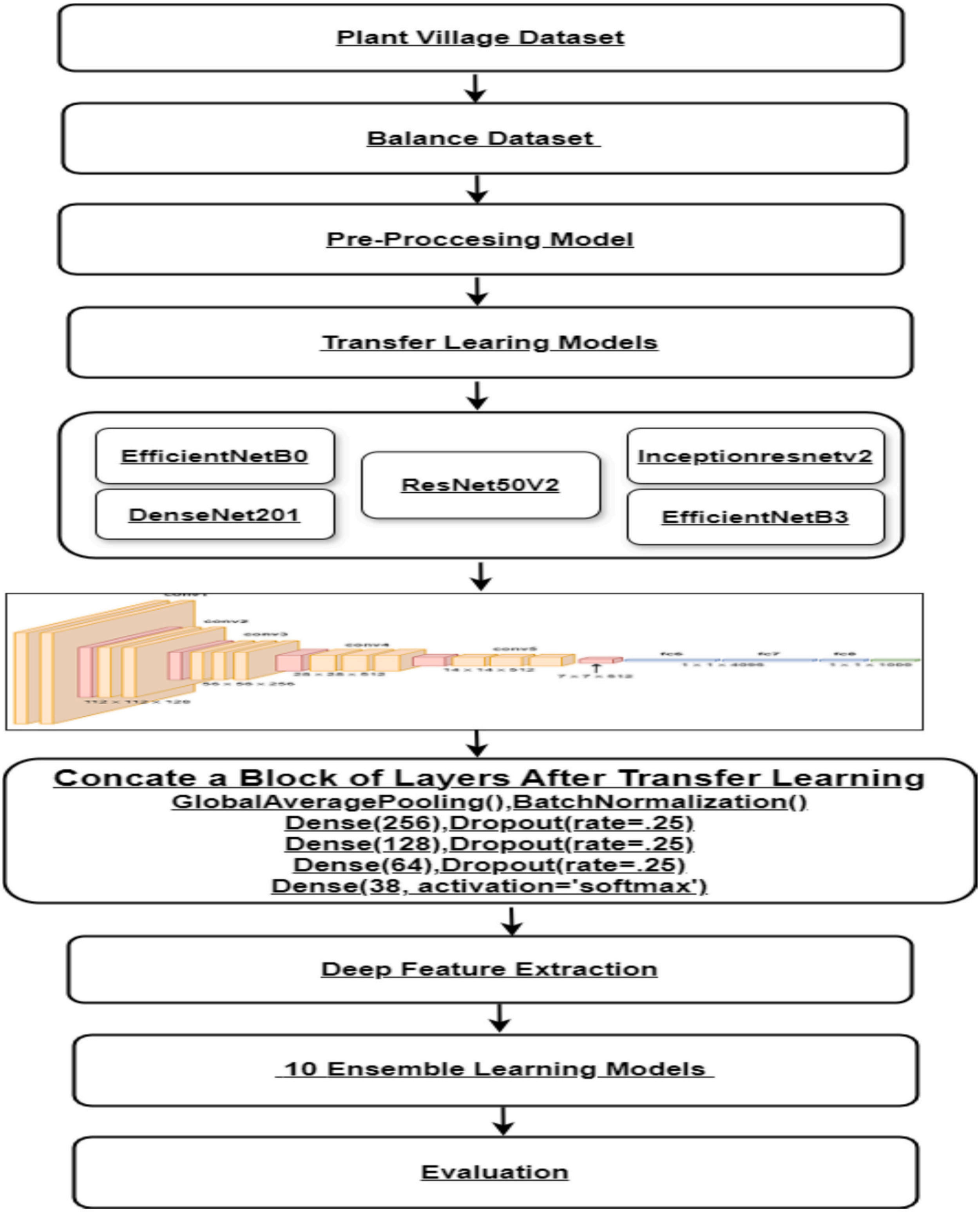

Fig. 3. The proposed model.

minority classes during training.

This is implemented in two steps. First we use a function called compute_class_weights from sklearn library in Python. This function is responsible of calculating the weights that should be given to each class depending on the class numbers.

The second step involves using these class weights as inputs to train the model. This is done by setting the class weight parameters using the previously calculated weights. The fit function inside sklearn is implemented in a way to accept such parameters and it applies them during training.

3.4. *Deep learning models*

In this work, we propose deep-learning architectures to classify plant diseases. We also propose using different ensembles to increase classification accuracy. To this end, five different deep-learning models are proposed to classify plant diseases. The chosen models are:

- DenseNet201
  DenseNet is a network proposed by (Huang et al., 2017). In this network, each Layer has a connection with subsequent Layers. Densenet201 is a CNN that has 201 deep layers.
- EfficientNetB0 EfficientNet is a CNN architecture and scaling method that uniformly scales all dimensions of the CNN. EfficientNetB0 (Tan and Le, 2019) is a network that is trained on more than a million images from the ImageNet database.
- EfficientNetB3
  EfficientNetB3 (Tan and Le, 2019) is another network of the EfficientNet family. EfficientNetB3 comes from scaling EfficientNetB0 which was generated using AutoML.
- InceptionResNetv2
  InceptionResNetv2 (Szegedy et al., 2017) is a CNN network that builds on the inception family of architectures. However, the model incorporates residual connections instead of the filter concatenation stage of the inception family.
- ResNet50v2
  (Gomez et al., 2017) ResNet50 is a deep learning architecture that is 50 layers deep. ResNet50v2 is also trained on a million images from the ImageNet dataset.

Different ensembles from these models are tested to achieve high accuracy of classification. The models proposed and ensembles are given in Table 4. The table also shows the accuracy reported for each ensemble model highlighting the highest accuracy.

3.5. Transfer learning

In this work, transfer learning is used to train the deep learning models. These deep-learning models are combined to form ensembles to achieve better classification accuracy. Transfer learning is a technique in ML that allows ML models trained on a certain dataset to be retrained on new data while preserving the knowledge gained from the old data (Weiss et al., 2016).

3.6. Ensemble models

As stated earlier, in this work, the use of an ensemble of models is proposed. In ensemble learning, several ML models are combined to reach higher accuracies from single models (Sagi and Rokach, 2018).

3.7. Performance metrics

In this work, we measure the accuracy of the deep-learning models by using the following metrics:

- Accuracy
  accuracy which is a simple metric for deep-learning classification models. This metric measures the overall percentage of correct predictions. However, accuracy does not take into account the different types of errors. That is the reason we employ other metrics for measuring the efficiency of deep learning models.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

- Precision
  Precision is a metric that measures the number of positive predictions that are actually correct viz. a viz. correct and incorrect positive predictions.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

- Recall
  recall is a metric that measures the number of positive predictions that are actually correct viz. a viz. correct positive predictions and incorrect negative predictions. Recall measures how many positive predictions were correctly spotted by the model.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

- F1-score
  F1-score is a metric that combines precision and recall in a single metric.

$$F1 - score = \frac{TP}{TP + .5(FP + TP)} \tag{4}$$

TP stands for the number of samples that were correctly predicted to be positive. FP stands for the number of samples that were falsely predicted to be positive. FN stands for the number of samples that were falsely predicted to be negative. TN stands for the number of samples that were correctly predicted to be negative.

4. Results and discussion

The main aim of this study was to build a new ensemble model for the automatic classification of plant leaf diseases. Building an automatic detection and classification system for plant diseases is a critical research area. These systems can save a substantial amount of wasted agricultural produce.

In literature, many researchers have focused on developing ML models for plant disease classification (Shrivastava and Pradhan, 2021). However, these ML systems require manual engineering of features and need a feature selection algorithm to achieve high accuracy. The application of deep-learning eliminates the need for feature engineering to solve such problems.

Along with the classification system, a novel image pre-processing algorithm is proposed in this research to increase the classification accuracy. The main steps in the image processing algorithm that contribute to image enhancement are the CLAHE algorithm applied to the L-channel of the image and the adaptive median filter. Also, a class-weighted method is used to solve the problem of imbalanced data found in the dataset. This balancing method is simple but yet necessary to train the model without bias towards certain classes and different models have been tested and evaluated in this work to achieve better accuracy. DenseNet201 (Huang et al., 2017), EfficientNetB0 (Tan and Le, 2019), InceptionResNetV2 (Szegedy et al., 2017), EfficientNetB3 (Tan and Le, 2019), and ResNet50v2 (Gomez et al., 2017) were the deep-learning models that were utilized in this work. Ten different deep learning models from the state-of-the-art deep learning models were initially experimented with. The results of the best five models are reported. Moreover, different ensemble learning models (chosen randomly) are utilized to get better accuracy. We developed a model that distinguishes between healthy and diseased crop leaves and, in the event that the crop has a disease, can identify which illness it is. To the best of the authors' knowledge, this is the first work that proposes an in-depth study on the use of different ensemble deep-learning models in solving the problem of plant disease classification. In this study, ten different ensembles are tested to evaluate their performance. Offline data balancing was applied to the source dataset. The dataset used in this work is made up of 38 distinct classes – shown in Table 2 – that have been created from around 87 K RGB images of both healthy and damaged crop leaves.

By removing unwanted distortion, such as spores and dust, this preprocessing aids in enhancing the image quality. The goal of this step is to

Table 2
Classes in the Dataset and Their Respective Numbers of Samples.

Type of diseases	NS
Tomato Late_blight	1851
Tomato healthy	1926
Grape healthy	1692
Orange Haunglongbing_(Citrus_greening)	2010
Soybeanhealthy	2022
Squash Powdery_mildew	1736
Potato healthy	1824
Corn_(maize) Northern_Leaf_Blight	1908
Tomato Early_blight	1920
Tomato Septoria_leaf_spot	1745
Corn_(maize) Cercospora_leaf_spot Gray_leaf_spot	1642
Strawberry Leaf_scorch	1774
Peach healthy	1728
Apple Apple_scab	2016
Tomato Tomato_Yello_Leaf_Curl_Virus	1961
Tomato Bacterial_spot	1702
Apple Black_rot	1987
Blueberry healthy	1816
Cherry_(including_sour) Powdery_mildew	1683
Peach Bacterial_spot	1838
Apple Cedar_apple_rust	1760
Tomato Target_Spot	1827
Pepper,bell healthy	1988
Grape Leaf_blight_(Isariopsis_Leaf_Spot)	1722
Potato Late_blight	1939
Tomato Tomato_mosaic_virus	1790
Strawberry healthy	1824
Apple healthy	2008
Grape Black_rot	1888
Potato Early_blight	1939
Cherry (including_sour)healthy	1826
Corn_(maize) Common_rust_	1907
GrapeEsca_(Black_Measles) _	1920
Raspberr healthy	1781
Tomato Leaf_Mold	1882
Tomato Spider_mites Two_spotted_spider_mite	1741
Pepper_bell Bacterial_spot	1913
Corn_(maize) healthy	1859

provide crystal-clear images for further examination. The added image enhancement technique enabled the proposed systems to reach high accuracy values.

This dataset is used after these pre-processing steps with the suggested models. Table 3 presents a comparison between different trained deep learning models proposed in this paper. The table shows the five different deep learning models chosen for this application and their reported accuracy and F1-score. The table also shows ten different ensembles chosen from combinations of the previous five models and their corresponding accuracy and F1-score. It can be seen from the table that InceptionResNetv2 model has the highest accuracy among the other deep learning models. Also, it can be seen that ResNet50v2 has the lowest accuracy of all the deep learning models. This is possibly the reason that the proposed ensemble model with higher accuracy doesn't include ResNet50v2 model. It is also clear that DenseNET, EfficiecntNetB0, EfficientNetB3 and InceptionResNetv2 have comparable accuracies. On the other hand, ResNet50v2 has a small accuracy compared to them. The table also shows the F1-score. It can be seen from the table that there are around six models that have an F1-score of 0.999. One of these models is the ensemble that has the best accuracy. The results in Figs. 7 and 6 show the accuracy and loss along with the epochs for the EfficientNetB3 model. Fast convergence is noticed in the model's response as it converges at the fourth epoch. Figs. 4 and 5 show the confusion matrix of classification of plant diseases by the ensemble model and by EfficientNetB3 respectively. A confusion matrix shows the number of correct classifications for each class in its diagonal cells. It also shows the number of misclassifications in other cells than the diagonal. It can be seen from the confusion matrix that EfficientnetB3 model misclassified class 19 with class 22 23 time. This classification error is removed with ensemble modelling. It can be seen from the confusion matrix of the ensemble model that there is a small number of misclassifications compared to the EfficientNetB3 model. Figue 5 shows a small number of misclassification. It can be seen that the model biggest error is a misclassifcations of class 7. These misclassifications are reduced from 6 using EfficientNetB3. This proves the robustness and efficiency gain of the ensemble model.

Table 3
Comparison between different trained deep learning models proposed.

model	Accuracy	F1-score
DenseNet201	0.9929	0.9948
efficientNetB0	0.9964	0.9979
efficientNetB3	0.9952	0.9979
Inceptionresnetv2	0.9972	0.9948
ResNet50v2	0.9764	0.9917
ensemble of EfficientNetB0, Inceptionresnetv2, resnet50v2	0.9978	0.9979
ensemble of EfficientNetB0, DesNet201	0.9977	0.9979
ensemble of EfficientNetB0, Desnet201, ResNet50V2	0.9981	0.9979
ensemble of EfficientNetB0, EfficientNetB3	0.9986	0.9990
ensemble of EffieicntNetB0, EffieicntNetB3, inceptionresnetv2, Desnet201, ResNet50V2	0.9986	0.9990
ensemble of EfficientNetB0, inceptionresnetv2, Desnet201	0.9986	0.9990
ensemble of EfficientNetB0, inceptionresnetv2, Desnet201, EfficientNetB3	0.9989	0.9990
ensemble of EfficientNetB0, inceptionresnetv2,Desnet201, ResNet50V2	0.9987	0.9990
ensemble of EfficientNetB0, inceptionresnetv2, EfficientNetB3	0.9987	0.9979
ensemble of inceptionresnetv2 + Desnet201 + ResNet50V2	0.9978	0.9990

The highest accuracy is achieved by using an ensemble model combining four deep-learning models (i.e., EfficientNetB0, InceptionResNetv2, Desnet201, EfficientNetB3). The accuracy obtained with this ensemble is 99.89%. In our review of the literature related to our work, most of the works have targeted small data sets with a single type of plants (Salamai, 2023). This dataset is widely used in research and is known for its challenging images. The ability of our models to classify unseen data with high accuracy proves its ability to generalize. The work presented here, however, targets a general and a much broader dataset for plant disease (i.e., new PlantVillage). Moreover, we have noticed that most of the work on the general dataset uses some deep-learning models or ensembles. This work provides an in-depth study of five different deep-learning models along with ten different ensembles of these models and provides a comparison between their performance. One limitation of this work is that proposing to work with an ensemble of deep learning models needs computational power which may not be available in smartphones for example. In future work, the authors propose investigating light models using the same novel image-processing algorithm. One other possible future work for this paper is to add explainability to the system. Explainability will help the user be more confident in the system.

5. Comparison with the state-of-the-art

In this section, a comparison with the State-of-art algorithms is presented. The main metric used in the comparison is the accuracy of the models. As accuracy is a simple metric, it is easy to interpret. Table 4 shows the comparison table. In (Gokulnath et al., 2021) a novel loss-fused CNN model is proposed that can classify plant diseases. The model takes advantage of two loss functions. The system was tested on PlantVillage dataset. The system achieved 98.9% accuracy and F-score of 96.65%. Our proposed system using augmented data and the proposed image processing technique achieves F-score of 99.9%.

In (Mohanty et al., 2016) the authors proposed using DCNNs to classify different plant diseases. The proposed model achieved 99.31% accuracy on PlantVillage dataset. Our proposed model using our proposed image processing algorithm was able to achieve 99.89% accuracy.

In (Atila et al., 2021) a plant disease classification system was built.

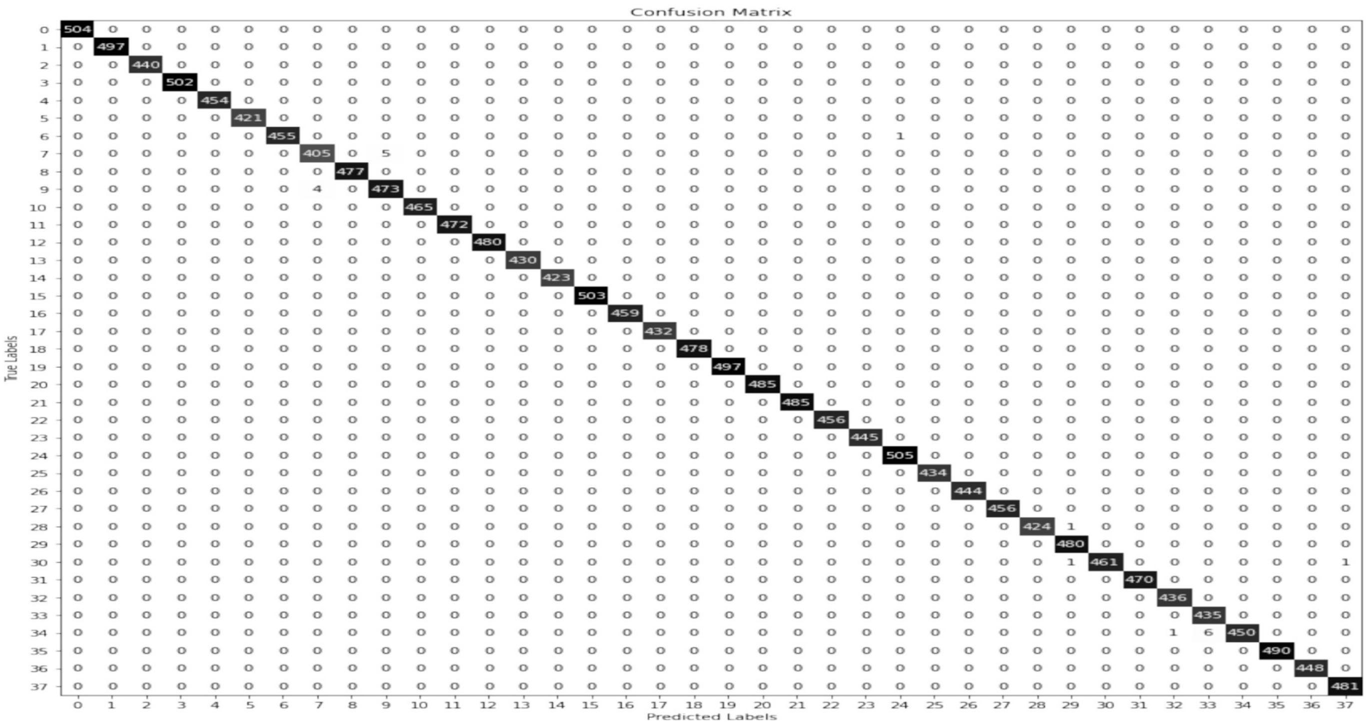

Fig. 4. Confusion Matrix related to the proposed ensemble.

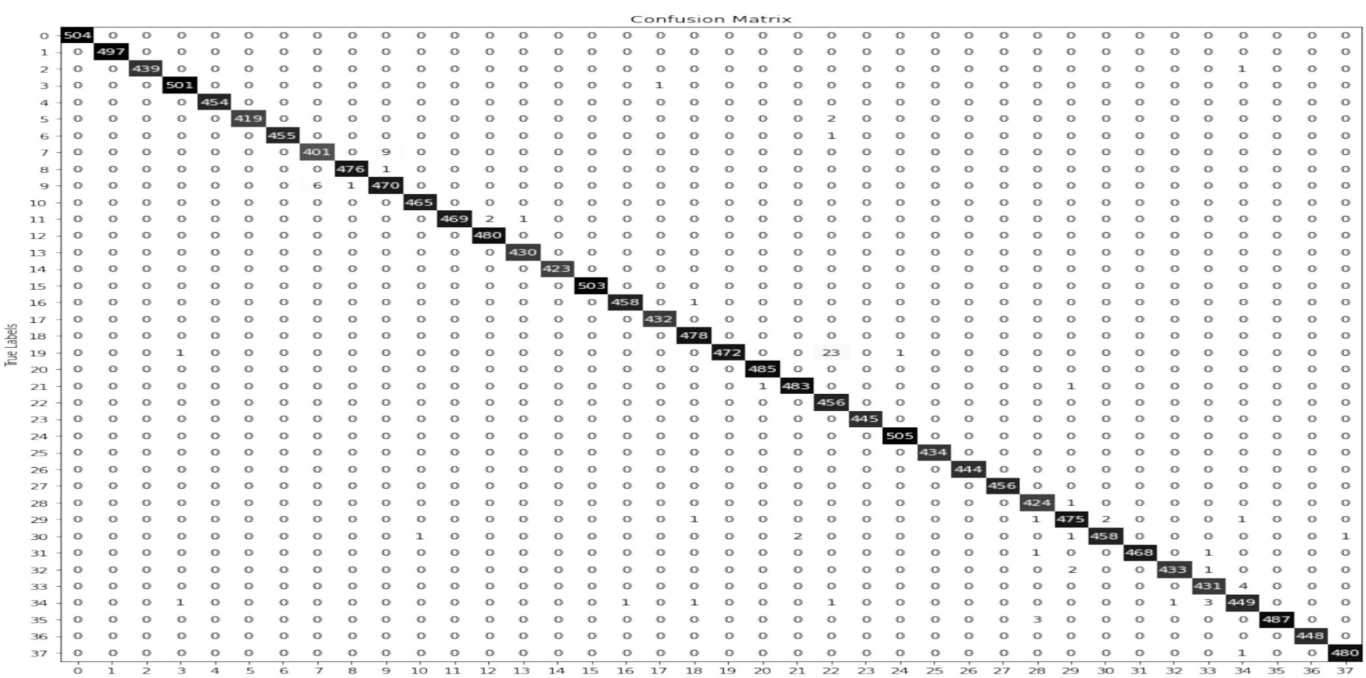

Fig. 5. Confusion Matrix of classification of plant diseases by EfficientNetB3.

Different deep learning models are tested to find the model with the highest accuracy. The results show that EfficientNet achieves the highest accuracy. A data augmentation technique is added to increase the number of images from 55,448 to 61,486 images. Results of the proposed deep learning models on the dataset with and without augmentation are reported. This work augments the dataset using a complex data augmentation technique to produce a new dataset. However, in our proposed work we use an augmented version of PlanVillage that can be used by different researchers for comparison. Also, we propose using a balancing method to balance the dataset. In (Too et al., 2019) a comparative study on applying different deep learning algorithms on the PlantVillage dataset is proposed. The proposed methodology achieved 99.75% accuracy on the PlantVillage dataset proposed in this work. However, our proposed ensemble using the augmented dataset achieves higher accuracy.

In (Ashwinkumar et al., 2022) a novel optimal MobileNet model is proposed to solve the problem of plant disease classification. The proposed model and preprocessing techniques are tested on a tomato

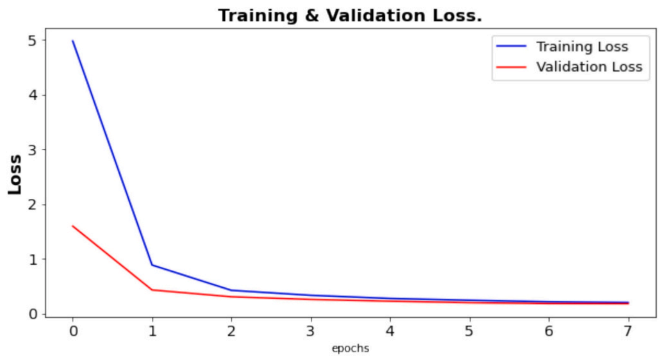

Fig. 6. Progression of loss of the deep learning model (EfficientNetB3) as a function of epochs.

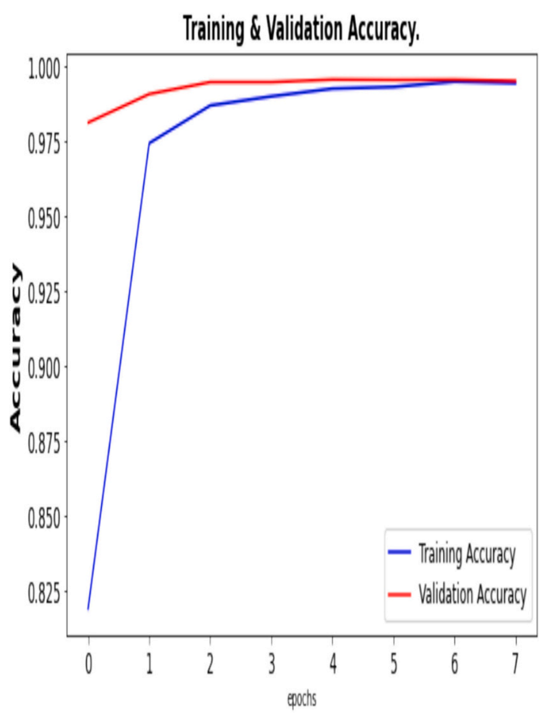

Fig. 7. Accuracy of the deep learning model (EfficientNetB3) as a function of epochs.

Table 4
Comparison between proposed model and previous models from literature.

reference	Dataset	accuracy	contribution
(Too et al., 2019)	PlantVillage	99.75%	DenseNets compared to other DL models
(Gokulnath et al., 2021)	PlantVillage	98.93%	loss-fused CNN
(Atila et al., 2021)	PlantVillage	99.91%	augmentation the dataset to build bigger dataset
(Vo et al., 2023)	PlantVillage	99.7%	building an ensemble of two deep ML models
(Kaya and GÜrsoy, 2023)	PlantVillage	98.17%	DenseNet DL network and fusing segmented images with RGB
(Mohanty et al., 2016)	PlantVillage	99.31%	DCNN
Proposed	New PlantVillage	99.89%	proposed novel Digital Image Processing (DIP) and Ten ensemble models

dataset which is a much less challenging dataset than PlantVillage. The proposed model was able to achieve only 98.7% accuracy. In comparison, our proposed model was able to achieve 99.89% accuracy on a much more challenging dataset.

In (Abd Algani et al., 2023) a novel deep learning model is introduced to solve the plant disease classification problem. The proposed algorithm combines ACO and CNN to achieve higher accuracy from the classification model. It can be seen from the results that this model only achieves 0.01% increase in accuracy over regular CNN models. However, our proposed image processing technique along with data balancing makes our models achieve comparable accuracies. This is achieved without the need to implement any additional optimization algorithms. Our proposed ensemble increases the accuracy by about 0.2% above the highest model accuracy. In (Vo et al., 2023) An ensemble model is built from two transfer learning models (EfficientNetB0 and MobileNetV2). The proposed models were trained on PlantVillage Dataset which contains 54,305 images from 38 different plant disease classes. The proposed ensemble achieves 99.77%. The work introduced did not propose a new image processing technique like the one proposed in this work. The previous paper only tests one ensemble combination, while our work tests more than ten ensemble combinations to find the best ensemble.

6. Conclusion

Plant disease identification and classification are now frequently based on deep earning techniques. In this paper, a novel methodology for addressing the plant disease classification problem is proposed. The paper proposes a novel image processing algorithm (CLAHE with AMF) that was developed based on extensive testing of different image processing techniques. The paper proposes a certain type of data balancing (class weight) algorithm. This suggestion is considered novel for this application. Also, in this work, we propose testing five different state-of-the-art deep learning models and their ensembles based on previously proposed algorithms and report their results. First, the dataset is balanced using the class weighting algorithm. This step is important since the dataset is slightly imbalanced which will bias the classification. Then a novel image pre-processing algorithm is applied to enhance images and increase efficiency. As discussed earlier this image processing technique is essential for enhancing the image quality and hence for increasing the model accuracy. After that, a number of pre-trained models are tuned with the cited dataset. Lastly, a number of ensembles of these models are tested and compared to find the best ensemble. Future work of this research shall be further testing other deep learning architectures or ensembles on more challenging datasets. Future work may also include training a deep learning model from scratch and comparing the results with the proposed work. It may also include testing the fusion between handcrafted and non-handcrafted features to increase the model's accuracy. Another extension for this work shall be the introduction of explainability for the model results.

CRediT authorship contribution statement

Ali Hussein Ali: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ayman Youssef:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Investigation, Conceptualization. **Mahmoud Abdelal:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Muhammad Adil Raja:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Data availability

We used the famous PlantVillage dataset for this work and we have provided a hyperlink for this open source dataset.

Acknowledgments

This research was partly funded by the Technological University Transfer Fund (TUTF) of the Higher Education Authority (HEA) of Ireland.

References

Abd Algani, Y.M., Caro, O.J.M., Bravo, L.M.R., Kaur, C., Al Ansari, M.S., Bala, B.K., 2023. Leaf disease identification and classification using optimized deep learning. Measurement 25, 100643.
Anandhakrishnan, T., Jaisakthi, S., 2022. Deep convolutional neural networks for image based tomato leaf disease detection. Sustain. Chem. Pharm. 30, 100793.
Ashwinkumar, S., Rajagopal, S., Manimaran, V., Jegajothi, B., 2022. Automated plant leaf disease detection and classification using optimal mobilenet based convolutional neural networks. Mater. Today Proc. 51, 480–487.
Atila, Ü., Uçar, M., Akyol, K., Uçar, E., 2021. Plant leaf disease classification using efficientnet deep learning model. Eco. Inform. 61, 101182.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.
Geetharamani, G., Pandian, A., 2019. Identification of plant leaf diseases using a nine-layer deep convolutional neural network. Comput. Electr. Eng. 76, 323–338.
Ghazi, M.M., Yanikoglu, B., Aptoula, E., 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing 235, 228–235.
Gokulnath, B., et al., 2021. Identifying and classifying plant disease using resilient lf-cnn. Eco. Inform. 63, 101283.
Gomez, A.N., Ren, M., Urtasun, R., Grosse, R.B., 2017. The reversible residual network: backpropagation without storing activations. Adv. Neural Inf. Proces. Syst. 30.
Hartigan, J.A., Wong, M.A., 1979. Algorithm as 136: a k-means clustering algorithm. J. R. Stat. Soc.: Ser. C: Appl. Stat. 28 (1), 100–108.
He, D.-C., He, M.-H., Amalin, D.M., Liu, W., Alvindia, D.G., Zhan, J., 2021. Biological control of plant diseases: an evolutionary and eco-economic consideration. Pathogens 10 (10), 1311.
Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708.
Iqbal, Z., Khan, M.A., Sharif, M., Shah, J.H., ur Rehman, M. H., and Javed, K., 2018. An automated detection and classification of citrus plant diseases using image processing techniques: a review. Comput. Electron. Agric. 153, 12–32.
Jain, A.K., Mao, J., Mohiuddin, K.M., 1996. Artificial neural networks: a tutorial. Computer 29 (3), 31–44.
Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90.
Kaya, Y., GÜrsoy, E., 2023. A novel multi-head cnn design to identify plant diseases using the fusion of rgb images. Eco. Inform. 75, 101998.
Madhogaria, S., Schikora, M., Koch, W., Cremers, D., 2011. Pixel-based classification method for detecting unhealthy regions in leaf images. In: GI-Jahrestagung, p. 482.
Marr, D., Hildreth, E., 1980. Theory of edge detection. Proc. Royal Soc. Lond. Ser. B Biol. Sci. 207 (1167), 187–217.
Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7, 1419.
Mzoughi, O., Yahiaoui, I., 2023. Deep learning-based segmentation for disease identification. Eco. Inform. 75, 102000.
Oerke, E.-C., 2006. Crop losses to pests. J. Agric. Sci. 144 (1), 31–43.
Pandian, J.A., Geetharamani, G., Annette, B., 2019. Data augmentation on plant leaf disease image dataset using image manipulation and deep learning techniques. In: 2019 IEEE 9th International Conference on Advanced Computing (IACC). IEEE, pp. 199–204.
Pandian, J.A., Kumar, V.D., Geman, O., Hnatiuc, M., Arif, M., Kanchanadevi, K., 2022. Plant disease detection using deep convolutional neural network. Appl. Sci. 12 (14), 6982.
Peterson, L.E., 2009. K-nearest neighbor. Scholarpedia 4 (2), 1883.
Sagi, O., Rokach, L., 2018. Ensemble learning: a survey. Wiley Interdiscip. Rev. 8 (4), e1249.
Salamai, A.A., 2023. Enhancing mango disease diagnosis through eco-informatics: a deep learning approach. Eco. Inform. 77, 102216.
Sangeetha, R., Rani, M.M.S., 2021. Tomato leaf disease prediction using transfer learning. In: Advanced Computing: 10th International Conference, IACC 2020, Panaji, Goa, India, December 5–6, 2020, Revised Selected Papers, Part II 10. Springer, pp. 3–18.
Sankaran, S., Mishra, A., Ehsani, R., Davis, C., 2010. A review of advanced techniques for detecting plant diseases. Comput. Electron. Agric. 72 (1), 1–13.
Sharma, V., Tripathi, A.K., Mittal, H., 2023. Dlmc-net: deeper lightweight multi-class classification model for plant leaf disease detection. Eco. Inform. 75, 102025.
Shen, D., 2007. Image registration by local histogram matching. Pattern Recogn. 40 (4), 1161–1172.
Shijie, J., Ping, W., Peiyi, J., Siping, H., 2017. Research on data augmentation for image classification based on convolution neural networks. In: 2017 Chinese Automation Congress (CAC). IEEE, pp. 4165–4170.
Shrivastava, V.K., Pradhan, M.K., 2021. Rice plant disease classification using color features: a machine learning paradigm. J. Plant Pathol. 103, 17–26.
Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 31.
Tan, M., Le, Q., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning. PMLR, pp. 6105–6114.
Tiwari, V., Joshi, R.C., Dutta, M.K., 2021. Dense convolutional neural networks based multiclass plant disease detection and classification using leaf images. Eco. Inform. 63, 101289.
Too, E.C., Yujian, L., Njuki, S., Yingchun, L., 2019. A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. 161, 272–279.
Vo, H.-T., Quach, L.-D., Hoang, T.N., 2023. Ensemble of deep learning models for multi-plant disease classification in smart farming. Int. J. Adv. Comput. Sci. Appl. 14 (5).
Weiss, K., Khoshgoftaar, T.M., Wang, D., 2016. A survey of transfer learning. J. Big Data 3 (1), 1–40.
Xu, X., Xu, S., Jin, L., Song, E., 2011. Characteristic analysis of otsu threshold and its applications. Pattern Recogn. Lett. 32 (7), 956–961.
Zadeh, L.A., 1988. Fuzzy logic. Computer 21 (4), 83–93.
Zhang, X., Qiao, Y., Meng, F., Fan, C., Zhang, M., 2018. Identification of maize leaf diseases using improved deep convolutional neural networks. Ieee Access 6, 30370–30377.