Taylor and Francis Book Chapter

LOC Page

Contents

CHAPTER 1

Large Ensemble of Transfer-Learned Models for Plant Disease Recognition from Diverse Leaf Images

Maria Chowdhury

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: mariachowdhury831@gmail.com

Azmain Yakin Srizon

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: azmainsrizon@gmail.com ***(Corresponding Author)***

Tahsen Islam Sajon

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: sajon.tahsen@gmail.com

Md. Farukuzzaman Faruk

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: faarukuzzaman@gmail.com

S. M. Mahedy Hasan

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: mahedycseruet@gmail.com

Ananya Sarker

Bangladesh Army University of Engineering & Technology, Dept. of Computer Science & Engineering, Natore, Bangladesh. Email: ananya.ruet@gmail.com

Md. Al Mamun

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: almamun00350@gmail.com

Md. Rakibul Haque

Rajshahi University of Engineering & Technology, Dept. of Computer Science & Engineering, Rajshahi, Bangladesh. Email: rakibulhaq56@gmail.com

Md. Rakib Hossain
Rajshahi University of Engineering & Technology, Dept. of Electronics & Telecommunication Engineering, Rajshahi, Bangladesh. Email: rakib.ete12.ruet@gmail.com

CONTENTS

PLant diseases caused by microorganisms like fungi, bacteria, and insects can have devastating effects on crop productivity and quality. Detecting and identifying plant diseases is essential for effective management and sustainable agriculture. Computer vision-based image analysis employing deep learning techniques has emerged as a promising solution for efficient and timely disease identification in crops. This study aimed to enhance the classification performance of plant leaf disease images from the PlantDoc dataset by employing an ensemble CNN technique. Seven pre-trained CNN models were trained on the dataset, and a SoftMax Averaging ensemble technique was used to achieve superior plant disease classification performance. The proposed approach outperformed previous studies by a great margin, achieving an accuracy of 98.66%. The SoftMax Averaging ensemble technique allowed for a more diverse

range of feature extraction and improved classification accuracy which can effectively address the challenge of plant disease classification.

1.1 INTRODUCTION

Plant diseases refer to ailments that affect the health and productivity of crops, often caused by insidious microorganisms like fungi, bacteria, and insects. These diseases can result in devastating consequences such as yield losses and reduced crop quality[10]. Timely detection and accurate identification are crucial for effective management [12]. Globally conducted surveys have shed light on the pervasive impact of plant diseases on agricultural production. Despite farmers contributing over 80% of the world's agricultural output, more than 50% of this output is lost due to pests and diseases, resulting in disruptions to the food supply and widespread food insecurity. The effects of climate change have further exacerbated this situation, with plant pests causing increasing damage to crucial cash crops, posing threats to both food security and the environment. The Food and Agricultural Organization (FAO) estimates that up to 40% of the world's crop yield is lost to pests annually, resulting in losses of at least $70 billion and an annual economic loss of more than $220 billion due to plant diseases [11].

Conventional methods of manually surveying fields and employing traditional strategies to control plant diseases have proven to be inefficient and time-consuming, often resulting in delays in disease analysis. Additionally, visual analysis by qualified plant pathologists can introduce subjective biases into the process. For expeditious and efficient disease identification in crops, computer vision-based image analysis employing cutting-edge techniques has emerged as a promising solution to these restrictions [14]. Previous studies in the field of plant disease detection relied on methods such as shape, color, and texture feature extraction along with traditional machine learning classifiers. However, more recent studies have extensively utilized deep learning techniques, with convolutional neural networks (CNNs) showing promising results [13].

In light of the aforementioned considerations, this research endeavours to address the challenge of plant disease classification through the application of advanced deep learning networks. The main goal of this study is to enhance the classification performance of plant leaf disease images from the PlantDoc dataset by employing an ensemble CNN technique. To achieve this, seven cutting-edge pre-trained CNN models, namely, DenseNet201, Xception, EfficientNetB6, InceptionResNetV2, InceptionV3, ResNet50, and MobileNetV2, were selected and individually trained on the PlantDoc dataset [2]. Subsequently, these models were utilized for a SoftMax Averaging ensemble technique, aiming to achieve superior plant disease classification performance. We perform a detailed analysis of our model's performance, including a comparison with other state-of-the-art methods, and provide insights into the strengths and limitations of our proposed approach.

1.2 LITERATURE REVIEW

Previous research conducted by Saanidhya et al. employed deep learning techniques, training six standard pre-trained networks using a modified PlantVillage dataset [15]. Aanis et al. divulged intriguing insights into the accuracy of the Xception model when subjected to images from the PlantDoc dataset achieving a testing accuracy of 81.53%. However, during subsequent generalization tests, the accuracy dropped to 61%, underscoring the need for further improvements in model performance [14]. Sumaiya et al. embarked on a comprehensive evaluation of three distinct CNN image classifiers—Mobilenet, InceptionResnetV2, and Xception—to ascertain their performance on the PlantVillage dataset [17]. Xin et al. opted for ResNet50, MobileNetV2, and DenseNet121 as their models of choice, with DenseNet121 demonstrating superior performance by achieving an accuracy of 86.36% on the PlantDoc dataset [20]. Thipwimon et al. trained and evaluated five CNN models, including MobileNetV1, MobileNetV2, NASNetMobile, DenseNet121, and Xception, on multiple leaf disease datasets [18]. Thararat et al. analyzed multiple deep learning methods, including VGG16, MobileNetV2, InceptionResNetV2, and DenseNet201, on the PlantDoc dataset for plant leaf and plant leaf disease image classification obtaining an accuracy of 69.51% [16].

It is important to note that these groundbreaking works are not without limitations. They are confined either to specific crops or datasets that are not publicly available. As of now, the PlantVillage dataset (PVD) and the PlantDoc dataset [2] remain the only publicly accessible datasets for plant disease detection. However, it is worth mentioning that the PlantVillage dataset comprises images taken in controlled laboratory setups, whereas the PlantDoc dataset contains authentic images captured from real-world cultivation fields, thus lending greater credibility to its efficacy in practical applications.

1.3 MATERIALS AND METHODS

In this section, dataset preparation and proposed methodologies have been discussed.

1.3.1 Dataset Selection and Description

Researchers have produced numerous datasets related to plant diseases over the years. One of the most well-known is the PlantVillage dataset [1]. However, the images in the PlantVillage dataset were taken in a controlled laboratory setting, making it easier for transfer-learned approaches to achieve high accuracy. In contrast, obtaining laboratory-quality images in real-world scenarios is often impractical due to various device and human limitations. Therefore, for this study, we chose the PlantDoc dataset, which features real-life leaf images with high diversity [2].

The PlantDoc dataset is an object detection and image classification dataset containing 2,548 images of 13 different plant species, each classified as either diseased or healthy, resulting in a total of 27 classes. The dataset was originally published by researchers at the Indian Institute of Technology. One example from each class has

Figure 1.1 Image samples for each of the classes of PlantDoc dataset

Figure 1.2 10 diverse image samples from the ‘Apple Leaf’ class

been showcased in Figure 1.1. Table 1.1 shows distribution of classes in the PlantDoc dataset i.e., number of sample images in each of the classes.

1.3.1.1 Challenges of the Dataset

The PlantDoc dataset is known for its diversity as each class contains a wide variety of leaf images. The variations in the images make it difficult for the CNNs to generalize their learned features to new, unseen data. Moreover, differentiating between leaves of same class can also be challenging as the appearance of leaves can vary significantly for a single plant. As a result, developing a model that can accurately classify the images in the PlantDoc dataset is a challenging task. The diversity of the “Apple Leaf” class has been illustrated in Figure 1.2.

1.3.1.2 Augmentation

In this study, three types of augmentation techniques were applied, including zoom random function, rotation, and horizontal and vertical flip, with a probability of 70%. Zoom random function was used with a maximum factor of 1.3 to randomly zoom into the image. The maximum rotation applied was 10 degrees to the left or right. Horizontal and vertical flip were used to create mirror images of the original image. Only the training set underwent augmentation, generating 300 images per class. From this, 240 images per class were selected to form the training set, while the remaining 60 images were placed in the validation set for each class.

Table 1.1 Distribution of classes in the PlantDoc dataset i.e., number of sample images in each of the classes

Serial	Name of the Class	Number of Samples
1.	Apple Leaf	91
2.	Apple Rust Leaf	89
3.	Apple Scab Leaf	87
4.	Bell Pepper Leaf	61
5.	Bell Pepper Leaf Spot	71
6.	Blueberry Leaf	116
7.	Cherry Leaf	57
8.	Corn Gray Leaf Spot	67
9.	Corn Leaf Blight	192
10.	Corn Rust Leaf	113
11.	Grape Leaf	69
12.	Grape Leaf Black Rot	64
13.	Peach Leaf	112
14.	Potato Leaf Early Blight	117
15.	Potato Lead Late Blight	105
16.	Raspberry Leaf	116
17.	Soyabean Leaf	65
18.	Squash Powdery Mildew Leaf	129
19.	Strawberry Leaf	96
20.	Tomato Early Blight Leaf	83
21.	Tomato Leaf	62
22.	Tomato Leaf Bacterial Spot	107
23.	Tomato Leaf Late Blight	111
24.	Tomato Leaf Mosaic Virus	54
25.	Tomato Leaf Yellow Virus	75
26.	Tomato Mold Leaf	91
27.	Tomato Septoria Leaf Spot	148

1.3.2 Transfer Learning

Transfer learning is a machine learning technique where a pre-trained model is used as a starting point for a new task. Transfer learning is accomplished by freezing the pre-trained layers and adding a new classification layer to the network. The output of the last convolutional layer is flattened and fed into the new layer. The weights of the pre-trained layers are kept fixed, and only the weights of the new classification layer are trained on the new dataset. This approach helps to overcome the overfitting problem and can achieve high accuracy with less training data.

1.3.2.1 Justification of Utilizing Transfer Learning

In the context of plant disease recognition from leaf images, transfer learning allows the use of pre-trained convolutional neural networks (CNNs) on large image datasets like ImageNet to extract features from the images, instead of training a CNN from scratch on a small dataset. This approach is particularly useful in this case where the dataset is both diverse and similar to the original dataset that the pre-trained model was trained on.

1.3.3 Utilized Transfer Learned Models

This research utilized seven transfer learning models, including DenseNet201, Xception, EfficentNetB6, InceptionResNetV2, InceptionV3, ResNet50, and MobileNetV2, to improve the accuracy of plant disease recognition from leaf images.

1.3.3.1 DenseNet201

The architecture of DenseNet201 is based on the idea of dense connectivity, where each layer is connected to every other layer in a feedforward fashion [3]. This connectivity pattern enables feature reuse and promotes feature propagation throughout the network. The equation for the output of a dense block in DenseNet201 can be written as:

$$\mathbf{H}_i = [\mathbf{x}_0, \mathbf{x}_1, \cdots, \mathbf{x}_{i-1}] \tag{1.1}$$

where $\mathbf{x}_i$ is the output of the i^{th} layer in the dense block and $[\cdot]$ denotes concatenation. The equation for the output of the transition layer can be written as:

$$\mathbf{H}_{i+1} = \mathrm{BN}(\mathrm{Conv}(\mathrm{Pool}(\mathbf{H}_i))) \tag{1.2}$$

where BN is batch normalization, Conv is convolution, Pool is pooling, and $\mathbf{H}_{i+1}$ is the output of the transition layer.

1.3.3.2 Xception

The core idea behind Xception is to split the depthwise separable convolution into two separate layers: one for spatial convolution and another for channel convolution [4]. The key equation for Xception is the depthwise separable convolution, which can be defined as:

$$Y_{i,j,k} = \sum_{r,s} X_{i+r,j+s,k} \cdot K_{r,s} \tag{1.3}$$

where X is the input tensor, K is the kernel, and Y is the output tensor.

1.3.3.3 EfficentNetB6

EfficientNetB6 is designed to have an optimal balance between accuracy and computational efficiency, achieved by scaling the depth, width, and resolution of the network in a principled manner [5]. The equations used in EfficientNetB6 are related to scaling the network architecture:

$$d = \alpha^{\phi}, w = \beta^{\phi}, r = \gamma^{\phi} \tag{1.4}$$

where d is the depth, w is the width, r is the resolution, α, β, and γ are scaling coefficients, and ϕ is the scaling parameter that controls the overall size of the network.

1.3.3.4 InceptionResNetV2

The architecture of InceptionResNetV2 includes multiple Inception modules, which are composed of parallel branches with different filter sizes, followed by a pooling layer and concatenation of the output feature maps [6]. In addition, residual connections are used to improve the flow of information between layers and mitigate the problem of vanishing gradients. Suppose, the input to the network is x, and the output is y. The residual connection for a layer i can be written as:

$$r_i(x) = x - f_i(x) \tag{1.5}$$

where f_i is the output of layer i. The Inception module output can be defined as:

$$Inception(x) = Concatenate([branch_1(x), branch_2(x), ..., branch_n(x)]) \tag{1.6}$$

where $branch_n$ is the output of the nth branch in the Inception module. The final output of the network is obtained through a global average pooling layer followed by a softmax activation function:

$$y = Softmax(AvgPool(x)) \tag{1.7}$$

1.3.3.5 InceptionV3

InceptionV3 uses a series of convolutional layers with varying filter sizes to capture both fine-grained and high-level features in the image [7]. The equation for the inception module is $Inception(x) = concat(conv(x, 1 \times 1), conv(x, 3 \times 3), conv(x, 5 \times 5), maxpool(x, 3 \times 3, stride = 1, padding = 1))$, where x is the input to the module, $conv(x, k \times k)$ is a convolutional layer with kernel size $k \times k$, and $maxpool(x, k \times k, stride, padding)$ is a max pooling layer with kernel size $k \times k$, stride $stride$, and padding $padding$. The concatenated output is then passed through a nonlinearity, such as a rectified linear unit (ReLU), before being used in subsequent layers of the network.

1.3.3.6 ResNet50

ResNet50 is a variant of ResNet that consists of 50 layers, and it uses skip connections to enable the training of very deep networks [8]. The residual block of ResNet50 can be represented mathematically as:

$$\mathbf{x}_l = \sigma\left(\mathbf{x}_{l-1} + \mathcal{F}\left(\mathbf{x}_{l-1}; \mathbf{W}_l\right)\right) \tag{1.8}$$

where σ is the ReLU activation function, $\mathbf{x}_{l-1}$ is the input to the l-th layer, $\mathcal{F}$ is the residual function that computes the layer's output, $\mathbf{W}_l$ are the layer's weights, and $\mathbf{x}_l$ is the output of the layer. The skip connection in ResNet50 is represented by the addition of $\mathbf{x}_{l-1}$ and $\mathcal{F}(\mathbf{x}_{l-1}; \mathbf{W}_l)$, which allows the gradients to flow directly through the network.

1.3.3.7 MobileNetV2

MobileNetV2 is a lightweight deep neural network architecture designed for efficient mobile and embedded vision applications [9]. The depthwise convolution applies a single filter to each input channel separately, and the pointwise convolution combines the outputs of the depthwise convolution with a linear transformation:

$$y = PW(DW(x)) \tag{1.9}$$

where x is the input feature map, DW is the depthwise convolution with a depthwise filter of size $K \times K$, PW is the pointwise convolution with 1×1 filters, and y is the output feature map.

1.3.3.8 Justification of Choosing Models

The selection of these seven models for plant disease recognition from leaf images is justified based on their strengths in computer vision tasks and their ability to extract relevant features from the input images. DenseNet201 is a deep neural network that excels at learning feature representations in densely connected blocks, making it highly effective for image classification. Xception is another deep neural network that uses depthwise separable convolutions to capture spatial dependencies between features, which leads to better accuracy and fewer parameters.

EfficientNetB6 is a scalable convolutional neural network that achieves state-of-the-art accuracy with fewer parameters and less computational cost than other models. InceptionResNetV2 combines the strength of both Inception and ResNet architectures by using residual connections and multiple parallel convolutional layers. InceptionV3 is known for its effectiveness in image classification and feature extraction, as it uses a combination of convolutional layers with different filter sizes to capture different levels of detail. ResNet50 is a widely used neural network that uses residual connections to overcome the problem of vanishing gradients and improve accuracy in deep networks. MobileNetV2 is a lightweight neural network that uses depthwise separable convolutions to achieve high accuracy with fewer parameters and lower computational cost than other models.

Each of these models has different strengths in feature extraction and image classification, and by using them in combination, we were able to achieve high accuracy in plant disease recognition from leaf images. The features extracted by each model were combined using an ensembled approach for final classification. The use of these models and their extracted features highlights the importance of selecting appropriate models and feature extraction techniques for image recognition tasks.

1.3.4 Proposed Ensembling Method

Ensemble methods combine multiple models to achieve better accuracy and robustness in prediction. SoftMax Averaging is a popular ensemble method that combines the SoftMax outputs of multiple models to make predictions. In this study, the SoftMax Averaging ensemble method has been used to combine the predictions of seven different models, namely DenseNet201, Xception, EfficentNetB6, InceptionResNetV2,

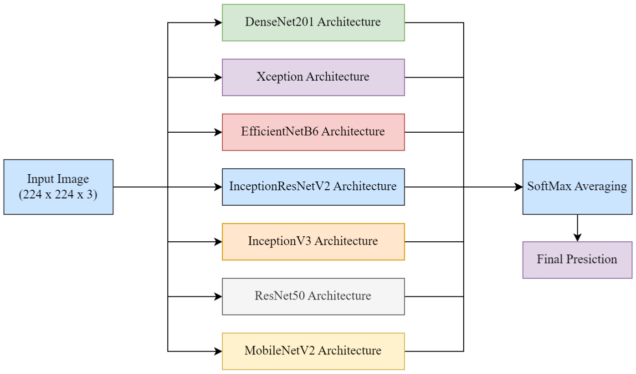

Figure 1.3 Proposed ensembling method based on SoftMax averaging

InceptionV3, ResNet50, and MobileNetV2. Each of these models has its own strengths and weaknesses, and combining them allows for a more robust prediction. The SoftMax Averaging method works by taking the SoftMax output of each model for each image and averaging them across all the models to get the final prediction. This approach is particularly effective when the models have different architectures, as is the case in this study.

The SoftMax Averaging ensemble combines the predictions from all models by taking the average of the predicted probabilities. The ensembled prediction probability $\hat{p}_i$ for image x_i can be calculated as:

$$\hat{p}_i = \frac{1}{N} \sum_{j=1}^{N} \hat{p}_{i,j} \tag{1.10}$$

where N is the total number of models used in the ensemble. Finally, the ensembled prediction $\hat{y}_i$ for image x_i can be obtained by selecting the class with the highest probability among all classes. Figure 1.3 showcases the proposed ensembling method based on SoftMax averaging.

1.3.4.1 Justification of Proposed Ensembled Method

The ensembled method based on SoftMax averaging offers several justifications for its effectiveness. Firstly, it improves generalization by combining predictions from multiple models, resulting in improved accuracy and robustness on unseen data. Secondly, it helps reduce overfitting by averaging predictions of different models, thereby decreasing the impact of overfitting. Thirdly, it leverages the diversity in model predictions, which captures a wider range of patterns in the data, leading to a more accurate and robust model. Fourthly, this approach is particularly effective for complex problems, where a single model may struggle to capture all the nuances in the data. Finally, the SoftMax function is used to convert model outputs into probabilities, making ensembled predictions easily interpretable and useful for informed decision-making.

Table 1.2 Class-wise accuracy, precision, recall, f1-score and support for the considered 27 classes of the PlantDoc dataset

Class	Accuracy	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	1.00	18
1	1.00	1.00	1.00	1.00	19
2	1.00	1.00	1.00	1.00	18
3	1.00	1.00	1.00	1.00	13
4	1.00	1.00	1.00	1.00	15
5	1.00	1.00	1.00	1.00	24
6	1.00	1.00	1.00	1.00	12
7	1.00	0.88	1.00	0.93	14
8	0.95	1.00	0.95	0.97	39
9	1.00	1.00	1.00	1.00	23
10	1.00	1.00	1.00	1.00	23
11	1.00	0.92	1.00	0.96	24
12	1.00	0.95	1.00	0.98	21
13	1.00	1.00	1.00	1.00	24
14	1.00	1.00	1.00	1.00	13
15	1.00	1.00	1.00	1.00	26
16	1.00	1.00	1.00	1.00	20
17	1.00	0.94	1.00	0.97	17
18	0.97	1.00	0.97	0.98	30
19	1.00	1.00	1.00	1.00	13
20	1.00	1.00	1.00	1.00	22
21	0.87	1.00	0.87	0.93	23
22	1.00	1.00	1.00	1.00	11
23	0.93	1.00	0.93	0.97	15
24	1.00	0.95	1.00	0.97	19
25	1.00	1.00	1.00	1.00	14
26	1.00	1.00	1.00	1.00	13

Table 1.3 Comparison among the performance obtained by previous researches and the proposed ensembled model for the PlantDoc dataset

Methods / Approaches	Obtained Accuracy
Xception [14]	81.53%
DenseNet121 [20]	86.36%
Snapshot Ensemble Technique of 4 CNNs [16]	69.51%
Proposed Ensemble (SoftMax Averaging) of 7 CNNs	98.66%

1.4 EXPERIMENTAL ANALYSIS

Initially, the PlantDoc dataset was divided into two sets: the training set (80%) and the test set (20%). After the augmentation process was completed, the augmented data was further divided into two sets: the validation set and the train set.

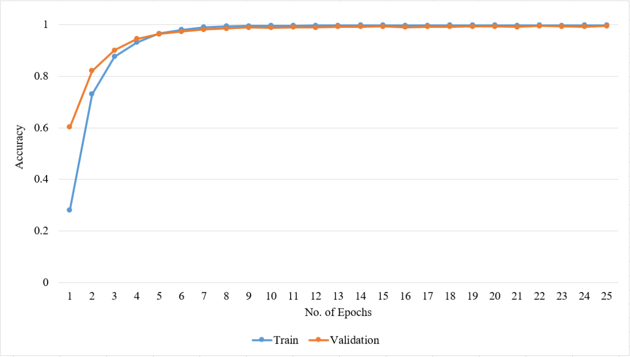

Figure 1.4 Training vs. Validation accuracy curve for PlantDoc dataset (DenseNet201)

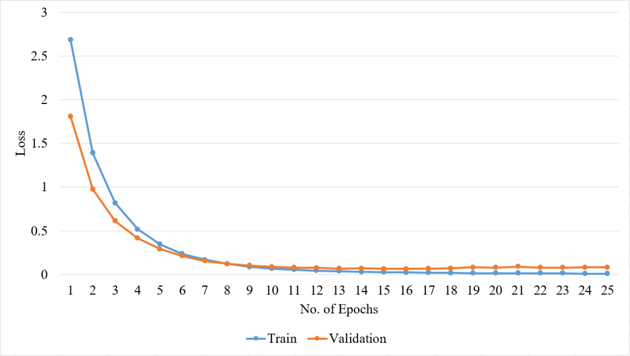

Figure 1.5 Training vs. Validation loss curve for PlantDoc dataset (DenseNet201)

1.4.1 Experimental Settings

The experimental setup used for the study involved training the model for 25 epochs with a batch size of 24. The learning rate was set to 0.000001, and the optimizer used was Adam. Accuracy matrices were used to evaluate the performance of the model, and the loss function used during training was Categorical Crossentropy. The experimental setup was designed to ensure that the model was trained effectively and that its performance was accurately assessed using appropriate evaluation metrics.

1.4.2 Obtained Results

After applying the ensembled approach combining 7 CNNs by SoftMax averaging technique, 98.66% accuracy was obtained. Among 523 samples of the test set, 516 samples were correctly classified and 7 samples were misclassified overall. Table 1.2 shows class-wise accuracy, precision, recall, F1-score, and support for the 27 classes of the PlantDoc dataset. Each row represents a particular class, and the columns show various evaluation metrics. For each class, the accuracy is 1.00, indicating that all samples belonging to that class were correctly classified. The precision and recall values are also 1.00 for most classes, indicating that the model correctly predicted all samples belonging to that class and did not misclassify any samples of other

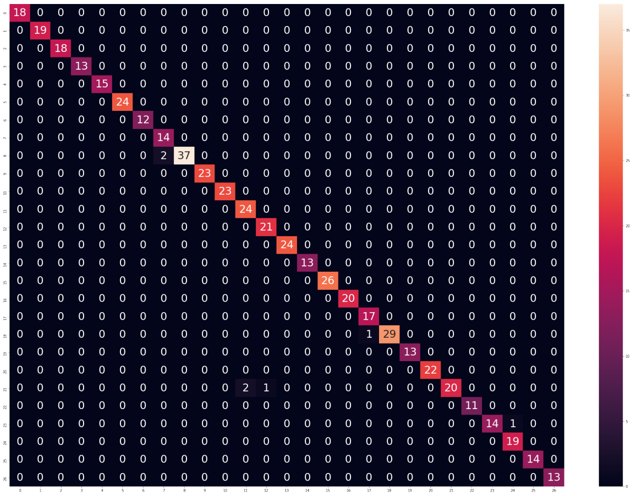

Figure 1.6 Confusion matrix for the proposed ensembled method based on SoftMax averaging

classes as belonging to that class. The F1-score is also 1.00 for most classes, which is the harmonic mean of precision and recall. However, a few classes have slightly lower values for precision, recall, and F1-score. For example, class 7 has a precision value of 0.88, indicating that a small fraction of samples belonging to this class were misclassified as belonging to other classes. Similarly, class 21 has a recall value of 0.87, indicating that a small fraction of samples belonging to this class were misclassified as belonging to other classes. Figure 1.4 and Figure 1.5 showcases the train vs. validation accuracy and loss curves for DenseNet201 which indicates that there is no overfitting. The obtained curves for other models were similar also. Figure 1.6 illustrates the confusion matrix for the proposed ensembled method.

1.4.3 Comparison and Discussion

In this study, the proposed work addresses the challenge of plant disease classification by utilizing an ensemble CNN technique, which combines seven pre-trained CNN models trained on the PlantDoc dataset. The SoftMax Averaging ensemble technique aims to achieve superior plant disease classification performance, resulting in an obtained accuracy of 98.66%. In comparison to previous researches, the proposed work significantly outperforms them, demonstrating its ability to enhance classification performance. Table 1.3 showcases the comparison among previous works and the proposed ensembled approach. Xception achieves an accuracy of 81.53%, which is considerably lower than the proposed ensemble approach. DenseNet121, on the other hand, achieves a higher accuracy of 86.36%, but still falls short in comparison to the proposed approach. The Snapshot Ensemble Technique of 4 CNNs achieves an accuracy of 69.51%, which is considerably lower than all the other approaches. The superior performance of the proposed ensemble approach can be attributed to the combination of multiple pre-trained CNN models, which allows for a more diverse range of feature extraction and improved classification accuracy. Furthermore, the SoftMax Averaging ensemble technique enables the models to complement each other's strengths, leading to a more robust and accurate model.

1.5 CONCLUSION

In conclusion, plant diseases can have devastating consequences for agricultural production, resulting in significant yield losses and reduced crop quality. Traditional methods for detecting and controlling these diseases are often time-consuming and inefficient. However, recent advancements in deep learning and computer vision technologies have paved the way for the development of more effective and efficient disease diagnosis methods. This study addresses the challenge of plant disease classification by employing an ensemble CNN technique, utilizing seven pre-trained CNN models trained on the PlantDoc dataset. The proposed approach significantly outperforms previous research, achieving an accuracy of 98.66%. The SoftMax Averaging ensemble technique and the combination of multiple pre-trained CNN models are the key factors that contribute to the proposed approach's superior performance. In the future, the aim will be to focus on mobile computing so that recognition can be performed in low-cost environment.

Bibliography

[1] Hughes, D., Salathé, M. & Others An open access repository of images on plant health to enable the development of mobile disease diagnostics. *ArXiv Preprint ArXiv:1511.08060.* (2015)

[2] Singh, D., Jain, N., Jain, P., Kayal, P., Kumawat, S. & Batra, N. PlantDoc: A dataset for visual plant disease detection. *Proceedings Of The 7th ACM IKDD CoDS And 25th COMAD.* pp. 249-253 (2020)

[3] Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Densely connected convolutional networks. *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition.* pp. 4700-4708 (2017)

[4] Chollet, F. Xception: Deep learning with depthwise separable convolutions. *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition.* pp. 1251-1258 (2017)

[5] Tan, M. & Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. *International Conference On Machine Learning.* pp. 6105-6114 (2019)

[6] Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. *Proceedings Of The AAAI Conference On Artificial Intelligence.* **31** (2017)

[7] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. *Proceedings Of The IEEE Conference On Computer Vision And Pattern Recognition.* pp. 2818-2826 (2016)

[8] Wu, Z., Shen, C. & Van Den Hengel, A. Wider or deeper: Revisiting the resnet model for visual recognition. *Pattern Recognition.* **90** pp. 119-133 (2019)

[9] Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. & Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *ArXiv Preprint ArXiv:1704.04861.* (2017)

[10] Vallabhajosyula, S., Sistla, V. & Kolli, V. Transfer learning-based deep ensemble neural network for plant leaf disease detection. *Journal Of Plant Diseases And Protection.* **129**, 545-558 (2022)

[11] Li, M., Cheng, S., Cui, J., Li, C., Li, Z., Zhou, C. & Lv, C. High-Performance Plant Pest and Disease Detection Based on Model Ensemble with Inception Module and Cluster Algorithm. *Plants.* **12**, 200 (2023)

[12] Yu, S., Xie, L. & Huang, Q. Inception convolutional vision transformers for plant disease identification. *Internet Of Things.* **21** pp. 100650 (2023)

[13] Turkoglu, M., Yanikoğlu, B. & Hanbay, D. PlantDiseaseNet: Convolutional neural network ensemble for plant disease and pest detection. *Signal, Image And Video Processing.* **16**, 301-309 (2022)

[14] Ahmad, A., El Gamal, A. & Saraswat, D. Toward Generalization of Deep Learning-Based Plant Disease Identification Under Controlled and Field Conditions. *IEEE Access.* **11** pp. 9042-9057 (2023)

[15] Vats, S. & Chivukula, V. Plant Disease Detection Using DeepNets and Ensemble Technique. *2022 IEEE International Conference On Machine Learning And Applied Network Technologies (ICMLANT).* pp. 1-6 (2022)

[16] Puangsuwan, T. & Surinta, O. Enhancement of plant leaf disease classification based on snapshot ensemble convolutional neural network. *ICIC Express Letters.* **15**, 669-680 (2021)

[17] Begum, A., Savitha, S., Shahila, S. & Sharmila, S. Diagnosis of leaf disease using enhanced convolutional neural network. *Int. J. Innov. Res. Appl. Sci. Eng.* **3**, 579-586 (2020)

[18] Chompookham, T. & Surinta, O. Ensemble methods with deep convolutional neural networks for plant leaf recognition. *ICIC Express Letters.* **15**, 553-565 (2021)

[19] Mohanty, S., Hughes, D. & Salathé, M. Using deep learning for image-based plant disease detection. *Frontiers In Plant Science.* **7** pp. 1419 (2016)

[20] Zuo, X., Chu, J., Shen, J. & Sun, J. Multi-Granularity Feature Aggregation with Self-Attention and Spatial Reasoning for Fine-Grained Crop Disease Classification. *Agriculture.* **12**, 1499 (2022)