International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 12 | Dec 2024 www.irjet.net p-ISSN: 2395-0072

Effective Plant Disease Identification Using Few-Shot Learning

Sujay Maladi[1], Hrishikesh Akundi[2], Pallavi Vanamada[3], Rahul Bellamkonda[4]

[1234]*Department of Computer Science and Engineering, GITAM (Deemed to be University), Visakhapatnam, Andhra Pradesh, India*

Abstract – *Plant diseases present a major problem to the agriculture sector, significantly impacting crop yields. Timely detection of plant diseases with good precision and accuracy is crucial for controlling outbreaks and reducing their effects on agricultural production. Although there have been recent developments in the use of Artificial Intelligence and Machine Learning for the detection of plant diseases, these solutions are dependent on large datasets of annotated images which may not be readily available in real life. This study examines the application of Few-Shot Learning approach for plant disease identification in scenarios with limited datasets. Few-Shot Learning enables a model to recognize new classes with very small number of examples, making it highly efficient. It imitates human learning abilities to generalize from minimal examples. This contrasts traditional machine learning methods, which usually require huge amounts of labeled data to achieve good accuracy. In this paper, the PlantVillage dataset has been used and EfficientNet, a pre-trained Convolutional Neural Network which was trained on the ImageNet dataset has been used along with Siamese Networks and a Triplet Loss function for FSL. The proposed approach is tested and compared with the regular fine-tuning transfer learning. The outcomes reveal that Few-Shot Learning with EfficientNet and Siamese Networks outshines the conventional transfer learning methods for the detection of plant diseases especially in the scenarios where quick adaptability is needed.*

Key Words: Machine Learning, Deep Learning, Few-shot Learning, Transfer Learning, Fine-tuning, Siamese Networks, Triplet Loss, EfficientNet

1. INTRODUCTION

The agriculture sector is the backbone in providing food for human beings and animals, as well as in helping achieve sustainable development goals, especially SDG 2: Zero Hunger. The agriculture sector mainly deals with plants, and currently, plant diseases are a major problem. Plant diseases that are not detected and managed early on may spread throughout the entire crop. This not only reduces produce quality but also causes huge economic losses for farmers and may harm the ecosystem. Manual plant disease detection is susceptible to human error and demands extensive knowledge in the field.

So, there is a clear need for automated plant disease detection system [1]. Images of plant leaves can be used to easily detect the presence of a disease and identify its type [2]. Artificial Intelligence and Machine Learning have introduced many solutions for detecting and classifying plant diseases but most of these approaches are impractical due to the need of collecting and annotating large datasets of plant images. Hence, this study uses an approach called Few-Shot Learning.

1.1 Few-Shot Learning

Few-Shot Learning is a machine learning approach where a model learns to make predictions by training with only a very small number of examples. This approach is mainly used when data is limited and the model needs to adjust efficiently to new tasks. Few-Shot Learning is generally used in applications like face recognition, medical imaging, natural language processing, visual search, robotics, game AI, etc.

Few-Shot Learning can be studied using the N-way-K-shot structure where N indicates the number of classes and K indicates the number of examples provided for each of the N classes. For example, if it is 4-way-2-shot task, it means it contains 4 classes of images and 2 examples of each class are provided. The higher the N value, the more difficult the task is but the higher the K value, the easier the task gets because more supporting information is available to derive an inference.

In Few-shot learning, the K value is typically less than or equal to 10. Few-shot learning is given specific names when K = 0 and K = 1. If K = 0, it is called zero-shot learning, and if K = 1, it is called one-shot learning.

2. RELATED WORK

Deep learning approaches based on convolutional neural networks (CNNs) have proven highly effective for plant disease classification as highlighted in [3, 4, 5, 6]. However, with the rapid emergence of new plant diseases, there is a need for plant disease detection technologies that can quickly adjust to changes and operate effectively with limited data [7, 8]. In this context, few-shot learning appears to be a good solution.

In [9] the authors proposed a few-shot learning approach using Siamese Networks. Here, the features are extracted by a two way convolutional neural network (CNN) with shared weights, a spatial structure optimizer (SSO) enhanced the metric learning and a k-nearest neighbor (kNN) classifier is used for classification. In [10] a low-shot learning method was proposed, where support vector machine (SVM) was employed to segment disease spots while preserving edge

information, and improved conditional deep convolutional generative adversarial networks (C-DCGAN) generated augmented images for training VGG16 model. They achieved an average accuracy of 90%. In [11] the authors proposed a methodology for plant disease detection based on a semi-supervised few-shot learning approach. They improved the accuracy of the model by adaptively choosing the pseudo-labeled samples to fine-tune the model. As discussed in [12], the author proposed a few-shot learning strategy which involved a pre-trained model on the ImageNet dataset and refining it with the help of a similar dataset for plant-based features extraction. Here, three CNN baseline architectures were used: GoogleNet, DenseNet and MnasNet. M. H. Saad and A. E. Salman [13] proposed a one-shot learning approach utilizing Siamese network. They also provided a unique method for image segmentation. Also, M. Rezaei et al. [14] introduced a few-shot learning method integrated with a feature attention module. The feature learners used here are ResNet50 and Vision Transformers (ViT). All of these studies yielded very good results, but none of them used EfficientNet.

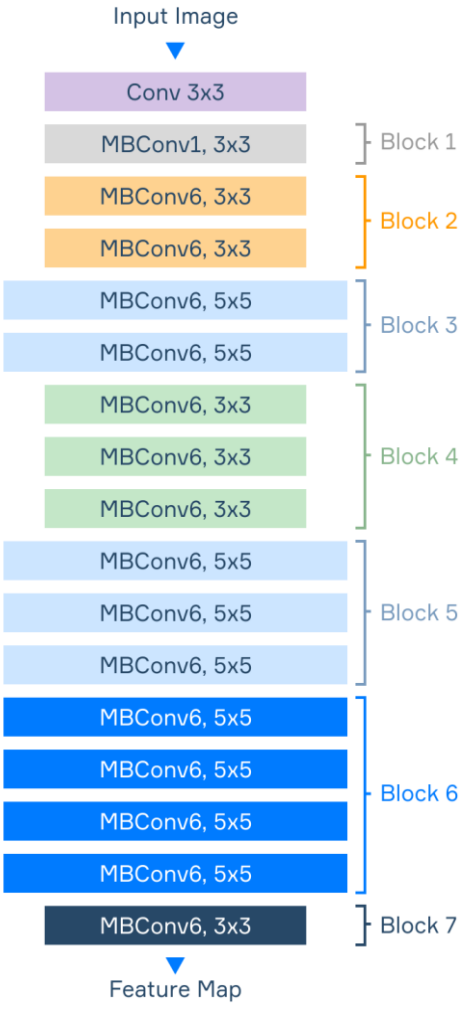

Fig-1: EfficientNet-B0 Architecture

EfficientNet is a CNN based architecture which is mainly designed for problems related to computer vision and image recognition. EfficientNet-B0 is the baseline architecture from which EfficientNet-B1 to EfficientNet-B7 are derived. This is done using compound scaling. Fig-1 shows the EfficientNet-B0 architecture. According to [15, 16] EfficientNet showcased better performance compared to other similar architectures such as ResNet, AlexNet, VGG16 and Inception V3. Hence, this study employs EfficientNet.

3. MATERIALS

This study used the PlantVillage dataset, which was launched as a part of the PlantVillage project, an initiative led by researchers from Pennsylvania State University. This dataset was selected for this study because it is widely recognized as a benchmark for plant disease classification problems. This dataset contains 54303 images of plant leaves with 38 classes as shown in Table-1. It includes 14 different crop species and 26 different diseases [2]. Some examples are shown in Fig-2.

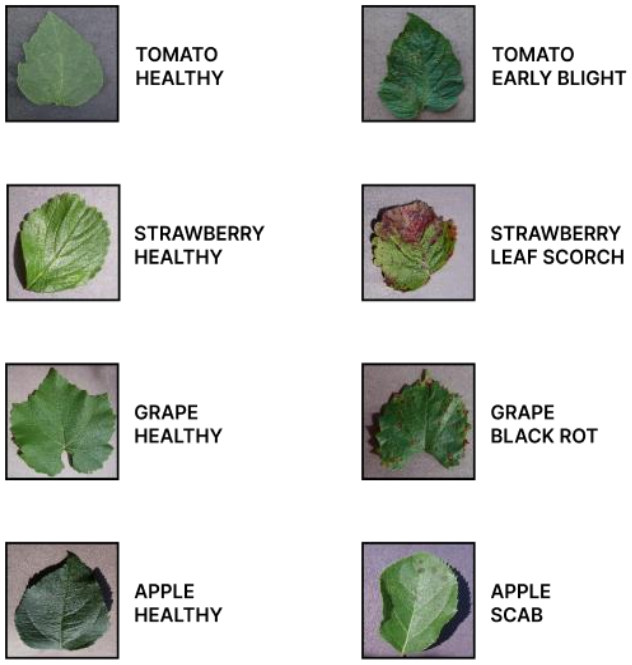

Fig-2: Some images from PlantVillage dataset

Data augmentation techniques including rotation, flipping, cropping and Gaussian blurring were applied to increase the diversity of the dataset. Also, all the images were resized to equal dimensions of 256 × 256. The dataset was divided into a source domain D_s of 34 classes and a target domain D_t of 4 classes. The images are split into 80% for training and 20% for testing.

This study utilized both local and cloud-based environments for implementation. The local setup included a computer with the following specifications: Intel(R) Core(TM) i7-9750H CPU, NVIDIA(R) GeForce(R) GTX 1660 Ti GPU (6GB GDDR6

memory), 32 GB of DDR4 RAM, and a 64-bit Windows 10 Pro operating system and the study was conducted using Python in Jupyter Notebook. Google Colab, a cloud-based platform provided access to advanced computational resources such as GPUs.

Table-1: PlantVillage Dataset Overview

Crop	Condition or Disease	Images
Cherry	Powdery Mildew	1052
Corn	Cercospora	513
Corn	Rust	1192
Corn	Healthy	1162
Corn	Northern Leaf Blight	985
Grape	Black Rot	1180
Grape	Black Measles	1384
Grape	Healthy	423
Grape	Isariopsis Leaf Spot	1076
Orange	Citrus Greening	5507
Peach	Bacterial Spot	2291
Peach	Healthy	360
Pepper	Bacterial Spot	997
Pepper	Healthy	1478
Potato	Early Blight	1000
Potato	Healthy	152
Potato	Late Blight	1000
Raspberry	Healthy	371
Soybean	Healthy	5090
Squash	Powdery Mildew	1835
Strawberry	Healthy	456
Strawberry	Leaf Scorch	1109
Tomato	Bacterial Spot	2127
Tomato	Early Blight	1000
Tomato	Healthy	1592
Tomato	Late Blight	1910
Tomato	Leaf Mold	952
Tomato	Septoria Leaf Spot	1771
Tomato	Spider Mites	1676
Tomato	Target Spot	1404
Tomato	Mosaic Virus	373
Tomato	Yellow Leaf Curl	5357
Apple	Scab	630
Apple	Black Rot	621
Apple	Cedar Rust	276
Apple	Healthy	1645
Blueberry	Healthy	1502
Cherry	Healthy	854

4. METHODOLOGY

The system architecture of this study is based on proposals given in [17, 18]. Concepts like pre-training, fine-tuning, transfer learning, few-shot learning, etc. were used in this study. By the end of the study, using performance metrics, the performance of a regular fine-tuning transfer learning model is compared with that of a few-shot learning model using Siamese networks. Fig-3 shows the complete system methodology.

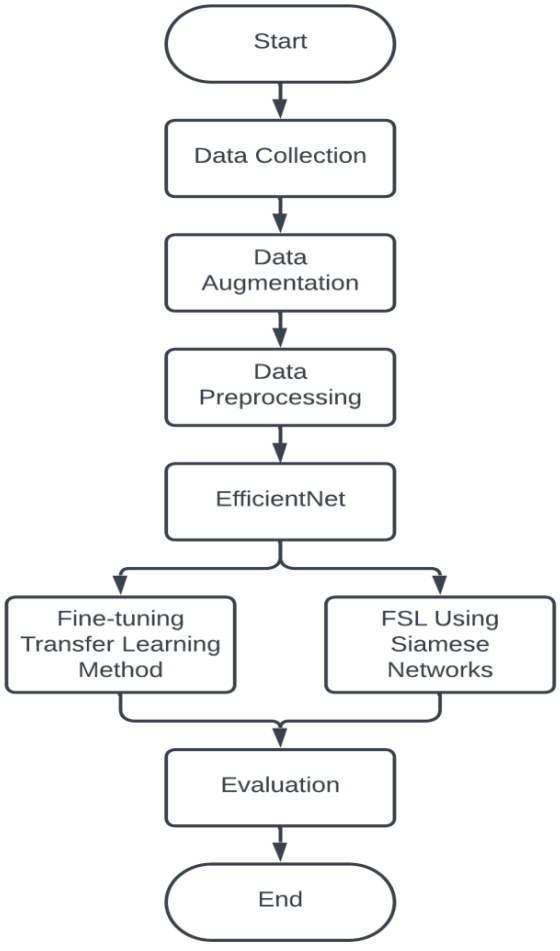

Fig-3: System Methodology

In studies like this, transfer learning is an important part [19]. Transfer learning is a technique where a model trained on one task (source) is reused or fine-tuned for another related task (target). The knowledge gained from a large, general dataset can be used to improve performance on a smaller or specific dataset. Knowledge can be transferred by freezing some layers in the network and fine-tuning other layers.

Fig-4 illustrates the procedure where a convolutional neural network based architecture is adapted for feature extraction of plant leaves. Here, the image x_i is embedded into a feature vector f_i. Then, an SVM classifier is employed to determine the label of input image. Fig-5 shows the Few-Shot Learning architecture. The upper section is a plant leaf classifier trained on source dataset and the lower section is achieved by freezing the CNN.

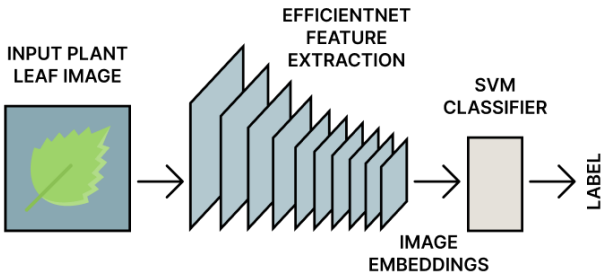

Fig-4: Plant Leaf Classification

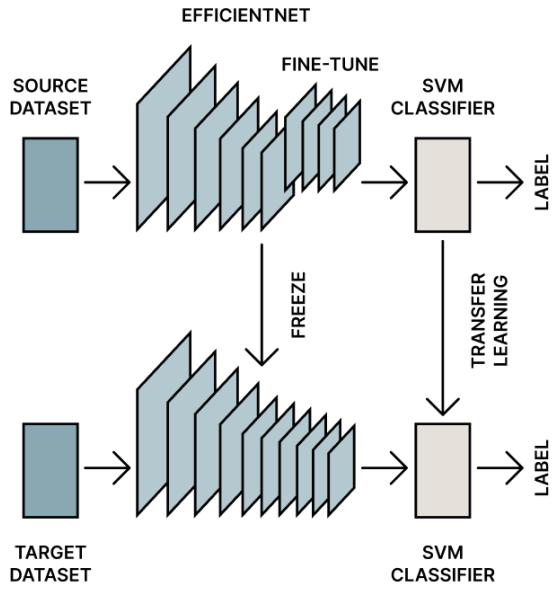

Fig-5: Few-Shot Learning Architecture

4.1 Pre-training

Pre-training is the process of training a model on a huge dataset before fine-tuning it on smaller datasets. These huge datasets are generic and the model learns general features from such datasets. This step is crucial as training models from scratch on such huge datasets can be computationally expensive. After pre-training the model can be further trained or fine-tuned on a smaller, task-specific dataset.

ImageNet is a large-scale dataset with images broadly used in computer vision and image recognition tasks for training and benchmarking machine learning models. ImageNet has over 14 million labeled images organized into 21 thousand categories. It has wide diversity as the images are taken from various real-world scenarios, covering a range of objects and environments. ImageNet serves as the main foundation for many popular pre-trained models like ResNet, VGG, AlexNet, Inception, MobileNet, DenseNet, EfficientNet, etc.

These models can be used for feature extraction purposes and they can act as a foundation for transfer learning tasks. This study mainly focused on EfficientNet because of its potential for superior performance compared to other similar models.

4.2 EfficientNet

EfficentNet was introduced in 2019, mainly designed to achieve high accuracy with fewer parameters and computational resources compared to other convolutional neural networks. EfficientNet uses a systematic compound scaling method to balance the depth, width, and resolution of the network effectively. EfficentNet has a range of variants from B0 to B7. Smaller variants like B0 are suitable for resource-constrained environments. Larger variants like B7 are optimized for maximum accuracy.

In this study, EfficientNet-B5 pre-trained on ImageNet is utilized as the baseline architecture because of its balanced performance and efficiency. ImageNet is a general dataset that does not exclusively focus on plants. So, the convolutional neural network is first used for feature extraction of plant leaves. It then learned to map images of plant leaves x_i into $f(x_i)$, which represents essential characteristics.

4.3 Fine-tuning

For task specific purposes, the network is then fine-tuned. Fine-tuning involves the process of adjusting a pre-trained convolutional neural network like EfficientNet to a specific task – in this case, plant leaves. Around 80% of the layers of EfficientNet network are frozen. So, these weights are not affected during training. General features like edges and shapes are preserved. The remaining layers are trained with the source dataset D_s. Then, the fully connected layers are replaced with support vector machine (SVM) classifier, which is trained to make predictions. Then, a dimensionality reduction is applied to simplify the representation while preserving essential features. An embedding dimension of 128 is selected for this task.

4.4 Siamese Network

For learning image embeddings, we can use Siamese network. In [17] the authors mentioned two architectures:

1. Siamese network with two subnets and contrastive loss.

2. Siamese network with three subnets and triplet loss.

The authors stated that the Siamese network with three subnets and triplet loss produced the best results compared to the baseline fine-tuned model and the Siamese network with contrastive loss. So, this study focused on the Siamese network with three subnets and triplet loss.

A three subnet Siamese network with triplet loss is an architecture often used for tasks related to similarity learning such as face recognition, object identification, etc. It can also be employed in plant disease classification. This Siamese network has three inputs: Anchor (X_a), Positive (X_p), Negative (X_n). An anchor is like a reference image, positive is an image similar to the anchor, and negative is an image different from the anchor. The triplet loss function is given by:

$$L = max\ (0, ||f_a - f_p||^2 - ||f_a - f_n||^2 + m) \quad (1)$$

Here, all the subnets in the Siamese network architecture share the same weights. This is illustrated in Fig-6.

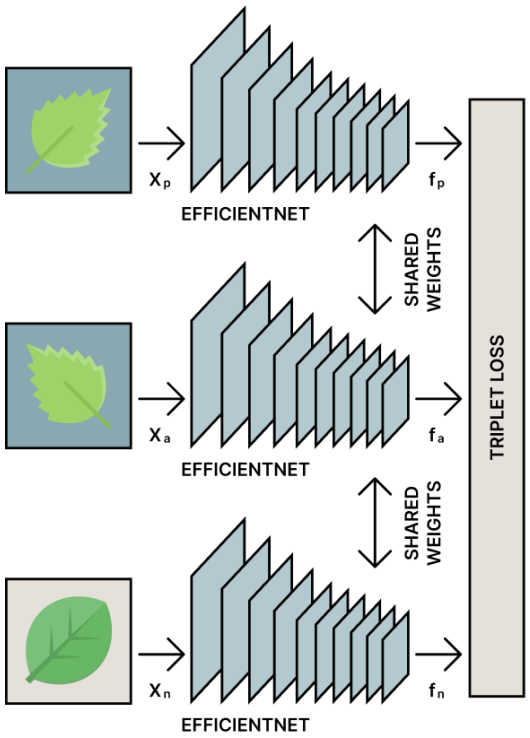

Fig-6: Siamese network with three subnets and triplet loss.

4.5 Support Vector Machine (SVM)

SVMs are algorithms used for classification and regression problems. They work well for tasks which require a clear margin of separation between classes. The aim of SVM classifier is to reduce misclassification and improve generalization on unseen data. SVMs can be applied to handle multiclass problems using techniques like one-vs-all or one-vs-one strategies. This study used a one-vs-all strategy. In this approach for a K class problem, K separate classifiers are trained. Each of them distinguishes one class from all other classes. Here, each embedding *(f_i)* is assigned to a class based on the SVM classifier with the highest confidence score.

4.6 Few-Shot Learning

Now, knowledge transfer from source domain D_s to target domain D_t is done. The goal is to adapt a model trained on one domain (D_s) to perform well on another domain (D_t). After training the convolutional neural network on the source domain D_s the SVM classifier is retrained using embeddings from the target domain D_t. The embeddings for both the methods were derived from the CNN trained on D_s. During this phase only the shallow classifier is retrained using the training data from D_t. This study used different amounts of training data per class, ranging from 1 image to 50 images. Comparing it with the N-way-K-shot structure, K value varies from 1 to 50. So, this number can be called as the number of shots.

For the baseline fine-tuning method, EfficientNet is first trained on the source D_s and few layers of the network are then fine-tuned using subsets of target domain D_t. The test set has been kept constant in both the methods and the experiment is repeated multiple times to ensure reliability.

For the experiment, 50 epochs were used with a batch size of 34. Adam optimizer with margin m = 1 was used in triplet loss function. The experiment was conducted with the help of Keras powered by TensorFlow.

4.7 Performance Metrics

Performance metrics for this experiment can be calculated using confusion matrix. These performance metrics help evaluate the quality of the model's predictions. Below are the commonly used metrics:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (2)$$

$$Precision = TP / (TP + FP) \quad (3)$$

$$Recall = TP / (TP + FN) \quad (4)$$

$$F1\text{-}Score = 2 * (Precision * Recall) / (Precision + Recall) \quad (5)$$

Where, TP refers to True Positives, TN refers to True Negatives, FP refers to False Positives, and FN refers to False Negatives.

Since, this is multiclass problem, the multiclass accuracy can be calculated using equation (6). In this study, the datasets classes were balanced making multiway accuracy an effective measure of performance [17, 20].

$$\text{Accuracy} = (\sum_i N_{ii}) / (\sum_{i,j} N_{ij}) \quad (6)$$

Where N_{ij} refers to the number of images of i classified by the algorithm as j. In this experiment, $i,j \in \{1, 2, 3, \ldots, 34\}$ for the main branch in the D_s dataset, and $i,j \in \{1, 2, 3, 4\}$ for the Few-Shot Learning in the D_t dataset.

5. RESULTS

Both the methods, i.e., regular fine-tuning transfer learning and the proposed methodology, Few-Shot Learning with EfficientNet and Siamese Networks were evaluated. The experiment was conducted using different amounts of examples per class, ranging from 1 image to 50 images. Results show that the proposed methodology, Few-Shot Learning with EfficientNet and Siamese Networks, achieved an accuracy of 90.2% using as few as 10 shots, whereas the regular fine-tuning transfer learning model consistently scored significantly lower accuracy than proposed model for every value of number of examples per class. Both methods struggled to differentiate between leaves and diseases that closely resemble each other. These methods are likely to achieve even better results on datasets where plant diseases are more visually distinct. The number of shots vs maximum accuracy is shown in Fig-7.

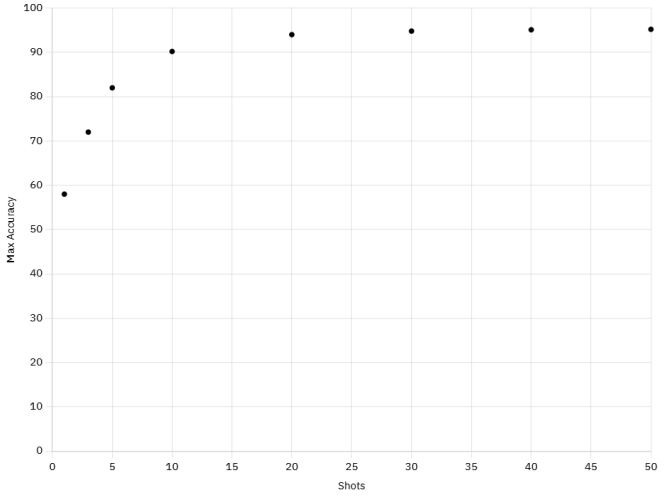

Fig-7: Accuracy Plot

As shown in Fig-7, the proposed methodology gives higher accuracy as the number of shots increases. As the number of shots increase, the challenge of learning with small number of examples diminishes. However, using larger number of shots defeats the primary purpose of employing Few-Shot Learning and it starts to resemble traditional supervised learning where larger datasets are available to use.

In recent years, Few-Shot Learning has been gaining significant attention in various domains. Many recent studies have applied Few-Shot Learning techniques to diverse tasks demonstrating its effectiveness [21, 22, 23, 24].

EfficientNet is one of the latest convolutional neural network architectures and can be considered as state-of-the-art. While other similar convolutional neural network architectures can also be used in this experiment, EfficientNet gives better accuracy with fewer parameters compared to other architectures [25]. This highlights its potential to use in similar studies and applications.

5. CONCLUSION

This study shows that Few-Shot Learning with EfficientNet and Siamese Networks using triplet loss gave better results compared to the regular fine-tuning transfer learning model. The study also highlights the potential of Few-Shot Learning for developing effective plant disease identification solutions with minimal annotated data. Also, the findings confirm that Few-Shot Learning is well-suited for tasks requiring quick adaptation and where data is limited.

REFERENCES

[1] 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, pp. 1–11, 2016, doi: 10.1155/2016/3289801.

[2] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," Frontiers in Plant Science, vol. 7, Sep. 2016, doi: 10.3389/fpls.2016.01419.

[3] G. G. and A. P. J., "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," Computers & Electrical Engineering, vol. 76, pp. 323–338, Jun. 2019, doi: 10.1016/j.compeleceng.2019.04.011.

[4] M. H. Saleem, J. Potgieter, and K. M. Arif, "Plant Disease Detection and Classification by Deep Learning," Plants, vol. 8, no. 11, p. 468, Oct. 2019, doi: 10.3390/plants8110468.

[5] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review," Plant Methods, vol. 17, no. 1, p. 22, Dec. 2021, doi: 10.1186/s13007-021-00722-9.

[6] J. Lu, L. Tan, and H. Jiang, "Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification," Agriculture, vol. 11, no. 8, p. 707, Jul. 2021, doi: 10.3390/agriculture11080707.

[7] J. B. Ristaino et al., "The persistent threat of emerging plant disease pandemics to global food security," Proceedings of the National Academy of Sciences, vol. 118, no. 23, Jun. 2021, doi: 10.1073/pnas.2022239118.

[8] R. W. Mwangi, M. Mustafa, K. Charles, I. W. Wagara, and N. Kappel, "Selected emerging and reemerging plant pathogens affecting the food basket: A threat to food security," Journal of Agriculture and Food Research, vol. 14, p. 100827, Dec. 2023, doi: 10.1016/j.jafr.2023.100827.

[9] B. Wang and D. Wang, "Plant Leaves Classification: A Few-Shot Learning Method Based on Siamese Network," IEEE Access, vol. 7, pp. 151754–151763, 2019, doi: 10.1109/ACCESS.2019.2947510.

[10] G. Hu, H. Wu, Y. Zhang, and M. Wan, "A low shot learning method for tea leaf's disease identification," Computers and Electronics in Agriculture, vol. 163, p. 104852, Aug. 2019, doi: 10.1016/j.compag.2019.104852.

[11] Y. Li and X. Chao, "Semi-supervised few-shot learning approach for plant diseases recognition," Plant Methods, vol. 17, no. 1, p. 68, Dec. 2021, doi: 10.1186/s13007-021-00770-1.

[12] P. Uskaner Hepsağ, "Efficient plant disease identification using few-shot learning: a transfer learning approach," Multimedia Tools and Applications, vol. 83, no. 20, pp. 58293–58308, Dec. 2023, doi: 10.1007/s11042-023-17824-2.

[13] M. H. Saad and A. E. Salman, "A plant disease classification using one-shot learning technique with field images," Multimedia Tools and Applications, vol. 83, no. 20, pp. 58935–58960, Dec. 2023, doi: 10.1007/s11042-023-17830-4.

[14] M. Rezaei, D. Diepeveen, H. Laga, M. G. K. Jones, and F. Sohel, "Plant disease recognition in a low data scenario using few-shot learning," Computers and Electronics in Agriculture, vol. 219, p. 108812, Apr. 2024, doi: 10.1016/j.compag.2024.108812.

[15] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," Ecological Informatics, vol. 61, p. 101182, Mar. 2021, doi: 10.1016/j.ecoinf.2020.101182.

[16] A. K. Bairwa, S. Joshi, and S. Chaudhary, "Machine Learning Based Plant Disease Detection Using EfficientNet B7," 2024, pp. 1–12. doi: 10.1007/978-3-031-47372-2_1.

[17] D. Argüeso et al., "Few-Shot Learning approach for plant disease classification using images taken in the field," Computers and Electronics in Agriculture, vol. 175, p. 105542, Aug. 2020, doi: 10.1016/j.compag.2020.105542.

[18] A. Medela et al., "Few Shot Learning in Histopathological Images:Reducing the Need of Labeled Data on Biological Datasets," in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, Apr. 2019, pp. 1860–1864. doi: 10.1109/ISBI.2019.8759182.

[19] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "Using deep transfer learning for image-based plant disease identification," Computers and Electronics in Agriculture, vol. 173, p. 105393, Jun. 2020, doi: 10.1016/j.compag.2020.105393.

[20] A. B. Rad et al., "ECG-Based Classification of Resuscitation Cardiac Rhythms for Retrospective Data Analysis," IEEE Transactions on Biomedical Engineering, vol. 64, no. 10, pp. 2411–2418, Oct. 2017, doi: 10.1109/TBME.2017.2688380.

[21] J. Wang, K. Liu, Y. Zhang, B. Leng, and J. Lu, "Recent advances of few-shot learning methods and applications," Science China Technological Sciences, vol. 66, no. 4, pp. 920–944, Apr. 2023, doi: 10.1007/s11431-022-2133-1.

[22] Y. Song, T. Wang, P. Cai, S. K. Mondal, and J. P. Sahoo, "A Comprehensive Survey of Few-shot Learning: Evolution, Applications, Challenges, and Opportunities," ACM Computing Surveys, vol. 55, no. 13s, pp. 1–40, Dec. 2023, doi: 10.1145/3582688.

[23] W. Li et al., "LibFewShot: A Comprehensive Library for Few-Shot Learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 12, pp. 14938–14955, Dec. 2023, doi: 10.1109/TPAMI.2023.3312125.

[24] W. Zeng and Z. Xiao, "Few-shot learning based on deep learning: A survey," Mathematical Biosciences and Engineering, vol. 21, no. 1, pp. 679–711, 2023, doi: 10.3934/mbe.2024029.

[25] M. Tan and Q. v. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," May 2019.