Data-Driven Solutions in Agriculture: Machine Learning for Apple Disease Classification

1st Hanish Acharla
Mountain House High School
Mountain House, United States of America
1079407@lammersvilleusd.net

***Abstract*—The agricultural sector faces significant challenges due to plant diseases, with apple orchards being particularly susceptible to various pathogens. Accurate and timely disease detection is crucial for mitigating crop loss and ensuring food security. This paper proposes an approach to apple disease detection by leveraging both - fine-tuning and feature extraction methods using advanced pre-trained Convolutional Neural Network (CNN) architectures. Specifically, we employ ResNet-50, a widely adopted deep learning model, fine-tuning it to optimize for apple disease classification. Additionally, we utilize ConvNeXt-Xlarge for feature extraction, integrating its outputs with shallow classifiers, such as Random Forest(RF), K-Nearest Neighbors (KNN), and Multilayer Perceptron (MLP), to optimize classification accuracy. Comparative evaluations were performed on the Plant Pathology dataset, featuring six categories of apple leaf disease. Results demonstrate that ConvNeXt-Xlarge coupled with MLP achieved a test accuracy of 87.50%, outperforming ResNet-50. Achieving a high degree of precision (88%) in identifying and classifying apple diseases enables early intervention and reduces manual inspection processes. The research findings highlight the potential of deep-learning strategies in transforming plant disease management and enhancing agriculture productivity.**

I. Introduction

Throughout mankind, agricultural sectors have been at the forefront of importance because they are necessary for human survival and yet plant diseases continue to pose serious challenges to crop productivity and food availability worldwide. With over 10% of global food production lost annually to disease, the agricultural sector faces increased strain, particularly in regions already vulnerable to food insecurity and famine [1]. Effective disease prevention is essential for mitigating these losses and supporting food security. However, misdiagnosis or delayed diagnosis of plant diseases often results in excessive pesticide use, which not only damages soil and water quality but also poses risks to human health and reduces overall crop yield [2]. Addressing plant disease accurately and promptly is thus crucial for sustaining agricultural productivity and minimizing the impacts of food scarcity.

This paper focuses on classifying apple diseases as apples are the third most produced fruit in the world with over 95.84 Million Metric Tons being produced yearly based on 2022 statistics [3]. They play a large part in the fruit industry and bring over 23 Billion Dollars worth of downstream revenue every year, showing just how important they are to the farming sector [4]. A large reason for this is the popularity of the product due to their nutritional benefits, being rich in fiber, Vitamin A, and antioxidants, while also being affordable and available in large quantities. However, their susceptibility to diseases like powdery mildew, scab, and rust makes them particularly vulnerable, with over 20% of apple production lost to disease annually. This leads to economic losses, increased pesticide use, and reduced biodiversity. For instance, the Fireblight outbreak in Michigan led to the loss of over 220,000 apple trees and $42 million in damages [5].

Diseases significantly affect apple trees primarily due to limited methods available for detecting them in orchards. The problem is compounded by the fact that apple trees have weak defense mechanisms against diseases, creating challenging situations for farmers [5]. At present, the most common method for identifying disease is visual inspection, which is not always reliable and can lead to several issues. First, the level of experience among farmers varies, leading to some possible mistaking diseases for mere blemishes or healthy trees. Furthermore, diagnosing the exact disease often requires farmers to send samples for lab testing, before they can take corrective actions. This not only reduces efficiency but also increases the cost and efforts involved in orchard management. In recent years, improvements in AI have opened the doors for using machine learning in the agricultural sector and have greatly helped farmers. AI helps with spotting and managing plant diseases by using smart image recognition to find problems early. Early diagnosis and timely interventions improve crop production and sustainability.

The paper is organized as follows - related work section surveys existing research on the use of deep learning models in disease classification, identifying key trends and gaps. The materials and methods section details the dataset, preprocessing techniques, and the deep learning architectures used. This is followed by the results section, where we present the evaluation metrics for accessing the models performance, analyze and interpret the results illustrated through figures and tables. Finally, the paper concludes with a summary of our findings, and acknowledge the limitations of our approach and suggestions for future research directions.

II. Related Work

Over recent decades, artificial intelligence (AI) has transformed numerous industries, including agriculture, with significant advancements [6]. Machine learning (ML), a subset of AI, uses algorithms that can learn from data, recognize patterns, and make predictions [7]. However, conventional ML methods often require manual feature extraction, a process that can be time-consuming and may not always result in optimal accuracy [2]. Deep learning (DL) offers a solution by automatically extracting features from raw data, which enhances predictive accuracy and efficiency [8].

One of the most successful DL architectures for image-related tasks is the Convolutional Neural Network (CNN), which captures complex data patterns through its layered structure [9]. CNNs have proven particularly effective for plant disease classification, as they can detect subtle disease symptoms in leaf images, enabling early diagnosis and intervention [10]. DL models' scalability further supports their application to large agricultural datasets, allowing them to process vast quantities of plant images and recognize disease patterns efficiently [11].

Numerous studies have demonstrated the effectiveness of CNNs in accurately identifying apple leaf diseases, which helps farmers make informed decisions and improves crop health and yield [12]. For instance, DL-based approaches have shown high accuracy in detecting apple leaf diseases such as scab, rust, and powdery mildew, significantly contributing to precision agriculture [13].

In [14], a Deep Convolutional Generative Adversarial Network(DCGAN) is used to synthesize images to combat the issue of having a relatively small dataset, consisting of 319 images. AlexNet [15] is used for classifying 4 different apple leaf diseases. The study [2] employs DenseNet-121 [16] as a backbone model. A notable drawback of this study is that the dataset is very skewed, where the ratio between Healthy Apples and Serious Cedar Apple Rust is 1:30, potentially leading to misclassification. These approaches face challenges due to a limited dataset size, hindering broader disease identification.

Chao et al. [17] collected a custom dataset and trained a combination of DenseNet and Xception [18] models on the dataset. In [19], the EfficientNetB7 [20] model with the Noisy Student Weights [21] was trained on the custom-created dataset. The study [22], employs a custom CNN model on an augmented subset of the Plant Village Dataset. A customized VGG16 [23] model is trained to classify apple leaf diseases, resulting in low accuracy. Yu et al. employed the ResNet-50 [24] model trained on an augmented apple leaf disease dataset [25]. However, these studies do not address the issue of apples having multiple diseases simultaneously. In practical situations, farmers might find these models challenging to use effectively, as many of their apple diseases could be incorrectly identified, causing more harm than good.

The study [26] used MobileNet [27], Inception V3 [28], and ResNet-152 [29] on an augmented dataset, but it faced the issue of low classification accuracy. Khan et al. [30] employed an Xception model to train on a custom dataset of about 6000 images, classifying apple leaf diseases into 10 labels. Diseased images are analyzed with EfficientNet [31] to segment the diseased portion of the leaves, but this approach has a main drawback of relatively low accuracy [32]. Even though these approaches combat the issues of insufficient datasets that consist of a small number of diseases, their classification results are relatively low. This leads to problems in real-world use cases where the model misclassifies a disease, causing more issues for farmers. Even with such large amounts of research being placed in Deep Learning Classification for Disease Detection, many drawbacks are still apparent and need ratification.

In this study, we have focused on classifying apple leaf diseases using various Deep Learning Algorithms. The dataset used in the study is the Plant Pathology 2021 dataset which consists of over 17,000 RGB images of apple leaf diseases. In the first approach, we fine tuned ResNet-50 model for disease classification task. Second, using ConvNeXt-Xlarge, we extracted features from the dataset and used shallow models like K Nearest Neighbors(KNN) [33], Random Forest(RF) [34], and Multilayer Perceptron(MLP) [35] for the classification of the disease. Metrics such as Accuracy, Precision, Recall, and F-1 score were used to evaluate the performance of the models. From the experimental results, it is evident that featurization approach effectively addressed the challenges of classifying multiple apple leaf diseases.

III. Materials and Methods

The key stages of the methodology include data preprocessing, model selection, hyperparameter tuning, and evaluation on separate training, validation, and test sets. The following subsections describe the dataset, model architectures, and implementation techniques employed to achieve reliable disease classification.

A. Dataset

The dataset used in this study is the Plant Pathology Dataset [36] publicly available on Kaggle. The dataset consists of 17,826 RGB images split into 12 categories. Originally the 12 categories were Scab, Healthy, Frog Eye Leaf Spot, Rust, Complex, Powdery Mildew, Scab & Frog Eye Leaf Spot, Scab & Frog Eye Leaf Spot & Complex, Frog Eye Leaf Spot & Complex, Rust & Frog Eye Leaf Spot, Rust Complex, and Powdery Mildew Complex. The samples that contained multiple diseases were labeled as Complex. The issue that arises with this dataset is that it is imbalanced with the ratio of samples between Scab & Powder Mildew and Complex being 55:1. The imbalanced dataset leads to misclassification of the minority classes which are underrepresented in the dataset. In order to counter the class imbalance problem, the samples belonging to the minority classes(Scab Frog Eye Leaf Spot complex, Frog Eye Leaf Spot complex, Rust complex, and Powdery Mildew complex) were combined into the Complex Class. The Plant Pathology dataset was split into Train, Validation, and Test datasets. Each split contained 14,288, 1,790, and 1,748 images respectively.

B. Methodology

Primarily two deep architectures were used in this study. The first one is ResNet-50, which is a 50-layer deep convolutional neural network (DCNN) from the Residual Network family, designed to address vanishing gradient issues via residual connections that enable effective gradient flow. Its architecture is an effective and widely used model for image classification tasks.

Using pre-trained models in Image Classification tasks offers various advantages. However, many pre-trained models are large and computationally heavy, requiring substantial computational power and memory, which hinders its deployment in an environment with limited resources. Featurization provides a way to overcome some of these challenges. By leveraging pre-trained models as a feature extractor, the model's ability to capture complex patterns in the images is utilized to extract the intricate features of the image. These extracted features are then fed into simpler models like KNN and RF to make classifications, which can address the domain mismatch problem of the pre-trained model. Since feature extraction only involves fine-tuning a few model layers, it reduces the chance of overfitting. Furthermore, instead of fine-tuning the entire model, which is computationally expensive, this approach only requires running the pre-trained model once to extract features which reduces the computational demand of the pre-trained model which makes it suitable for practical applications.

Using ConvxNets for featurization is a popular approach in Image Classification tasks [37]. One of the key features of ConvNeXt-Xlarge is its use of Depth Wise Convolutionals which reduce the computational load and increase the model's efficiency without sacrificing efficiency [38]. Each Convolutional Layer employs a variety of kernel sizes and strides to capture different levels of detail and spatial hierarchies. In ConvNeXt-Xlarge, depthwise convolutions enhance efficiency by applying separate kernels to each input channel, reducing computational cost and memory requirements without sacrificing the feature quality. The feature maps produced by each layer progressively shrink spatially while increasing in depth, resulting in a compact, high-dimensional representation. After passing through all CNN layers, a global average pooling layer reduces the feature map to a final 2048-dimensional vector (1x1x2048). This vector is a concise summary of the image's most salient features.

This output vector effectively forms a high-dimensional "tabular" dataset where each row represents an image, and each column represents a unique feature extracted from the original input. These features can then be fed into lightweight classifiers, such as K-Nearest Neighbors (KNN), Random Forest (RF), or Multilayer Perceptron (MLP). Using this tabular format allows a range of experiments with minimal computational resources, as feature extraction is performed only once. The extracted feature vectors enable a flexible, resource-friendly pipeline, facilitating rapid experimentation with various classifiers without requiring repeated passes through the CNN layers.

The methodology used in the study is shown in Figure 1.

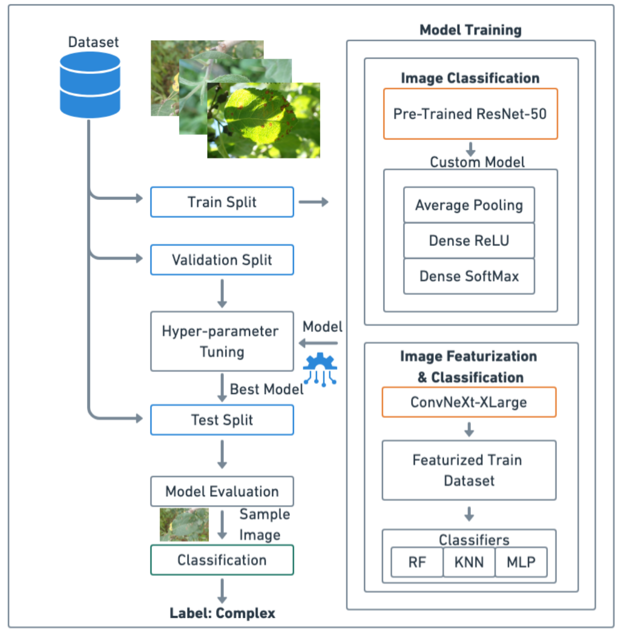


Fig. 1: Block Diagram of Proposed Methodology

First, the dataset was split into Train, Validation, and Test datasets. The training dataset was fed as an input to both approaches where ResNet-50 was used for Fine-tuning and ConvxNeXT-Xlarge was used for feature extraction. In the fine-tuning approach, most of the layers from the pre-trained model were retained and new layers specific to the apple leaf disease classification task were added. In the featurization approach, the feature maps that are extracted from ConvxNeXT-Xlarge are then fed into shallow classifiers like Random Forest(RF), K-Nearest-Neighbors(KNN), and Multilayer Perceptron(MLP). In KNN, the feature vector distances to training samples determine the label based on majority class voting among nearest neighbors. RF uses the vectors to train multiple decision trees across subsets, combining predictions through majority voting. MLP leverages the vectors through a network of hidden layers, learning complex patterns with non-linear activation functions and backpropagation.

After training the models on the training dataset, we used the validation dataset to assess it's performance. The validation data helps in fine-tuning the hyper parameters which control the model's behavior by evaluating how well the model performs on the validation dataset. We can adjust the hyper-parameters to enhance it's predictive accuracy. This step is essential because it helps prevent over-fitting, ensuring that the model generalizes well to new, unseen data. Once the optimal hyper-parameters are found, the model's final performance is tested on the test dataset.

IV. Results and Discussion

The ResNet-50 model, pre-trained on the ImageNet dataset, was adapted by appending a fully connected classification layer matching the number of target classes, utilizing a softmax activation function to enable multiclass classification. For model fine-tuning, the initial layers of ResNet-50 were frozen to retain previously learned features, while the appended classification layer was trained with various learning rates ($learning\_rate = \{0.01, 0.001, 0.0001\}$) and epoch counts ($epochs = \{10, 20, 30, 40, 50\}$) to optimize performance. Early stopping criteria based on validation set performance were used to halt training when additional epochs provided minimal improvement. The results from ResNet50 show clear trends across different learning rates and epochs, indicating distinct impacts on model accuracy. For the highest learning rate of 0.01, the accuracy remains consistently low and fails to improve significantly with increased training. This suggests that 0.01 is too high for this model and appears to be unsuitable for this classification task as it leads to erratic learning without improvement in accuracy. In contrast, the lower learning rates of 0.001 and 0.0001 demonstrate substantially better performance. For 0.001, accuracy initially rises from 0.8435 at 10 epochs to a peak of 0.8639 at 30 epochs before slightly decreasing to 0.8500 at 50 epochs. This pattern suggests that while 0.001 performs well initially, prolonged training may lead to slight overfitting, as indicated by the slight drop in accuracy after 30 epochs. The lowest learning rate, 0.0001, consistently achieves the highest accuracy across most epochs. Starting at 0.8650 at 10 epochs, the model's accuracy decreases slightly to 0.8618 by 20 epochs and stabilizes in the range of 0.8545 to 0.8579 through to 50 epochs. The trend at 0.0001 suggests that it allows for steady learning, maintaining accuracy without overfitting, but with a slight adjustment after the initial epochs. As shown in Figure 2, the configuration with $learning\_rate = 0.0001$ and $epochs = 10$ achieved the highest validation accuracy and was selected as the optimal setup, balancing accuracy and generalization. Overall, these findings indicate that the optimal learning rate for this model is 0.0001, providing the best balance between convergence and generalization. The results suggest that early stopping around 10 epochs for 0.0001 may preserve the highest accuracy while minimizing overfitting risks, making it the most effective configuration for achieving robust performance in this image classification task. The optimized model was subsequently evaluated on the test set, with output class probabilities mapped to discrete labels for classification.

All the images pass through through all CNN layers of the ConvxNets pre-trained model with ImageNet weights, followed by a global average pooling layer reduces the feature map to a final 2048-dimensional vector (1x1x2048). This vector serves as a feature summary, which can be utilized by classifier models such as K-Nearest Neighbors (KNN), Random Forest (RF), and Multilayer Perceptron (MLP) for classification.

Hyperparameter tuning is performed to optimize KNN, RF, and MLP performance. For KNN, the primary parameter is the number of neighbors ($k = \{2, 3, ..., 14\}$). RF tuning focuses on tree count ($n\_estimator = \{20, 40, ..., 100\}$) and depth ($max\_depth = \{1, 2, ..., 7\}$). MLP tuning includes learning rate ($learning\_rate\_init = \{0.01, 0.05, 0.001, 0.0001, 0.00001\}$) and epochs ($max\_iter = \{100, 120, ..., 180\}$). The optimal number of neighbors, ($k = 13$), was selected because it resulted in the highest validation accuracy. Too few neighbors can make the model sensitive to noise in the training data, while too many neighbors lead to over-smoothing the decision boundary. ($k = 13$) strikes a balance, providing a model that is both robust and accurate for the given dataset. For the Random Forest algorithm, the combination of ($max\_depth = 7, n\_estimators = 100$), emerged as optimal. Setting a maximum depth of 7 ensures that each decision tree in the forest is constrained from becoming too deep, thus preventing overfitting while maintaining sufficient complexity to capture the underlying data structure. Meanwhile, using 100 estimators ensures a diverse ensemble of trees, which enhances the robustness and accuracy of the model predictions through effective averaging. In the case of MLP, the parameters ($learning\_rate\_init = 0.001, max\_iter = 160$) were found to yield the best performance. A learning rate of 0.001 provides a stable convergence during training, avoiding the risks of overshooting the minimum loss function that might occur with a higher learning rate. Furthermore, ($max\_iter = 160$) ensures the model has enough epochs to fully learn from the data without excessive computation, preventing the model from underfitting due to insufficient training. The best model configuration is then evaluated on the test dataset to validate generalization on unseen data.

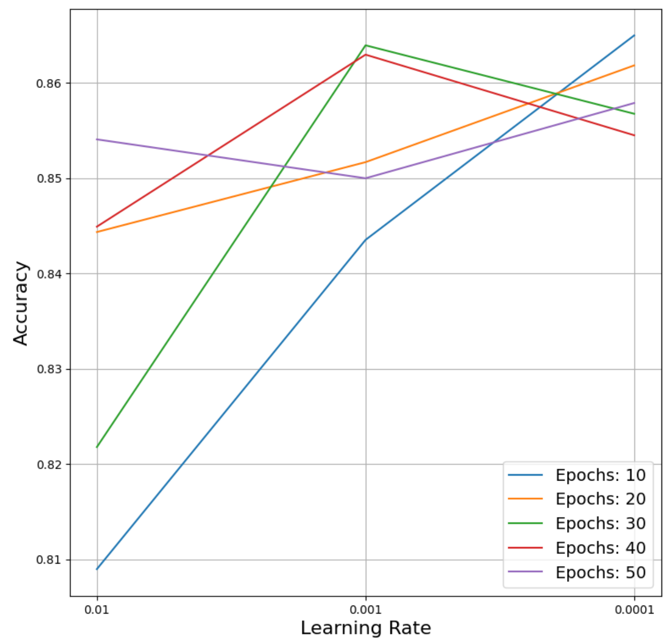
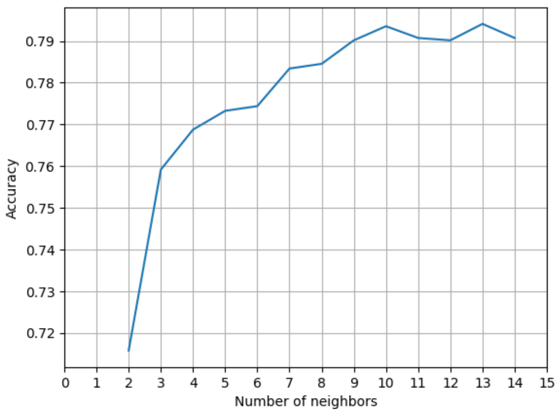
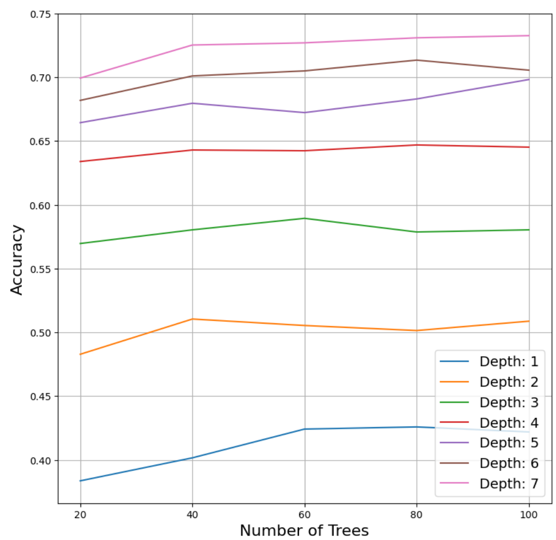


Fig. 2: Hyper Parameter Tuning of ResNet-50

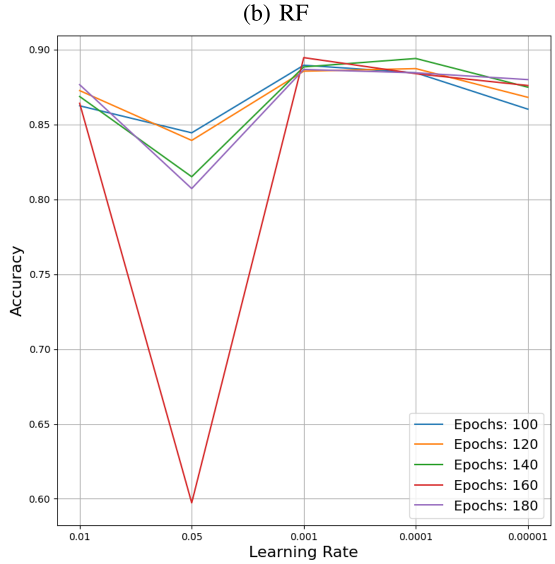
This study evaluates the performance of pre-trained models in image classification by comparing fine-tuning of ResNet-50 and the featurization approach of ConvNeXt-Xlarge. Perfor-



(a) KNN



(b) RF



(c) MLP

Fig. 3: Hyper Parameter Tuning of Shallow Classifiers

mance was assessed using classification accuracy Accuracy = $\frac{\text{TP + TN}}{\text{TP + TN + FP + FN}}$, where TP and TN represent correct predictions of positive and negative classes, and FP (Type I Error) and FN (Type II Error) are incorrect predictions. Precision Precision = $\frac{\text{TP}}{\text{TP + FP}}$, Recall = $\frac{\text{TP}}{\text{TP + FN}}$, and F1 Score F1 = $\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$ were also used to evaluate model performance comprehensively in terms of accuracy and error minimization.

The performance comparison Table I highlights the effectiveness of different model architectures. The primary goal in classification tasks is to achieve a high degree of accuracy, precision, recall, and F1 score, indicating a model's capability to correctly identify positive instances while minimizing false positives and false negatives.The finetuning of ResNet-50 leads to a notable performance gain. By adjusting the pre-trained weights to the target dataset, the model demonstrated enhanced capabilities in capturing intricate patterns, thereby improving classification accuracy and providing robust precision and recall values. ResNet-50, on the other hand, holds a competitive accuracy of 86.50%, with precision and recall both at 0.85 and an F1 score of 84%. However, the process requires significant computational resources and time for the large dataset which might pose a limitation for the real-time deployment with constrained resources.

The ConvNeXt-Xlarge featurization approach presents a compelling alternative with respect to computational efficiency and response time. ConvNeXt-Xlarge produces high-quality embedding that provides a strong foundation for subsequent classification tasks. In our experiments, ConvNeXt-Xlarge achieved a competitive accuracy and F-1 score, highlighting its effectiveness in exploiting pre-trained knowledge to deliver efficient results. ConvNeXT+KNN exhibits a reasonable accuracy of 79.41%, with a balance between precision (0.80) and recall (0.79), but its F1 score is relatively low at 69%, pointing to some imbalance between precision and recall. This suggests the model successfully identifies true positives but struggles with distinguishing these from false positives. Conversely, ConvNeXT+RF presents lower accuracy at 73.27%, with closely matched precision (0.75) and recall (0.73), reflected in a consistent F1 score (73%). This indicates the model is balanced but lacks the strengthening power required for better discrimination in more challenging classification tasks. ConvNeXT+MLP achieves the highest accuracy at 87.50%, with both precision and recall at 0.88 and an F1 score of 87%. ConvNeXT+MLP is justified as the best model due to its superior balance across accuracy, precision, recall, and F1 score. Moreover, the precision and recall metrics being equal ensure that the model is well-tuned for both detecting true positives and minimizing false positives. Overall, the comparisons revealed that the integration of the ConvNeXT architecture, known for its ability to extract rich features, with the MLP's flexibility in handling non-linearity and complex decision boundaries, results in an optimized model and yielded superior performance when compared to ResNet-50 on the validation dataset.

The testing was done on MLP models and ResNet-50 to eval-

TABLE I: Summary of Deep CNN Model Evaluation on test data

Algorithm	Accuracy	Precision	Recall	F1 Score
ConvNeXt+KNN	79.41%	0.80	0.79	69%
ConvNeXt+RF	73.27%	0.75	0.73	73%
ConvNeXt+MLP	87.50%	0.88	0.88	87%
ResNet-50	86.50%	0.85	0.85	84%

uate their robustness on new unseen data. The model ResNet-50 demonstrated an accuracy of 85% on the test dataset. The MLP model trained on features extracted from ConvNeXt-Xlarge achieved a test accuracy of 90.32%, marking a significant improvement over ResNet-50. The high precision, recall, and F-1 scores demonstrate that the model is proficient in accurately identifying true positives while minimizing false positives (Type I Error) and false negatives (Type II Error), thus ensuring effective classification performance across various situations.

We utilized confusion matrices to provide a comprehensive understanding of the model's performance across individual classes. By comparing confusion matrices of finetuned ResNet-50 and ConvxNeXT-Xlarge, we were able to gain insights into each model's limitations in handling specific classes within the dataset. The confusion matrix for the finetuned ResNet-50 reveals a strong overall performance with a high number of True Positive predictions across Healthy, Powdery Mildew, and Rust classes. This highlights the model's ability to effectively learn and adapt to unique patterns in the data. However, certain misclassifications were observed in the Complex, Frog Eye Leaf Spot, and Scab classes with subtle feature differences, which could be attributed to similarities in visual characteristics that challenge the model's ability. Notably, ResNet-50 exhibited a high number of False Positives in these cases, suggesting areas for improvement in enhancing Recall, particularly for the more complex classes.

In contrast, the ConvxNext-Xlarge model's confusion matrix showcases a balanced accuracy across various classes. ConvxNext-Xlarge demonstrated a reduced number of False Positives compared to ResNet-50, indicating superior Precision in several instances. This highlights ConvxNext-Xlarge strength in feature extraction, allowing it to distinguish between classes with clear boundaries. However, certain classes with intricate details still pose challenges reflected in the modest rise in False Negatives. This could be due to the model's reliance on feature embeddings, which sometimes fell short without extensive finetuning. The ResNet-50 tends to show more variability in misclassification patterns while ConvxNext-Xlarge is generally more stable.

V. Conclusion

In this study, we explored a data-driven approach to apple disease detection by exploiting the fine-tuning capabilities of the ResNet-50 model and the advanced feature extraction of ConvNeXt-Xlarge architecture. Our results demonstrate that ConvNeXt-Xlarge effectively enhances classification accuracy

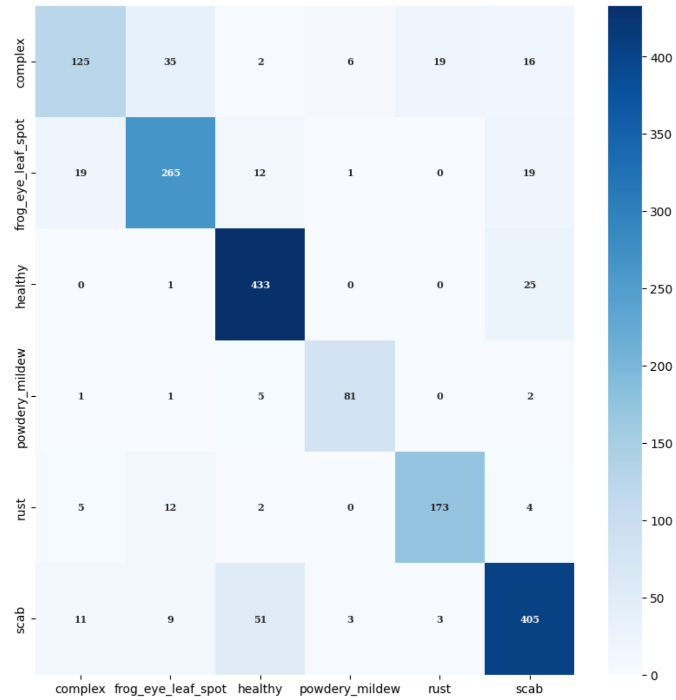


Fig. 4: Confusion Matrix for ResNet-50

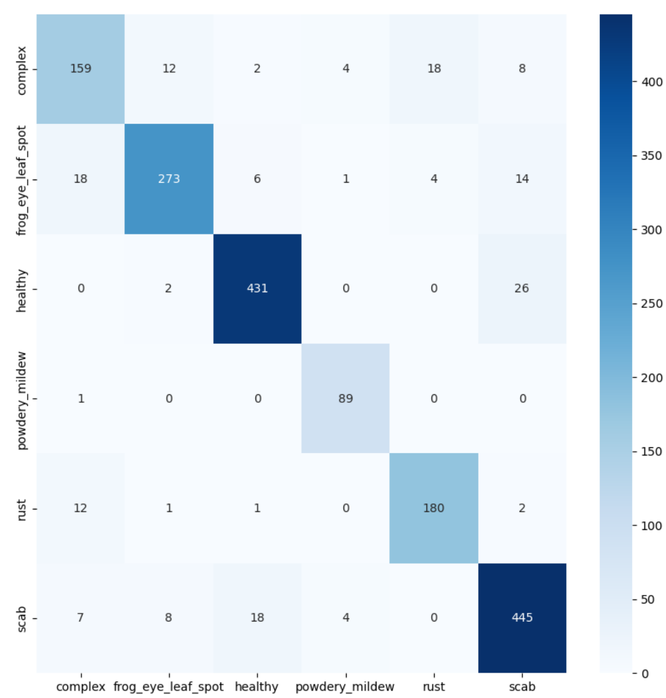


Fig. 5: Confusion Matrix for ConvNeXt-Xlarge+MLP

across diverse disease categories, outperforming ResNet-50. The implementation of our approach provides several advantages, like reduced training time and improved model generalization, showcasing the potential of using pre-trained models in agricultural domains. However, the feature extraction using a pre-trained model may result in sub-optimal performance if the extracted features are not well aligned with the characteristics necessary for different disease classifications which further requires tuning. Further research could focus on the real-time deployment of various state-of-the-art models in field conditions. Our findings suggest that applications of Deep-Learning techniques hold significant promise for crop health monitoring systems worldwide.

References

[1] R. N. Strange and P. R. Scott, "Plant disease: a threat to global food security," *Annu. Rev. Phytopathol.*, vol. 43, no. 1, pp. 83–116, 2005.

[2] Y. Zhong and M. Zhao, "Research on deep learning in apple leaf disease recognition," *Computers and electronics in agriculture*, vol. 168, p. 105146, 2020.

[3] M. Shahbandeh, "Global fruit production in 2022, by selected variety (in million metric tons)." Published by M. Shahbandeh, Statista, May 22, 2024, May 2024. Accessed: 2024-02-06.

[4] U. Apple, "Industry at a glance." Available at ://usapple.org/industry-at-a-glance.

[5] J. L. Norelli, A. L. Jones, and H. S. Aldwinckle, "Fire blight management in the twenty-first century: using new technologies that enhance host resistance in apple," *Plant Disease*, vol. 87, no. 7, pp. 756–765, 2003.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[7] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*. MIT press, 2018.

[8] I. Goodfellow, "Deep learning," 2016.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[10] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Computers and electronics in agriculture*, vol. 145, pp. 311–318, 2018.

[11] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.

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

[13] G. Wang, Y. Sun, J. Wang, and Q. Liu, "Automatic image-based plant disease severity estimation using deep learning," in *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*, pp. 60–64, IEEE, 2017.

[14] H. Ayaz, E. Rodríguez-Esparza, M. Ahmad, D. Oliva, M. Pérez-Cisneros, and R. Sarkar, "Classification of apple disease based on non-linear deep features," *Applied Sciences*, vol. 11, no. 14, p. 6422, 2021.

[15] Y. Zeng, "Enhancing image classification accuracy based on alexnet," *Highlights in Science, Engineering and Technology*, vol. 85, pp. 879–885, 2024.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

[17] X. Chao, G. Sun, H. Zhao, M. Li, and D. He, "Identification of apple tree leaf diseases based on deep learning models," *Symmetry*, vol. 12, no. 7, p. 1065, 2020.

[18] S. Li, H. Qu, X. Dong, B. Dang, H. Zang, and Y. Gong, "Leveraging deep learning and xception architecture for high-accuracy mri classification in alzheimer diagnosis," *arXiv preprint arXiv:2403.16212*, 2024.

[19] P. Bansal, R. Kumar, and S. Kumar, "Disease detection in apple leaves using deep convolutional neural network," *Agriculture*, vol. 11, no. 7, p. 617, 2021.

[20] M. Latha, P. S. Kumar, R. R. Chandrika, T. Mahesh, V. V. Kumar, and S. Guluwadi, "Revolutionizing breast ultrasound diagnostics with efficientnet-b7 and explainable ai," *BMC Medical Imaging*, vol. 24, no. 1, p. 230, 2024.

[21] A. Radhakrishnan, J. Davis, Z. Rabin, B. Lewis, M. Scherreik, and R. Ilin, "Design choices for enhancing noisy student self-training," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 1926–1935, 2024.

[22] V. K. Vishnoi, K. Kumar, B. Kumar, S. Mohan, and A. A. Khan, "Detection of apple plant diseases using leaf images through convolutional neural network," *IEEE Access*, vol. 11, pp. 6594–6609, 2022.

[23] S. Mascarenhas and M. Agarwal, "A comparison between vgg16, vgg19 and resnet50 architecture frameworks for image classification," in *2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON)*, vol. 1, pp. 96–99, IEEE, 2021.

[24] S. Kim, J. Lee, K. Jeong, J. Lee, T. Hong, and J. An, "Automated door placement in architectural plans through combined deep-learning networks of resnet-50 and pix2pix-gan," *Expert Systems with Applications*, vol. 244, p. 122932, 2024.

[25] H. Yu, X. Cheng, C. Chen, A. A. Heidari, J. Liu, Z. Cai, and H. Chen, "Apple leaf disease recognition method with improved residual network," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 7759–7782, 2022.

[26] C. Bi, J. Wang, Y. Duan, B. Fu, J.-R. Kang, and Y. Shi, "Mobilenet based apple leaf diseases identification," *Mobile Networks and Applications*, pp. 1–9, 2022.

[27] M. A. B. Abbass and Y. Ban, "Mobilenet-based architecture for distracted human driver detection of autonomous cars," *Electronics*, vol. 13, no. 2, p. 365, 2024.

[28] B. Karsh, R. H. Laskar, and R. K. Karsh, "miv3net: modified inception v3 network for hand gesture recognition," *Multimedia Tools and Applications*, vol. 83, no. 4, pp. 10587–10613, 2024.

[29] S. Chauhan, R. Chauhan, R. Rawat, S. Aluvala, and R. Singh, "Fracture detection in radiology with resnet-152 & cnn," in *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, pp. 1–6, IEEE, 2024.

[30] A. I. Khan, S. Quadri, S. Banday, and J. L. Shah, "Deep diagnosis: A real-time apple leaf disease detection system based on deep learning," *computers and Electronics in Agriculture*, vol. 198, p. 107093, 2022.

[31] M. M. Islam, M. A. Talukder, M. A. Uddin, A. Akhter, and M. Khalid, "Brainnet: precision brain tumor classification with optimized efficientnet architecture," *International Journal of Intelligent Systems*, vol. 2024, no. 1, p. 3583612, 2024.

[32] Q. Yan, B. Yang, W. Wang, B. Wang, P. Chen, and J. Zhang, "Apple leaf diseases recognition based on an improved convolutional neural network," *Sensors*, vol. 20, no. 12, p. 3535, 2020.

[33] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "Knn model-based approach in classification," in *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings*, pp. 986–996, Springer, 2003.

[34] S. J. Rigatti, "Random forest," *Journal of Insurance Medicine*, vol. 47, no. 1, pp. 31–39, 2017.

[35] H. Taud and J.-F. Mas, "Multilayer perceptron (mlp)," *Geomatic approaches for modeling land change scenarios*, pp. 451–455, 2018.

[36] Z. K. S. N. B. S. Thapa, Ranjita and A. Khan, "Plant pathology 2021 - fgvc8." Available at https://www.kaggle.com/c/plant-pathology-2021-fgvc8(2021/03/15).

[37] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of image classification algorithms based on convolutional neural networks," *Remote Sensing*, vol. 13, no. 22, p. 4712, 2021.

[38] Q. Hou, C.-Z. Lu, M.-M. Cheng, and J. Feng, "Conv2former: A simple transformer-style convnet for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.