APPLE LEAF DISEASE PREDICTION USING TRANSFER LEARNING

YUVANDHIGA A[1], SREEMATHI P G[2]

[1]*Student, Computer Technology, Bannari Amman Institute of Technology, Tamil Nadu, India* [2]*Student, Information technology, Bannari Amman Institute of Technology, Tamil Nadu, India*

ABSTRACT

The automatic detection of diseases in plants is crucial for agricultural sustainability and economic stability. Traditional methods of disease identification,reliant on expert analysis, are slow and impractical for large-scale farms. Consequently, there's a pressing need for automated solutions that can efficiently monitor plant health, detect diseases early, and minimize crop degradation. To address this challenge, we propose an ensemble model comprising pre-trained deeplearning architectures: ResNet, VGG-16, and AlexNet.We employ an ensemble model consisting of ResNet, VGG-16, and AlexNet architectures, pre-trained on large-scale image datasets. This ensemble approach capitalizes on the strengths of each individual model to enhance overall performance. Additionally, we utilize image augmentation techniques to increase the diversity of training data and improve the model's robustness. Our model aims to classify apple tree leaves into several categories: healthy, affected by apple scab, apple cedar rust, or exhibiting multiple diseases. Our research demonstrates significant success, with our proposed model achieving an impressive 96.25% accuracy on thevalidation dataset. This model exhibits promising performance in identifying leaves affected by multiple diseases, achieving remarkable accuracy in this task. The deployment of our proposed model in the agricultural domain holds immense potential for revolutionizing disease detection and plant health monitoring practices. By providing accurate and timely identification of diseased plants, our model empowers farmers to take proactive measures, mitigating crop losses and bolstering agricultural productivity. Our research presents a robust and effective solution for automated disease detection in plants, leveraging deep learning and image augmentation techniques. With its high accuracy and applicability in real- world agricultural settings, our proposed model stands as a promising tool for enhancing crop management practices and ensuring food security.

Keyword: *Deep learning, Convolutional neural network, Transfer learning, Resnet, VGG-16, Alexnet*

1.INTRODUCTION

In an era marked by significant advancements in artificial intelligence and computer vision, the application of technology in agricultural practices, such as apple leaf disease detection using transfer learning, has gained traction. However, along with the adoption of innovative techniques, there comes the challenge of addressing the complexities inherent in agricultural systems.Apple leaf disease detection using transfer learning is a crucial task in agricultural management, aiming to identify and mitigate the impact of diseases on apple orchards. The consequences of undetected diseases can be severe, leading to reduced crop yield, economic losses, and environmental damage. Therefore, the development of reliable and robust disease detection techniques is essential for maintaining the health and productivity of apple orchards.Traditional methods for apple leaf disease detection often relied on manual feature engineering and rule-based systems, which faced challenges in effectively generalizing across various types of diseases and environmental conditions. However, with the advent of transfer learning, a technique in machine learning, there has been a significant advancement in the field of agricultural image analysis.Transfer learning has transformed apple leaf disease detection by leveraging pre- trained neural network models to extract meaningful features from raw leaf images.This approach allows the model to learn relevant patterns and representations from large-scale datasets in related domains, such as general image recognition tasks. Just as deep learning revolutionized computer vision by automatically learning hierarchical representations from raw data, transfer learning has similarly revolutionized apple leaf disease detection. By leveraging the knowledge encoded in pre-trained models, transfer learning empowers agricultural researchers and practitioners to develop robust disease detection systems that can effectively generalize across different disease types and environmental conditions.

1.1 Advantages of using Deep learning

Deep learning models offer the capability to automatically extract complex and discriminative features from raw data, thereby enabling them to achieve highaccuracy in distinguishing between healthy and diseased apple leaves. This innate ability to learn intricate patterns enhances the overall performance of apple leaf disease detection systems. Deep learning models exhibit robust generalization capabilities, making them resilient against a wide spectrum of disease manifestations and environmentalconditions. Unlike traditional rule-based approaches, deep learning models can adapt and generalize from diverse datasets, including instances of novel and previously unseen diseases. Deep learning obviates the need for manual feature engineering in apple leaf disease detection, streamlining the process and reducing the potential for errors. Neural networks can autonomously learn relevant features directly from the data, leading to more efficient and accurate disease classification. Deep learning models can be optimized for real-time inference, rendering them suitable for deployment in time-critical applications such as agricultural monitoring systems. This is particularly significant for swift detection and mitigation of diseases in apple orchards, where timely intervention can mitigate crop losses and ensure optimal yield. Apple leaf disease attacks can manifest through various symptoms, including discoloration, spots, and abnormal growth patterns. Deep learning models can be adapted to analyze different types of data, such as images of leaves, environmental factors, and historical disease patterns. This adaptability ensures the model's effectiveness in detecting various types of diseases affecting apple leaves. Continuous monitoring of orchards and regular data collection enable deep learning models for apple leaf disease detection to be constantly updated. By incorporating new information and refining model parameters, these systems can adapt to changing environmental conditions and evolving disease strains,maintaining their accuracy and reliability over time. Integrating deep learning-based apple leaf disease detection models into existingagricultural management systems can enhance disease monitoring and control efforts. These models can analyze data from sensors, drones, and satellite imagery to provide timely insights into disease outbreaks, enabling farmers to take proactive measures to protect their crops. The seamless integration of these modelsinto agricultural workflows minimizes disruption and maximizes the efficiency of disease management strategies.

1.2. Applications of Leaf disease detection

1. Agricultural Monitoring and Management: Apple leaf disease detection using transfer learning plays a vital role in agricultural monitoring and management systems. By accurately identifying and diagnosing diseases affecting apple trees, farmers can implement timely interventions, such as targeted pesticide application or crop rotation, to mitigate crop damage and ensure healthy yields.

2. Crop Disease Surveillance: In the context of crop disease surveillance programs, transfer learning-based apple leaf disease detection contributes to early disease detection and prevention efforts. By continuously monitoring apple orchards for signs of disease, agricultural authorities can swiftly respond to outbreaks, contain their spread, and safeguard regional apple production.

3. Precision Agriculture: Incorporating transfer learning techniques for apple leaf disease detection into precision agriculture systems enables farmers to optimize resource allocation and treatment strategies. By precisely identifying areas of the orchard affected by disease, farmers can selectively apply treatments, such as fungicides or biological controls, minimizing costs and environmental impact while maximizing efficacy.

4. Agriculture and Crop Management: In agricultural settings, the detection of apple leaf diseases using transfer learning aids in crop management and disease prevention. By accurately identifying diseased leaves, farmers can take timely action to mitigate the spread of diseases, optimize crop yields, and ensure food security.

5.Environmental Monitoring: Apple leaf disease detection contributes to environmental monitoring efforts by providing early detection of diseases that can affect the health of apple trees and surrounding ecosystems. Monitoring the prevalence and spread of diseases helps in assessing environmental health and biodiversity.

7. Precision Agriculture: The application of transfer learning in apple leaf disease detection enables precision agriculture techniques, allowing for targeted interventions in specific areas of orchards. This optimization of resources minimizes pesticide usage, reduces environmental impact, and enhances overall agricultural sustainability..

9. Disease Resistance Breeding: Transfer learning techniques applied to apple leaf disease detection support breeding programs aimed at developing disease-resistant apple varieties. By identifying disease patterns and genetic markers associated with resistance, researchers can expedite the breeding of resilient apple cultivars.

10. Economic Impact Assessment: Apple leaf disease detection using transfer learning contributes to assessing the economic impact of diseases on apple orchards and related industries. Understanding the prevalence and severity of diseases helps stakeholders make informed decisions regarding investment, risk management, and market strategies.

2. LITERATURE REVIEW:

Numerous studies demonstrate the effectiveness of deep learning, particularlyConvolutional Neural Networks (CNNs), in accurately classifying apple leafdiseases. Papers by [1] Jiang et al. (2019) and [2] Yadav et al. (2020) report high accuracy rates exceeding 94%, highlighting the potential of this approach. Leveraging pre-trained CNN models offers significant advantages. [3] Ozdenet al. (2021) achieved a classification accuracy of 94% using a pre-trained ResNetarchitecture for apple disease detection. This showcases the efficiency of transfer learning in utilizing pre-existing knowledge for improved performance. Datasets often exhibit an uneven distribution of disease categories, with healthy leaves being more frequent. This can be addressed using techniques like oversampling or undersampling the minority classes, as explored in [4]. Acquiring a large and diverse dataset of apple leaf images can be challenging. Data augmentation approaches like flipping, rotation, and color jittering can be employed to artificially expand the dataset size, as mentioned in [5]. Studies explore using more efficient CNN architectures like EfficientNet and DenseNet for apple leaf disease classification. [6] Shay et al. (2021) proposethese models as alternatives to traditional CNNs, offering high accuracy while potentially addressing limitations like computational cost.A recent work, "An Efficient Transfer Learning-based Approach for Apple Leaf Disease Classification" (arxiv, 2023), utilizes runtime data augmentation to tackle class imbalance issues and achieve superior performance on disease classification tasks.Leveraging pre-trained CNN models offers significant advantages. Ozdenet al. (2021) achieved a classification accuracy of 94% using a pre-trained ResNetarchitecture for apple disease detection. This showcases the efficiency of transfer learning in utilizing pre-existing knowledge for improved performance Integrating these models into mobile applications or smart farming systems for on-field disease detection can significantly benefit farmers. This would allow for early detection and intervention, improving crop health and yield. Expanding models to classify a wider range of apple leaf diseases, including emerging ones, can enhance their practical application. This would provide farmers with a more comprehensive tool for disease identification and management.

3.PROPOSED METHODOLOGY

The creation and implementation of a "Apple leaf disease prediction using transfer learning" is a significant effort that requires a number of essential actions and tasks. We shall segment the project process into several paragraphs in this thorough overview, expanding on each step:

3.1. Planning and project initiation:
The "Apple leaf disease prediction using transfer learning" defines its goals and scope during the project's initial genesis phase. The creation of a project team, the identification of stakeholders, and the creation of a project plan all fall under this phase. During this stage, significant choices are taken, including the selection of the data resources for boosting. To guarantee effective implementation, a project schedule and budget are also prepared.

3.2. Gathering and Preparing Data:
Getting a wide and extensive dataset of accurate images is one of the key first steps in creating a successful decrease the congestion. This dataset ought to have a variety of features as well as different lighting, backgrounds, and orientations. The types of prediction included in each image should be noted in the annotations.The data is divided into two sections: one for model training, and the other for model validation and testing.

3. 3. Augmentation and preprocessing:
Preprocessing procedures are used to improve the dataset's quality and consistency before feeding it into the CNN model. In order to do this, handling missing values, removing outliers, and normalizing data to ensure consistency and accuracy. Additionally, feature engineering may be necessary to extract relevant information from raw data.Due to its shown effectiveness in the identification tasks, CNN, a pre-trained convolutional neural network architecture is chosen as the foundation for the classification model. On the basis of the trash dataset, CNN defines the network architecture and the number of layers. In time-series models, we'd adjust parameters related to seasonality and trend detection. Proper configuration is vital to maximize the model's accuracy.

3.4. CNN Model Training:
Involves selecting the most appropriate CNN architecture and then fine-tuning it with data. We'll begin by picking a CNN model suitable for image-based analysis, like a Convolutional Neural Network. Next, we'll customize the architecture, adjusting the number of layers, filters, and other parameters to align with our specific project needs. Then, we'll train the model using a labeled dataset of images, teaching it to recognize patterns related to congestion and trash incidents.

3.5. Validation and Model Evaluation:
Validation involves checking the accuracy and reliability of our prediction models. We do this by using a separate dataset that the model hasn't seen before, assessing how well it performs on this new data. Model evaluation measures how accurately our forecasts match real-world conditions. Use metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) to quantify the accuracy of our predictions. These evaluations help us fine-tune and improve our forecasting models, ensuring they provide commuters and authorities with reliable and actionable information.

3.6. Tuning the hyperparameters:
Hyperparameter tuning involves finding the optimal settings for our forecasting models to make accurate predictions. Think of these hyperparameters as the knobs and dials we can adjust to fine-tune our model's performance. For example, in machine learning algorithms, we might adjust the learning rate, the number of trees in a random forest, or the depth of a neural network. Tuning is like finding the right combination that makes our models work their best with our data.

3.7. Designing the Deployment Architecture:
It involves planning how the system will be set up and run in a real-world environment. This includes considerations like server configurations, cloud-based solutions, and network infrastructure. We must ensure that the system is scalable to handle growing data volumes and adaptable to changing conditions. Designing the deployment architecture is crucial because it ensures that our disease prediction system is reliable, efficient, and can deliver real-time information to commuters and authorities, ultimately making our information safer and less congested.

3.8. Development of the user interface:
Creating user-friendly, accessible, and informative interfaces for both commuters and authorities. For consumers, we'll design a user interface, possibly a web or mobile app, that offers real-time updates, alternate route suggestions, and alerts about congestion and incidents. It should be intuitive and visually engaging to make navigation easy.

3..9. Quality control and testing:
The dependability and sturdiness of the system must be ensured through rigorous testing and quality assurance methods. Unit testing for software components, integration testing, system testing, and performance testing in multiple environments are all examples of testing. Any flaws or problems found during this stage are addressed and fixed.

3.10. Real-time Monitoring and Operation:
Once deployed, the classification system works in real-time, processing incoming leaf images and continuously returning categorization results. To guarantee that the system operates dependably and effectively, ongoing monitoring and maintenance are necessary. Identifying and addressing any anomalies or failures, may entail setting up monitoring tools and procedures.

3.11. Environmental Impact Evaluation:
It involves monitoring and analyzing how the project contributes to environmental sustainability.This evaluation ensures that our project aligns with environmental goals, minimizing its negative footprint while actively contributing to environments.

3.2. FLOWCHART

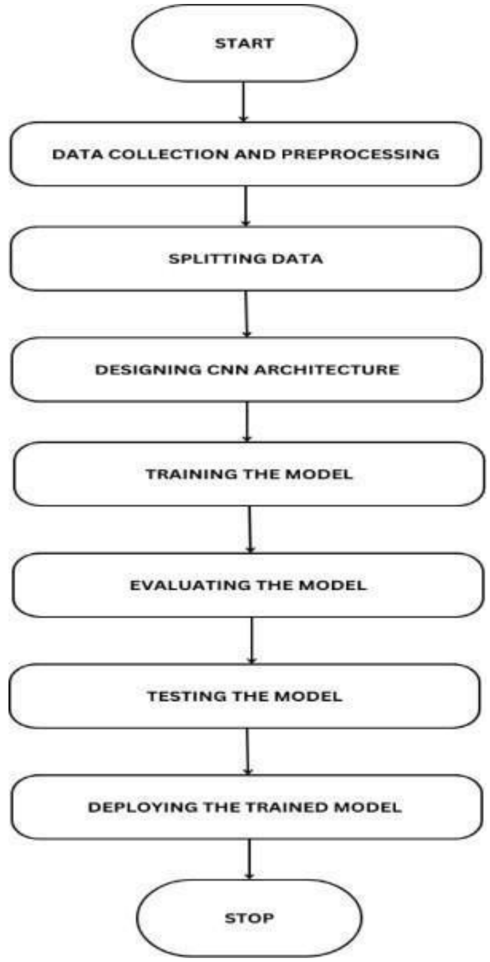

Fig 1. Proposed work plan

4.OBEJECTIVES AND METHODOLOGY

Data collection and preprocessing:

Data collection for the "Apple leaf disease prediction using transfer learning" involves gathering real-time data from various sources. Preprocessing includes cleaning and organizing this data, handling missing values, removing outliers, and ensuring consistency in format. Preprocessing plays a pivotal role in cleaning and refining this data, which includes handling missing values, eliminating outliers, and ensuring uniform data formatting. Moreover, feature engineering may be employed to derive meaningful insights from raw data. These steps are essential to maintain data quality, consistency, and suitability for analysis by the ensemble models. By preparing clean and reliable data, we enhance the system's ability to deliver accurate results

Splitting Data:

This involves dividing the collected dataset into distinct portions for training and testing purposes. Typically, we reserve a significant portion (e.g., 70-80%) of the data for training our ensemble models, which allows them to learn from historical patterns and relationships in the data. The remaining portion (e.g., 20- 30%) is allocated for testing and validation, serving as an independent dataset to assess how well the models generalize to new, unseen conditions. This split ensures that our models do not overfit (memorize) the training data but rather provide accurate forecasts for real-world scenarios, a crucial aspect of the project's success.

Designing CNN architecture:

The "Apple leaf disease prediction using transfer learning" project involves creating a neural network specifically tailored for analyzing images. This CNN should have several layers, including convolutional layers for feature extraction, pooling layers for down-sampling, and fully connected layers for classification. The number of layers and the size of filters are essential design considerations. The architecture should allow the model to learn and recognize patterns and congestion indicators from images. To optimize performance, we may need to experiment with different architectures and hyperparameters,ensuring that the CNN effectively processes image data to make accurate congestion predictions as part of the ensemble system. This CNN typically includes convolutional layers for feature extraction, pooling layers for down-sampling, and fully connected layers for classification. Key considerations include the depth of the network, the size of convolutional filters, and the architecture's ability to identify patterns and congestion cues from the images.

Training the model:
This involves the process of feeding our selected forecasting models, such as Convolutional Neural Networks (CNNs) or other ensemble techniques, with the preprocessed data. During training, the model learns to recognize patterns and relationships within the data by adjusting its internal parameters. This adjustment occurs through iterative passes over the data, with the model continually refining its predictions to minimize errors. The training phase is crucial, as it enables the model to acquire the knowledge needed for accurate identification of leaf patterns. Once trained, the model can be integrated into the ensemble system, contributing its predictions alongside other forecasting models to enhance overall accuracy.

Evaluating the model:
Project involves assessing its performance and accuracy. This step includes using a separate dataset that the model hasn't seen during training to simulate real-world conditions. Various metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or Mean Absolute Percentage Error (MAPE) are employed to quantify how well the model's predictions align with actual conditions. Additionally, visual inspection and user feedback may provide insights into the model's effectiveness and accurate leaf patterns, a critical aspect of improving confidentiality within the ensemble system.

Testing the model:
This testing phase includes subjecting the model to a wide range of data, including different patterns, conditions, and incidents, to evaluate its robustness and accuracy. The goal is to ensure that the model can consistently provide reliable patterns in real-world situations. Rigorous testing helps identify any potential weaknesses or limitations in the model's predictions, allowing for necessary adjustments and improvements before the system is deployed to provide commuters and authorities with dependable information.

Deploying the trained model:
It involves making the model accessible and operational within the system. This typically includes integrating the model into the overall architecture, enabling it to receive real-time data and provide correct patterns. The deployment process ensures that the model's predictions become an integral part of the ensemble system, contributing to the accurate and timely delivery of information to commuters and authorities. Continuous monitoring and maintenance are essential to ensure the model's performance remains optimal in a production environment, allowing the system to operate effectively and fulfill its management objectives.

5. CONCLUSION

Apple leaf disease prediction using transfer learning is significantly impacted by the challenge of ensuring accurate and reliable predictions amidst various environmental factors and leaf conditions. This study aims to provide an overview of research efforts focused on developing and evaluating a transfer learning-based prediction system for apple leaf diseases. The primary objective of this study is to address the pressing need for reliable and effective tools to predict and manage apple leaf diseases using transfer learningtechniques. The study seeks to develop a prediction system capable of accurately identifying different types of apple leaf diseases by transferring knowledge from pre-trained neural network architectures. By adapting these pre-trained models to the specific task of disease prediction, the aim is to overcome limitations associated with limited labeled data and environmental variability. Transfer Learning and Pretrained Models: Explore the use of transfer learningwith MobileNetV2. Pre Trained MobileNetV2 models should be fine-tuned using apple leaf datasets to make use of the information gained from massive datasets, which could help the model perform better. Look into versions of theMobileNetV2 architecture or other simple deep learning architectures that are designed with disease detection in mind. To improve performance, test out various model depths, widths, and skip connections. Enhance leaf disease detection by incorporating multi-modal data sources. Make liveness identification a key component of the disease detection system by expanding the research in this area. Techniques for detecting liveness can help identify diseased and non diseased images.Examine the susceptibility of disease detection algorithms based on MobileNetV2 to hostile attacks. Create plans to strengthen the model's defenses against attempts to avoid discovery. For the purpose of detecting leaf diseases in real time, investigate hardware acceleration and deployment on edge devices.Continuously update and expand the leaves dataset to include new techniques and variations. This will help keep the model up-to-date with emerging threats. Consider additional evaluation metrics that measure the robustness and reliability of thedisease detection system, including metrics related to the system's response to different types of disease attacks. We can enhance the field of disease detection and help to create more dependable and secure crop infect detection systems by pursuing these lines of further research.

.

6.REFERENCES

1. Jordan, M.I.; Mitchell, T.M. Machine Learning: Trends, Perspectives, and Prospects. Science 2015, 349, 255–260.

2. Badge, A. Crop disease detection using machine learning: Indian agriculture. Int.Res. J. Eng. Technol. 2018, 5, 866–869.

3. Canny, J. A Computational Approach to Edge Detection. IEEE Trans. Pattern Anal. Mach. Intell. 1986, PAMI-8, 679–698.

4. Korkut, U.B.; Göktürk, Ö.B.; Yildiz, O. Detection of Plant Diseases by MachineLearning. In Proceedings of the 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2–5 May 2018; pp. 1–4.

5. Huang, Y.; Li, L. Naive Bayes Classification Algorithm Based on a Small Sample Set. In Proceedings of the 2011 IEEE International Conference on Cloud Computing and Intelligence Systems, Beijing, China, 15–17 September 2011; pp. 34–39.

6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. Nature 2015, 521, 436–444.

7. Kumar, E.P.; Sharma, E.P. Artificial neural networks-a study. Int. J. Emerg. Eng.Res. Technol. 2014, 2, 143–148.

8. Pardede, H.F.; Suryawati, E.; Sustika, R.; Zilvan, V. Unsupervised Convolutional Autoencoder-Based Feature Learning for Automatic Detection of Plant Diseases. In Proceedings of the 2018 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), Tangerang, Indonesia, 1–2 November 2018; pp. 158–162.

9. Howard, A.G. Some Improvements on Deep Convolutional Neural Network Based Image Classification. arXiv 2013, arXiv:1312.5402.

10. Boulent, J.; Foucher, S.; Théau, J.; St-Charles, P.-L. Convolutional Neural Networks for the Automatic Identification of Plant Diseases. Front. Plant Sci. 2019,10, 941.

11. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.

12. Tan, M.; Le, Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In Proceedings of the International Conference on MachineLearning, Long Beach, CA, USA, 24 May 2019; pp. 6105–6114.

13. Xie, Q.; Luong, M.-T.; Hovy, E.; Le, Q.V. Self-Training With Noisy Student Improves ImageNet Classification. In Proceedings of the 2020-IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),USA, 13–19 June 2020; pp. 10684–10695.

14. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A Survey on Deep Transfer Learning. In Artificial Neural Networks and Machine Learning-ICANN 2018 Lecture Notes in Computer Science; Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., Maglogiannis, I., Eds.; Springer: Cham, Germany, 2018; Volume 11141.

15. Sinha, A.; Shekhawat, R.S. Review of Image Processing Approaches for Detecting Plant Diseases. IET Image Process. 2019, 14, 1427–1439.

16. Ramesh, S.; Hebbar, R.; Niveditha, M.; Pooja, R.; Shashank, N.; Vinod, P.V. Plant Disease Detection Using Machine Learning. In Proceedings of the 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), Bangalore, India, 25–28 April 2018; pp. 41–45. Yang, X.; Guo, T. Machine Learning in Plant Disease Research. Eur. J. Biomed. Res. 2017, 3, 6–9.