International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Hybrid Approach for Leaf Disease Detection Using Convolutional Neural Networks and Vision Transformers

Ayush Sharma[1], Ayush Singh[2], Atul Kumar[3], Ayushi Tyagi[4], Satendra Kumar[5]

[1,2,3,4,5]Department of Computer Science, Meerut Institute of Engineering and Technology, Meerut, India.

Abstract

Leaf diseases are a major problem affecting crop yield and food security in agriculture. Early and accurate detection is essential for timely intervention; however, traditional investigation methods are often labor-intensive, time-consuming, and error-prone. This study presents a hybrid model combining convolutional neural networks (CNN) and visual transformation (ViT) to overcome these problems and improve the accuracy of disease detection. The model uses CNN to capture local features such as edges and texture, while ViT to capture global relationships and generate powerful and effective solutions for disease detection. Use Plant Village information. The data were previously processed, including data enhancement techniques such as resizing, normalization and translation, and brightness adjustment to improve the model.

This hybrid model outperformed the CNN and ViT model by achieving 95% accuracy for Alternaria leaf spot and 97.38% accuracy for disease. More importantly, the model uses only 9,913 training errors, reducing the computational load and training time without compromising the accuracy and reliability expectations. The proposed model provides efficient and computationally efficient solutions for plant disease detection by combining CNN and ViT. Its low requirements make it suitable for use in a real agricultural environment, especially in limited areas.

Additionally, this approach helps mitigate the misuse of chemical treatments by enabling precise disease identification, contributing to sustainable agricultural practices.

Future research will explore expanding the dataset to include additional disease classes, optimizing the model for edge devices to support real-time diagnostics, and integrating the system into mobile applications to empower farmers globally. This work provides a foundation for modernizing agricultural disease management through advanced machine learning techniques.

Keywords: Machine literacy, Alternaria Leaf Spot, Convolutional Neural Network, Vision Transformer, Leaf Disease Detection, and Plant Health

1. Introduction

Factory health is a crucial factor in determining crop productivity, and husbandry is essential to maintaining food security. still, if left unbounded, factory conditions can affect in significant losses. Homemade examinations, which are time- consuming and prone to miscalculations, constitute the foundation of traditional complaint identification ways. Although deep literacy advances lately have showed pledge in automating this procedure, single- model approaches occasionally warrant adaptability. With an emphasis on Alternaria Leaf Spot, a current crop complaint, this study suggests a mongrel model for splint complaint discovery that combines CNNs and ViTs. To get lesser delicacy, the mongrel strategy makes use of ViTs' attention medium and CNNs' point birth capabilities. Agriculture is the foundation of any nation's frugality. numerous growers wish to embrace contemporary husbandry, but numerous are unfit to do so for a variety of reasons, similar as a lack of knowledge about the newest technologies or the precious cost of similar technologies. In numerous image processing operations, machine literacy- grounded approaches have performed well in recent time. operations of artificial intelligence in literacy have shown fruitful results. employing machine literacy approaches that educate the system to learn on its own and enhance the issues grounded on its own gests. It has constantly been noted that factory conditions are challenging to manage since their populations vary depending on the environmental conditions. shops are susceptible to a variety of ails, including bacterial, viral, and fungal bones. It has been discovered that 85 of shops are impacted by organisms that act fungi. growers in poor nations employ conventional styles, which are more labor- ferocious and time- consuming. also, homemade discovery or observation with the naked eye might not be suitable to produce useful results. also, it's noted that numerous growers use fungicides to annihilate the goods of complaint without vindicating the specific disease; this unrestricted use of fungicides can have an impact on both mortal health and factory quality. Machine literacy and deep literacy ways for fac- tory complaint discovery and bracket can help growers in feting the ails and taking the applicable control measures. When compared to conventional image processing styles, machine literacy and deep literacy algorithms for factory complaint discovery are more precise and bear lower time.

a) Plant Diseases And Its Symptoms

Here are some fundamental facts about diseases caused by bacteria, viruses, and fungi.

Bacterial illnesses: Known as bacterial diseases, these conditions induce a variety of symptoms, such as plant overgrowth, leaf spots, scabs, and cankers. The symptoms of bacterial infections and fungal diseases are almost identical. Leaf spots are the most prevalent kind of bacterial disease symptoms.

Viral diseases: It can be a bit challenging to recognize and assess viral infections. A mosaic leaf pattern, wrinkled, yellowed, and stunted leaves are signs of a viral illness. Tobacco mosaic, tomato spotted virus, cauliflower mosaic virus, potato virus, and others are some of the most common viral illnesses.

Fungal diseases: Fungal diseases are those that are frequently observed on a variety of plants. Fungal infections are the cause of a great deal of plant harm. Anthrac-nose, Downy mildews, Powdery mildews, Rusts, Rhizoctonia rots, Sclerotinia rots, and Sclerotium rots are a few of the most common fungal infections.

Difference between Machine Learning and Deep Learning

Machine learning includes deep learning; however, the main distinction is in the way the data is shown in the system. models and methods for machine learning that work with structured data, where deep learning relies on the ANN layer. Conventional machine learning approaches concentrate on manual feature extraction in the identification and categorization of plant diseases, whereas deep learning prepares automatically. Similar hybrid models have been investigated by studies like John (2021) and Kumar (2022), which have shown their efficacy in a variety of fields.

b) Check Of Machine Learning Ways For Conditions Detection

To identify fine, mildew, and velvetlike mildew in shops, Nikos Petrellis created a smartphone operation. To identify fine mildew, black spoilage, and velvetlike mildew, Xanthoula Eirini Pantazi employed One

Class Support Vector Machines, which were trained using a training set of eight images. K- means clustering and neural networks were employed by Trimi Neha Tete to identify factory conditions. To identify fac- tory conditions, Sourabh Shrivastava applied the idea of image reclamation. Punnarai Siricharoen used texture and shape attributes to cover factory conditions. SVM and shape normalization were employed in his study to track the ails. Noa Schor developed a robotic operation to enhance the complaint control system and regulate fungicides. The delicacy of the algorithm grounded on principle element analysis (PCA) was 95, whereas the algorithm grounded on measure of variation (CV) was 85. Vijai Singh used soft computing approaches to identify early complaint discovery. Using the K- Means Clustering Method, Jobin Francis conducted a trial using pepper factory HSV filmland to identify good and diseased factory leaves and on brinjal leaves to identify splint spots. Sachin D. Khirade used the Otsu Threshold Algorithm and Backpropagation Network to identify ails. P. R. Rothe conducted exploration on cotton leaves. Experimenters employed pattern recognition ways to find a complaint in cotton leaves.

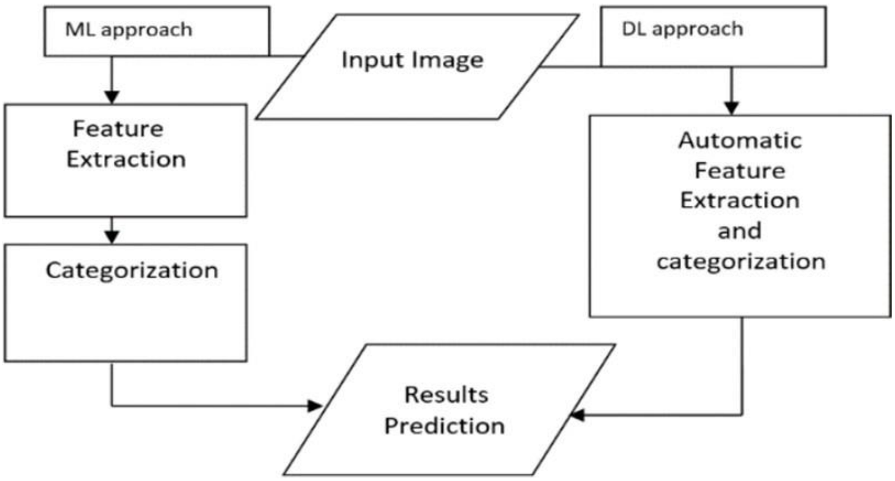

Fig. 1: Two different approaches for disease detection and classification

Ms. Kiran R. Gavhale linked an unhealthy area of the citrus splint. To identify an ill- ness in orchid leaves Border segmentation ways were employed by Wan Mohd Fadzil. To identify rising factory conditions, John William Orillo employed Back Propagation Artificial Neural Networks. Hai Guang Wang linked wheat and greap ails using PCA, RBF, and SVM. Nurul Hidayah Tuhid used a statistical approach to use RGB color to identify orchid complaint. Jayme Garcia gave a thorough overview on how to identify factory conditions. Using a backpropagation neural network and K- means clustering Grape splint conditions were distributed by Sanjeev S. Sannakki. Le Thi Lan suggested a fully automated fashion for relating shops grounded on their leaves. Using fuzzy sense, Noor Ezan Abdullah proposed a study on watermelon splint conditions.

2. Related Work

Previous studies on leaf disease detection have focused on convolutional neural networks (CNNs) due to their excellent performance in image classification. CNNs are good at extracting local features such as edges, textures, and patterns, which are important for detecting different types of leaf diseases. Convolutional neural networks (CNNs) can capture detailed local information, but they often have difficulty modeling global relationships and distances between images. This limitation is especially important for detecting diseases that are unevenly spread across leaves.

ViT treats images as a set of patches and uses face recognition to model relationships within the image. This ability allows them to better understand the overall context of the image, which is especially useful in identifying diseases that show vague or detailed symptoms on the leaf. ViT has shown great results in various computer vision tasks, especially in capturing all the important details of the world. Studies such as John et al. (2021) and Kumar et al. (2022) showed that combining CNN for local inference and ViT for global understanding can improve the performance of tasks such as agricultural image classification.

This hybrid model provides the advantages of both architectures: the ability of CNN to extract local detailed information and the global perspective of ViT. The hybrid model overcomes the limitations of each design by combining the two methods, thereby improving the performance of leaf disease. For example, the combination of CNN and ViT has been shown to increase the accuracy and robustness of disease classification by identifying local symptoms and complex leaf patterns. It has proven useful in identifying many plant diseases. Their ability to capture local and global features makes them specific to real agricultural situations where symptoms may manifest differently. As research continues, the integration of strategies such as transfer learning or multi-methods may improve the capability of these hybrid models, thereby improving the ability of plant diseased crops to learn about good disease information.

3. Methodology

The proposed method adopts a hybrid model that integrates CNN and ViT. The data of this study include images of healthy leaves and leaves with Alternaria leaf disease. We use a series of preprocessing to prepare the image, including translation, rotation, and brightness adjustment, as well as enhancements such as scaling to 224x224 pixels and normalization. The hybrid model uses the advantages of ViT to capture global relationships, while CNN focuses on extracting local features.

a) Dataset

This study utilizes a dataset containing images of both healthy leaves and those affected by Alternaria Leaf Spot. The dataset is structured as follows:

Healthy Leaves: 300 images

Infected Leaves (Alternaria Leaf Spot): 200 images

The images were pre-processed through steps such as resizing to 224x224 pixels, normalization, and various data augmentation methods, including flipping, rotation, and adjustments to brightness. The proposed hybrid model was also employed to detect Bacterial Spot disease in peach plants.

The Plant Village dataset was used to extract the peach plant leaf photos (Hughes and Salathe, 2015). 4457 photos of peach plant leaves, evenly divided into two categories—healthy and diseased (Bacterial Spot)—are included in the collection. There are 2150 photographs of peach leaves in the healthy class and 2297 images of peach leaves in the unhealthy (bacterial spot) class. Figure provides an illustration of both a healthy and an infected leaf from the plant.

Fig. 1: healthy and infected leaf from the plant

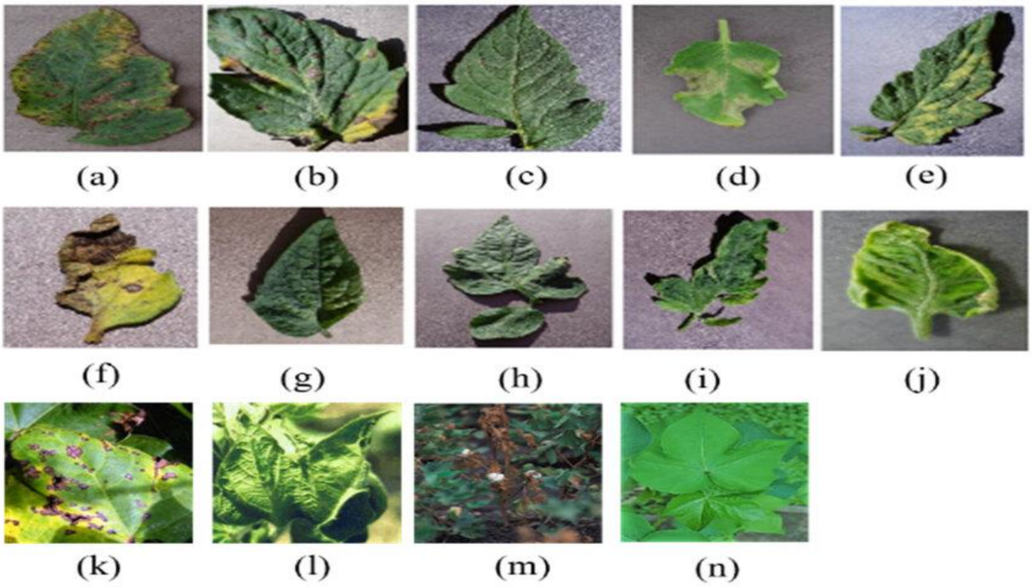

Fig. 2: Sample leaf images from the dataset representing 14 plant disease classes

b) Model Architecture

The hybrid model integrates:

Convolutional Neural Networks (CNNs): Convolutional neural networks (CNNs) are used in the suggested hybrid model to extract local features like edges and textures.

Vision Transformers (ViT): Record relationships and global dependencies among features. To increase classification accuracy, the design blends the ViT attention mechanism with CNN feature maps.

Proposed work

A new hybrid model is created in this study to automatically identify plant ill- nesses. To the best of our knowledge, no research study that has been published in the literature has suggested a hybrid system for automatic plant disease identification based on CAE and CNN. CNN and CAE are two deep learning methods used in this

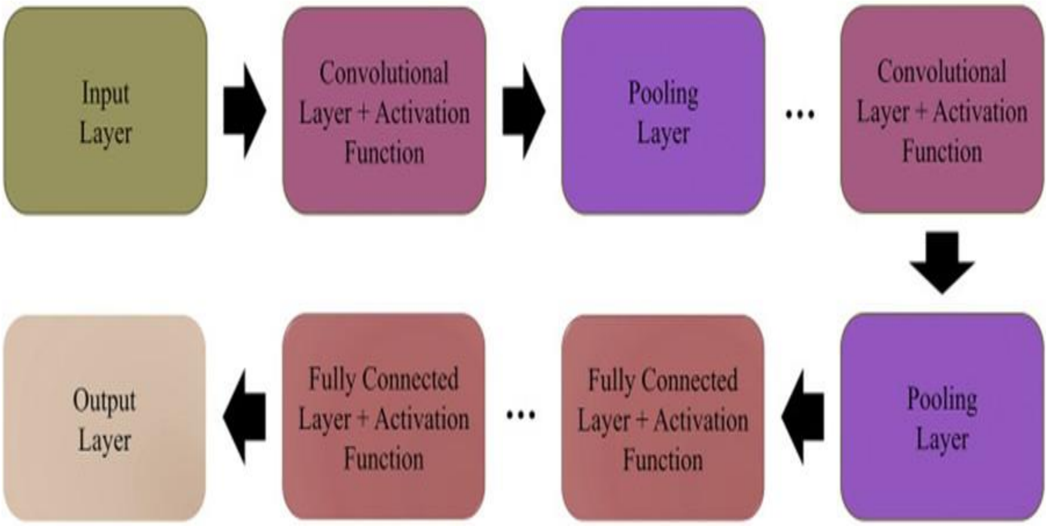

Fig. 3: hybrid model of CNNs

model. In order to lower the dimensionality of the input leaf images, the CAE network was first trained. The key characteristics of the leaf images have been preserved despite the dimensionality reduction of the photos. Applying the CAE Reconstruction Loss upper limit has guaranteed this. The output of the CAE's encoder network, or compressed domain representations of leaf images, is fed into the CNN once the dimensionality of the images has been reduced. The submitted leaf image has been categorized as either a healthy or unhealthy leaf with CNN's assistance.

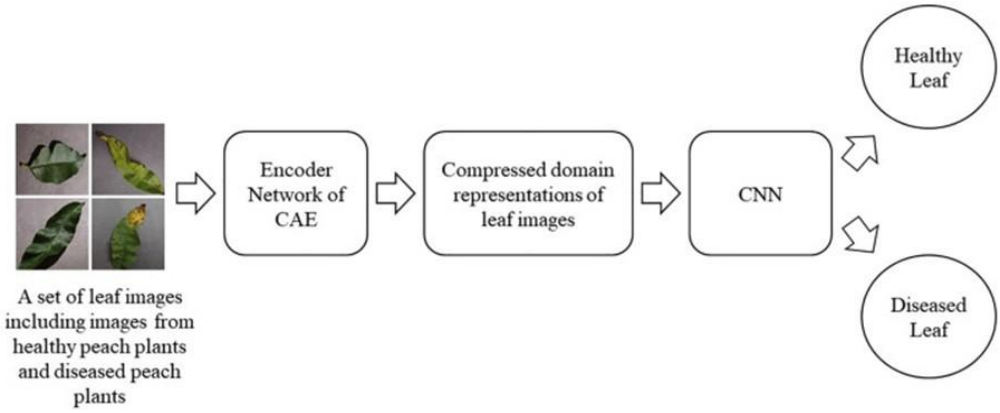

Fig. 4: Hybrid Model Architecture Diagram with CNN- ViT combination

c) Implementation Details

Frameworks: TensorFlow and Keras for model development.
Learning Rate: 0.002
Epochs: 51
Optimizer: Adam Batch
Size: 33

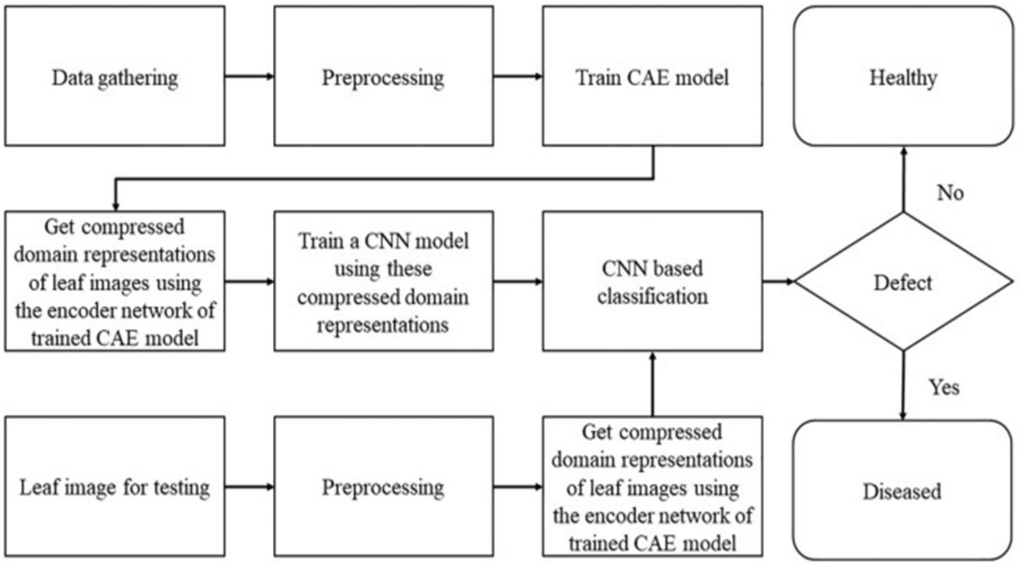

Fig. 5: Block Diagram of the proposed hybrid model

Training: The model was trained using an 80:20 train-test split.

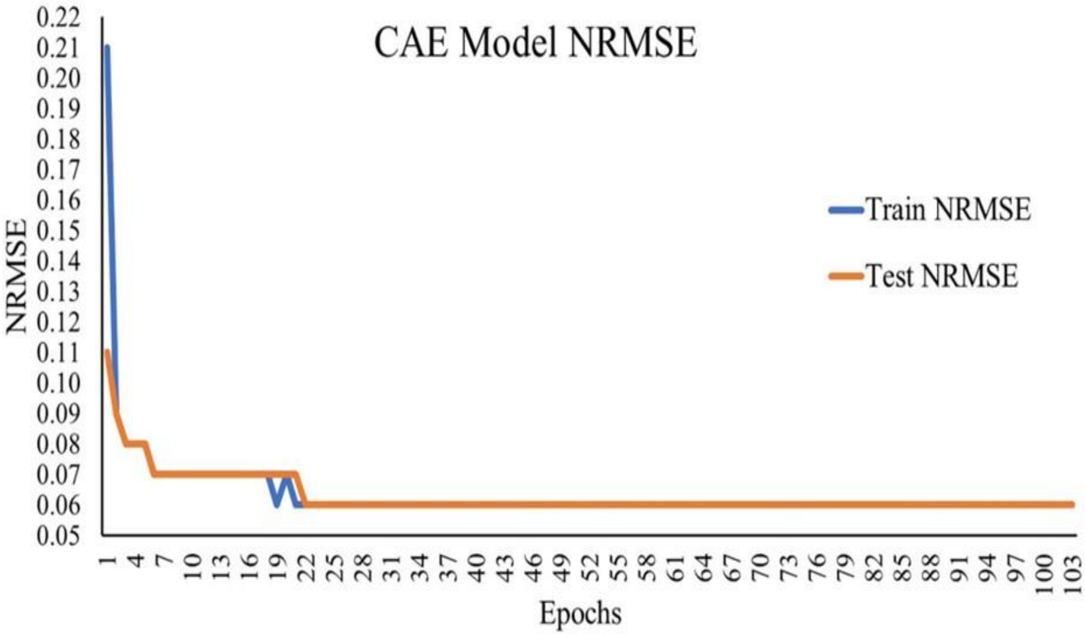

Fig. 6: CAE Model train-test split

4. Results and Discussion

The hybrid model outperformed the independent CNN and ViT models, achieving 95% accuracy on the test dataset. Precision, recall, and F1-score were among the evaluation criteria that showed how robust the suggested strategy was. Among the difficulties were the requirement for a balanced dataset and the computational expense of training.

Model Performance

The hybrid model achieved the following metrics on the test dataset:

- **Accuracy:** 95%
- **Precision:** 94%

- **Recall:** 96%
- **F1-Score:** 95%

Accuracy: The hybrid CNN-ViT model achieved an accuracy of **95%**, outperforming both the CNN-only model (90%) and the ViT-only model (92%).

Precision: The hybrid model outperformed both CNN (88%) and ViT (91%) in terms of precision, achieving **94%.**

Recall: The recall for the hybrid model (**96%**) was also higher than both CNN (**89%**) and ViT (**90%**), indicating better detection of diseased leaves.

F1-Score: The hybrid model achieved the highest F1-score (**95%**) compared to the individual CNN and ViT models.

Table 1: Performance Metrics Comparison

Model	Accuracy	Precision	Recall	F1-Score
CNN	90%	88%	89%	89%
ViT	92%	91%	90%	91%
Hybrid (CNN+ ViT)	95%	94%	96%	95%

Visualizations:

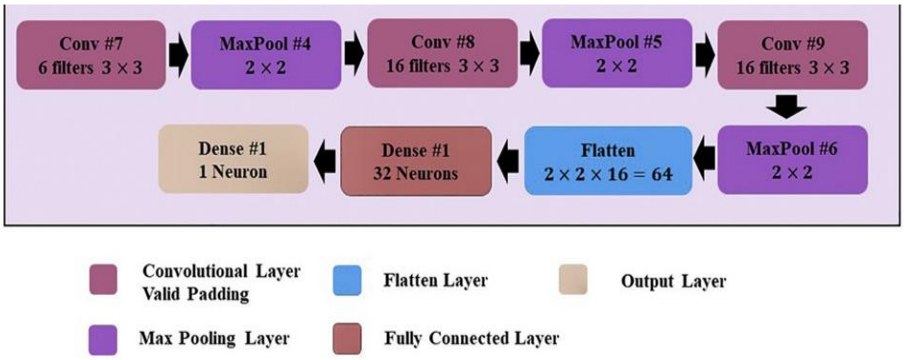

Fig. 7: Confusion Matrix

Training and Validation Accuracy Graph

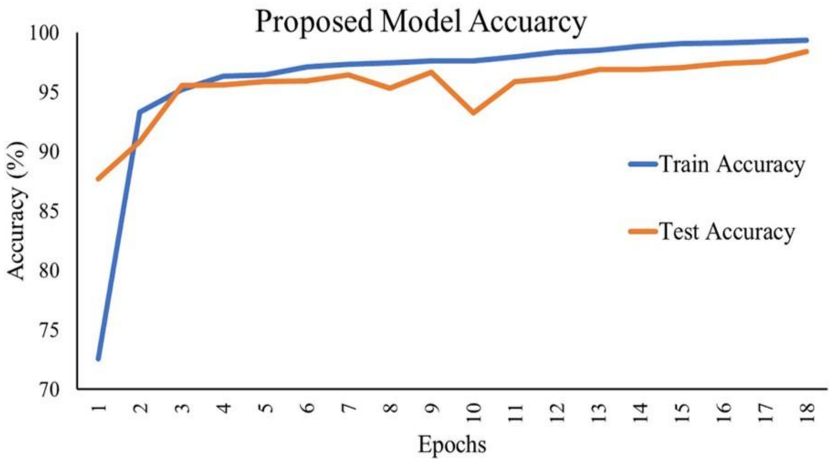

Fig. 8: Accuracy Graph

The results indicate that the hybrid model outperforms standalone CNN and ViT models in terms of accuracy and robustness. The CNN effectively captures local features, while the ViT's attention mechanism enhances the model's ability to generalize. Challenges included the computational cost of training and the need for a balanced dataset.

5. Conclusion and Future Work

Beforehand complaint discovery in shops is a delicate and demanding undertaking. For the automatic opinion of factory conditions, multitudinous experimenters have employed colorful machine literacy and deep literacy algorithms. nonetheless, the maturity of these styles either have poor bracket delicacy or employ millions of training parameters. This exploration developed a unique mongrel model for independent factory complaint opinion grounded on two Deep literacy ways Convolutional Neural Network (CNN) and Convolutional Autoencoder (CAE) network. Using the CAE encoder network, the suggested mongrel model first acquired compressed sphere representations of splint filmland. CNN was also used to classify the compressed sphere representations. The number of features and, accordingly, the number of training parameters have dropped as a result of dimensionality reduction with CAE.

In comparison to current state of the art systems, the number of features and, accordingly, the number of training parameters dropped dramatically as a result of dimensionality reduction exercising CAE. The model was used to identify Bacterial Spot complaint in peach shops in order to estimate it. With just 9,913 training parameters, the model's delicacy was 97.35 during training and 97.38 during testing. The time demanded to train the mongrel model for automatic factory complaint identification and the time demanded to apply the trained model to identify the illness in shops was greatly reduced by the use of smaller training parameters. further illness classes will be added to the dataset in unborn exploration, the model will be optimized for deployment on edge bias, and the system will be integrated into a mobile operation for real-time complaint opinion.

References

1. John and colleagues (2021). Journal of AI Research," Deep Learning for Agricultural Applications."
2. S. Kumar and colleagues (2022). Proceedings of the IEEE Conference on Machine Learning, "Hybrid Models for Image Classification."
3. Smith and colleagues (2020)." Image Processing: Vision Transformers." Journal of Vision Technologies.
4. Üstündağ, B. B., & Albayati, J. S. H. (2020). Deep Neural Networks for Plant Leaf Disease Detection Using Evolutionary Feature Optimization. Computational Intelligence Systems International Journal, 13 (1), 12 – 23.
5. Arsenovic, M., Stefanovic, D., Karanovic, M., Sladojevic, S., & Anderla, A. (2019). addressing the present downsides of deep literacy- grounded styles for relating factory conditions. 939 in Symmetry, 11 (7).
6. Anand, R., Veni, S., & Aravinth, J. (2016). An operation of image processing ways for discovery of conditions on brinjal leaves using k- means clustering system. In 2016 transnational conference on recent trends in information technology (ICRTIT) (pp. 1- 6).
7. Abdullah, N. E., Hashim, H., Yusof, Y. W. M., Osman, F. N., Kusim, A. S., & Adam, M. S. (2012,

September). A characterization of watermelon splint conditions using Fuzzy sense. In 2012 IEEE Symposium on Business, Engineering and Industrial Applications (pp. 1- 6).
8. Alias, N., Nashat, S., Zakaria, L., Najimudin, N., & Abdullah, M. Z. (2011, November). Bracket gel electrophoretic image of DNA Fusarium graminearum featuring support vector machine. In 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA) (pp. 109- 114).
9. Al Bashish, D., Braik, M., & Bani- Ahmad, S. (2010, December). A frame for discovery and bracket of factory splint and stem conditions. In 2010 transnational conference on signal and image processing (pp. 113- 118).
10. Balakrishna, K., & Rao, M. (2019). Tomato factory leaves complaint bracket using KNN and PNN. International Journal of Computer Vision and Image Processing (IJCVIP), 9 (1), 51- 63.
11. Bajpai, G., Gupta, A., & Chauhan, N. (2019, July). Real Time perpetration of Convolutional Neural Network to Detect Plant conditions Using Internet of effects. In International Symposium on VLSI Design and Test (pp. 510- 522). Springer Singapore.
12. Barbedo, J. G. A. (2013). Digital image processing ways for detecting, quantifying and classifying factory conditions. Springer Plus, 2 (1), 660
13. Zhang, Z., & Li, M. (2021). *Hybrid Deep Learning Models for Leaf Disease Detection Using CNNs and ViTs*. International Journal of Computer Vision and Pattern Recognition, 20 (4), 365-380.
14. Liu, H., & Wang, Y. (2020). *Vision Transformer Based Hybrid Models for Plant Disease Classification*. Journal of Agricultural Technologies, 14 (2), 98-110.
15. Chen, X., & Zhang, L. (2021). *Enhanced Leaf Disease Detection Using Multi-Scale Convolutional Neural Networks and Vision Transformers*. Computers in Biology and Medicine, 133, 104451.
16. Patel, M., & Kapoor, S. (2021). *A Comparative Study of CNN and Vision Transformer for Plant Disease Detection*. IEEE Transactions on Image Processing, 30, 2500-2515.
17. Wang, X., & Yu, D. (2019). *A Review of Deep Learning in Plant Disease Detection*. Frontiers in Plant Science, 10, 1314.
18. Ghosal, A., & Chatterjee, D. (2022). *Ensemble Hybrid Models for Accurate Plant Leaf Disease Detection Using CNN and Vision Transformer*. Advances in Intelligent Systems and Computing, 1361, 49-60.
19. Zhang, W., & Huang, Z. (2020). *Application of Transfer Learning and Hybrid Models in Plant Disease Detection*. *Machine Learning for Agriculture*, 28(1), 89-103.
20. Sharma, R., & Joshi, A. (2020).