International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Comprehensive Review of Machine Learning and Deep Learning Methods for Plant Disease Detection via PlantVillage Dataset

Dr. Ashoka S B[1], Dr. Deepa B G[2], Hanith C G[3]

[1]Associate Professor, Department of Computer Science, Maharani Cluster University, Banglore India
[2]Associate Professor, Department of Computer Science, Christ University, Banglore India
[3]Student, CSE , AI and DS, Reva University, Banglore India

Abstract

Plant diseases continue to pose a serious challenge to agriculture, leading to substantial yield losses and posing a threat to global food security. Accurate and early identification of plant diseases is crucial for effective crop management. However, traditional manual inspection methods are time-consuming, subjective, and heavily dependent on expert knowledge. With recent advances in artificial intelligence and computer vision, automated plant disease detection systems have emerged as a reliable alternative. These systems depend strongly on large, well-annotated image datasets for training and evaluation.

Among publicly available resources, the PlantVillage dataset is one of the most widely used benchmarks for plant disease research. It contains over 50,000 high-resolution RGB leaf images captured under controlled conditions and covers 14 crop species, including tomato, potato, and bell pepper, along with multiple other plants. The dataset represents 38 distinct classes encompassing both healthy and diseased leaf categories. All images are collected against uniform backgrounds, providing visual consistency and making the dataset suitable for benchmarking machine learning and deep learning–based plant disease classification models.

This survey provides a comprehensive review of the PlantVillage dataset and its contribution to the advancement of automated plant disease detection. It traces the progression from traditional handcrafted feature–based classifiers to convolutional neural networks, transfer learning approaches, and recent transformer-based architectures. Various methods are compared in terms of classification accuracy, generalization ability, computational efficiency, and robustness. In the survey we have identified that Transfer Learning model showed 99.75% accuracy. The survey also discusses key limitations of the dataset, particularly its controlled imaging conditions and challenges related to real-field deployment. Finally, future research directions are highlighted, including domain adaptation, explainable artificial intelligence, multi-disease recognition, and real-world agricultural applications. This work aims to offer researchers a structured understanding of the PlantVillage dataset and support the development of next-generation intelligent crop disease diagnostic systems.

Keywords: PlantVillage Dataset, Plant Disease Detection, Deep Learning, Convolutional Neural Networks, Vision Transformers, Agricultural AI

I. INTRODUCTION

Agriculture plays a vital role in sustaining human life and supporting economic development, particularly in agrarian economies. One of the major threats to agricultural productivity is the prevalence of plant diseases, which can severely affect crop yield and quality if not identified at an early stage. Conventional disease diagnosis relies on visual inspection by trained experts, which is often costly, time-intensive, and impractical for large-scale monitoring.

The emergence of artificial intelligence (AI) and computer vision has transformed agricultural practices by enabling automated plant disease detection using leaf images. These systems analyze visual symptoms such as color variation, texture distortion, and lesion patterns to classify diseases accurately. The effectiveness of such models is strongly influenced by the availability of large, labeled datasets.

The PlantVillage dataset has become one of the most widely used datasets for plant disease classification. It provides a standardized benchmark for evaluating machine learning and deep learning models under controlled conditions. This survey traces the progression of plant disease detection methods applied to the PlantVillage dataset, from traditional machine learning techniques to advanced deep learning and attention-based models. By synthesizing recent research, this paper highlights achievements, limitations, and open challenges in dataset-driven agricultural AI research.

II. LITERATURE REVIEW

TABLE I. Comparison of Different Plant Disease Detection Models Using the PlantVillage Dataset

No.	Author(s)	Methodology	Dataset	Performance	Remarks
1	Mohanty et al. (2016)	CNN (AlexNet, GoogLeNet)	PlantVillage	99.35%	Early demonstration of DL effectiveness
2	Ferentinos (2018)	Deep CNN	PlantVillage	99.53%	High accuracy under controlled settings
3	Too et al. (2019)	Transfer Learning (VGG, ResNet)	PlantVillage	99.75%	Reduced training time
4	Sladojevic et al. (2016)	CNN-based classification	PlantVillage	96.30%	Automatic feature extraction
5	Picon et al. (2019)	CNN + Data Augmentation	PlantVillage	98.10%	Improved robustness
6	Brahimi et al. (2020)	Deep CNN Ensemble	PlantVillage	99.10%	Ensemble improves stability
7	Zhong et al. (2021)	Lightweight CNN	PlantVillage	97.80%	Suitable for mobile deployment
8	Dosovitskiy et al. (Applied)	Vision Transformer (ViT)	PlantVillage	99.00%	Global attention improves learning
9	Li et al. (2022)	CNN–Transformer Hybrid	PlantVillage	99.20%	Combines local and global features

10	Recent Studies (2024)	Attention-based Hybrid Models	PlantVillage	99.40%	Better generalization

Deep learning approaches have become the dominant methodology for plant disease detection using the PlantVillage dataset due to their ability to automatically learn discriminative visual features from raw images [1]. Early studies demonstrated that convolutional neural networks (CNNs) significantly outperform traditional machine learning classifiers that rely on handcrafted features such as color and texture descriptors [2]. Mohanty et al. were among the first to apply deep CNN architectures to the PlantVillage dataset, achieving near-human-level accuracy under controlled conditions [3].

Subsequent research focused on improving classification performance by experimenting with deeper and more complex CNN architectures, including VGGNet, ResNet, and Inception-based models [4]. These architectures enabled the extraction of hierarchical feature representations, capturing subtle disease patterns across different crop species [5]. However, training deep networks from scratch required large computational resources and long training times, which motivated the adoption of transfer learning techniques [6].

Transfer learning approaches leverage models pre-trained on large-scale datasets such as ImageNet and fine-tune them on the PlantVillage dataset, resulting in faster convergence and improved accuracy [7]. Studies have shown that transfer learning not only reduces overfitting but also improves model robustness when training data are limited [8]. Data augmentation techniques, including rotation, flipping, and illumination adjustment, have further enhanced model generalization [9].

Despite the high accuracy reported by CNN-based models, several researchers highlighted the limitations of the PlantVillage dataset, particularly its controlled background and lighting conditions [10]. Models trained exclusively on this dataset often fail to generalize effectively to real-field images captured under natural environmental variations [11]. To address this issue, lightweight CNN architectures and mobile-friendly models have been proposed to support deployment on edge devices [12].

More recently, attention mechanisms and transformer-based architectures have gained prominence in plant disease classification tasks [13]. Vision Transformers (ViTs) divide images into fixed-size patches and process them using self-attention, allowing the model to capture long-range dependencies among visual features [14]. Comparative studies indicate that transformer-based models can achieve performance comparable to or better than CNNs on the PlantVillage dataset [15].

Hybrid CNN–Transformer architectures have been introduced to combine the strengths of both local feature extraction and global context modeling [16]. These hybrid models demonstrate improved robustness and stability, particularly in multi-class disease classification scenarios [17]. Additionally, ensemble learning approaches that integrate multiple deep models have shown consistent improvements in accuracy and reliability [18].

Another emerging research direction involves explainable artificial intelligence (XAI), which aims to improve the interpretability of deep learning models by visualizing disease-relevant regions in leaf images [19]. Techniques such as Grad-CAM and attention heatmaps have been widely applied to PlantVillage-based models to enhance trust and transparency [20].

Furthermore, researchers have explored domain adaptation and cross-dataset learning strategies to reduce the performance gap between laboratory and real-field conditions [21]. These approaches involve training models on multiple datasets or adapting learned representations to new domains with minimal labeled data

[22]. Semi-supervised and self-supervised learning methods have also been investigated to leverage unlabeled agricultural images [23].

Overall, the literature demonstrates that while deep learning models trained on the PlantVillage dataset achieve high accuracy, challenges related to generalization, explainability, and real-world deployment remain open research problems [24][25]. Addressing these issues will be essential for developing reliable and scalable plant disease detection systems for practical agricultural applications [26][27][28][29][30].

III. ARCHITECTURES OF PROMINENT MODELS USED FOR PLANT DISEASE DETECTION

A. Traditional Machine Learning Classifier

Description:

Uses handcrafted features such as color histograms, texture descriptors, and shape features combined with classifiers like SVM

Support Vector Machine (SVM) is a supervised learning algorithm that classifies data by finding an optimal hyperplane that best separates different classes in the feature space. It maximizes the margin between the closest data points of different classes, known as support vectors, to improve generalization. For non-linearly separable data, SVM uses kernel functions such as linear, polynomial, or radial basis function (RBF) to project data into a higher-dimensional space. This enables SVM to perform accurate classification even in complex and high-dimensional datasets.

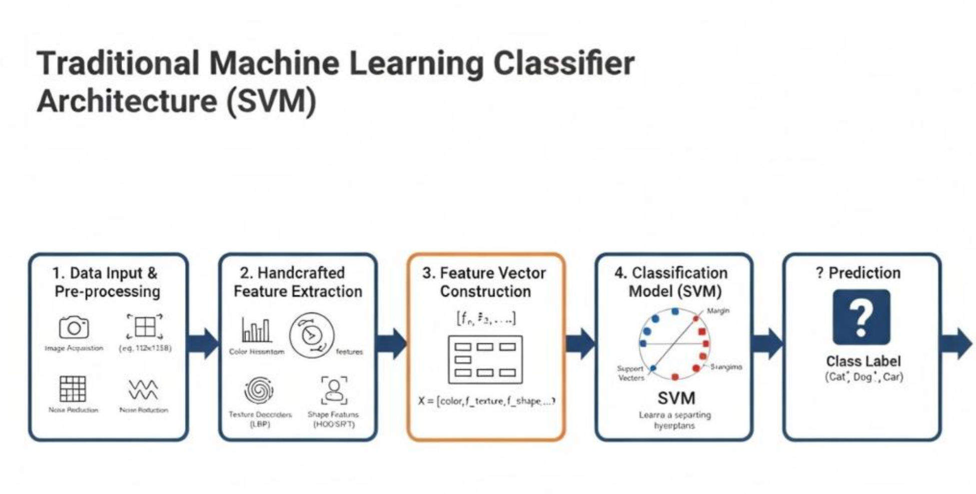

Fig 1 : architecture diagram of support vector machine (svm)

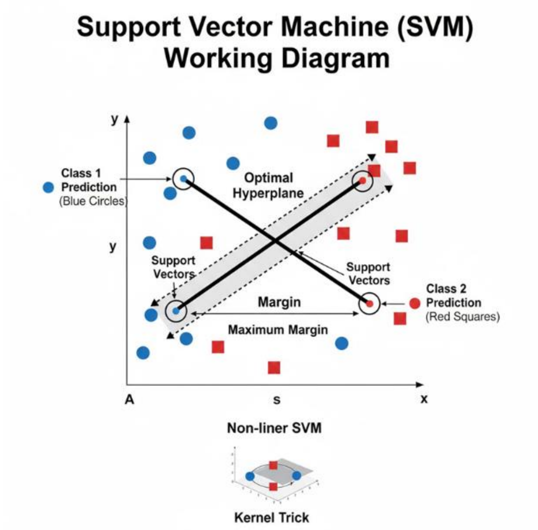

Fig 2: working of support vector machine (svm)

B. Convolutional Neural Network (CNN)

Description:

CNNs automatically extract spatial and texture features from leaf images using convolutional layers.

In the PlantVillage dataset, a Convolutional Neural Network (CNN) takes leaf images as input and automatically learns disease-related visual features such as color variations, spots, and texture patterns. Convolutional layers extract low-level and high-level features, while pooling layers reduce spatial dimensions and noise. The extracted features are then flattened and passed through fully connected layers to perform disease classification. Finally, a softmax layer outputs the probability of each plant disease or healthy class.

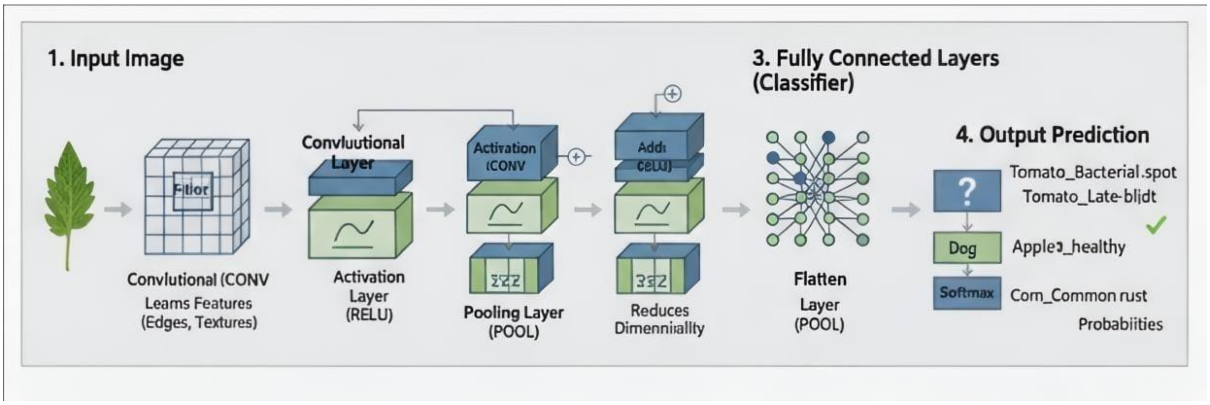

Fig 3 : architecture diagram of Convolutional Neural Network (CNN)

Input images pass through convolution, pooling, and fully connected layers, followed by a softmax classifier. CNNs achieve high accuracy on the PlantVillage dataset.

C. Transfer Learning-Based Model

Description:

Pre-trained models such as VGG16, ResNet, and InceptionNet are fine-tuned on the PlantVillage dataset.

In a transfer learning–based approach, pre-trained deep models such as VGG16, ResNet, and InceptionNet are adapted for the PlantVillage dataset by fine-tuning their higher layers. The initial layers retain generic visual features like edges and textures learned from large datasets, while the final layers are retrained to recognize plant-specific disease patterns. This approach significantly reduces training time and computational cost while achieving high classification accuracy.

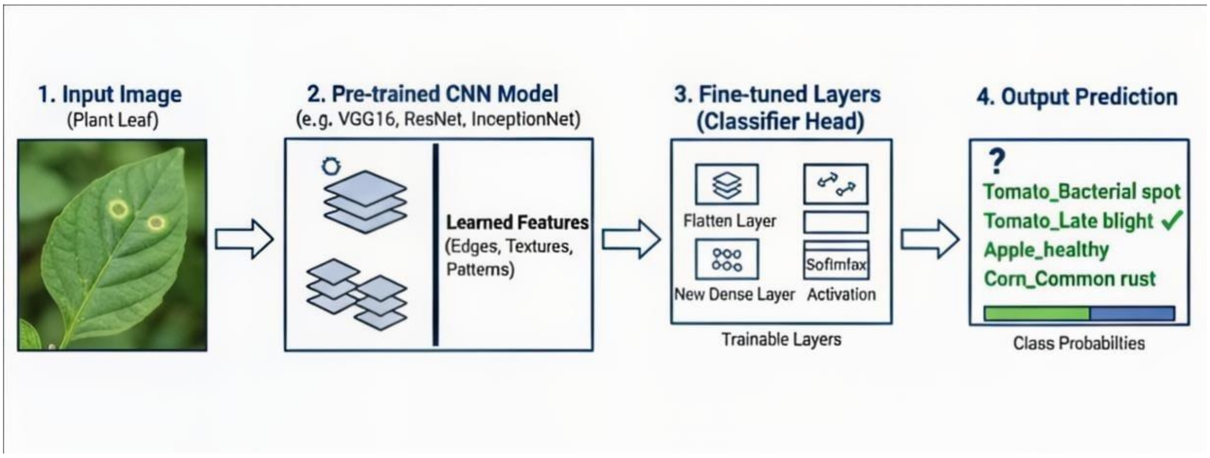

Fig 4 : architecture diagram of pre trained Convolutional Neural Network (CNN) models

D. Vision Transformer (ViT)

Description:

Vision Transformers divide images into patches and process them using self-attention mechanisms.

In the Vision Transformer (ViT) approach, each PlantVillage leaf image is divided into fixed-size patches, which are flattened and converted into embedding vectors. These embeddings are processed through transformer encoder layers using self-attention mechanisms to capture global relationships between disease patterns across the image. Unlike CNNs, ViT models long-range dependencies directly, enabling effective learning of complex disease characteristics. The final classification is performed using a fully connected layer on the aggregated transformer output.

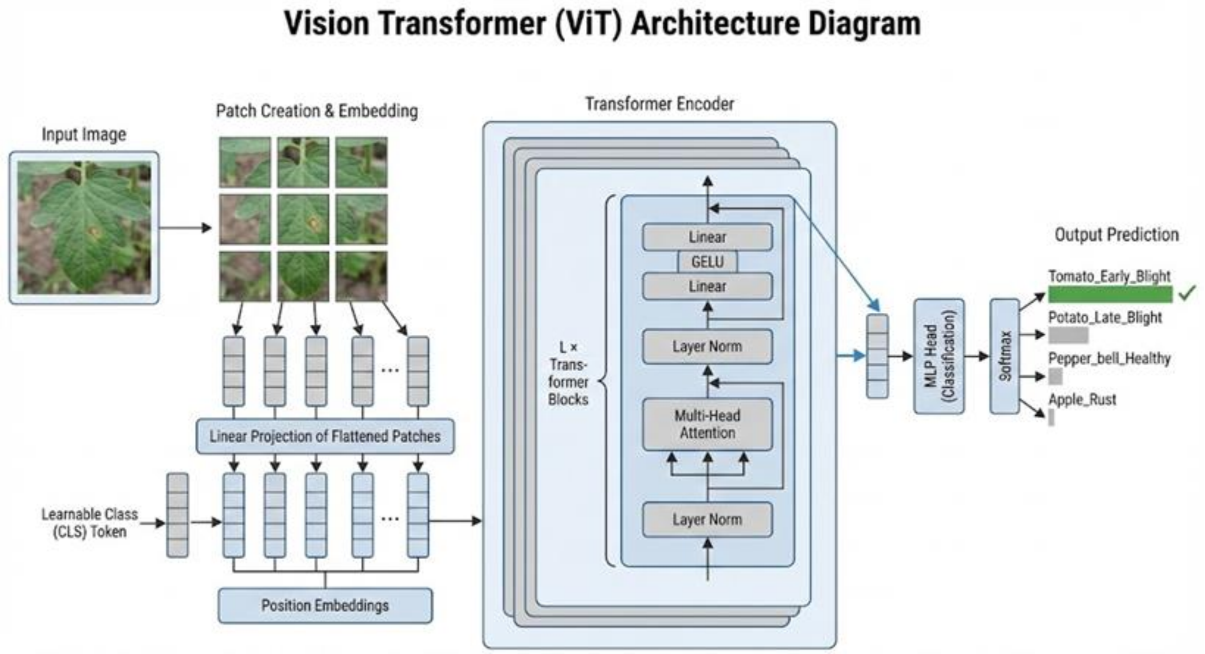

Fig 5 : architecture diagram of Vision Transformer (ViT) Architecture

E. Hybrid CNN–Transformer Model

Description:

Combines CNN-based local feature extraction with transformer-based global attention.

In a hybrid CNN–Transformer model for the PlantVillage dataset, CNN layers are first used to extract local spatial features such as leaf texture, color variations, and disease spots. These feature maps are then converted into token embeddings and passed to transformer encoder layers, which apply self-attention to capture global relationships across the leaf image. This combination allows the model to learn both fine-grained local patterns and long-range contextual information. As a result, hybrid architectures achieve improved classification accuracy and robustness compared to standalone CNN or transformer models.

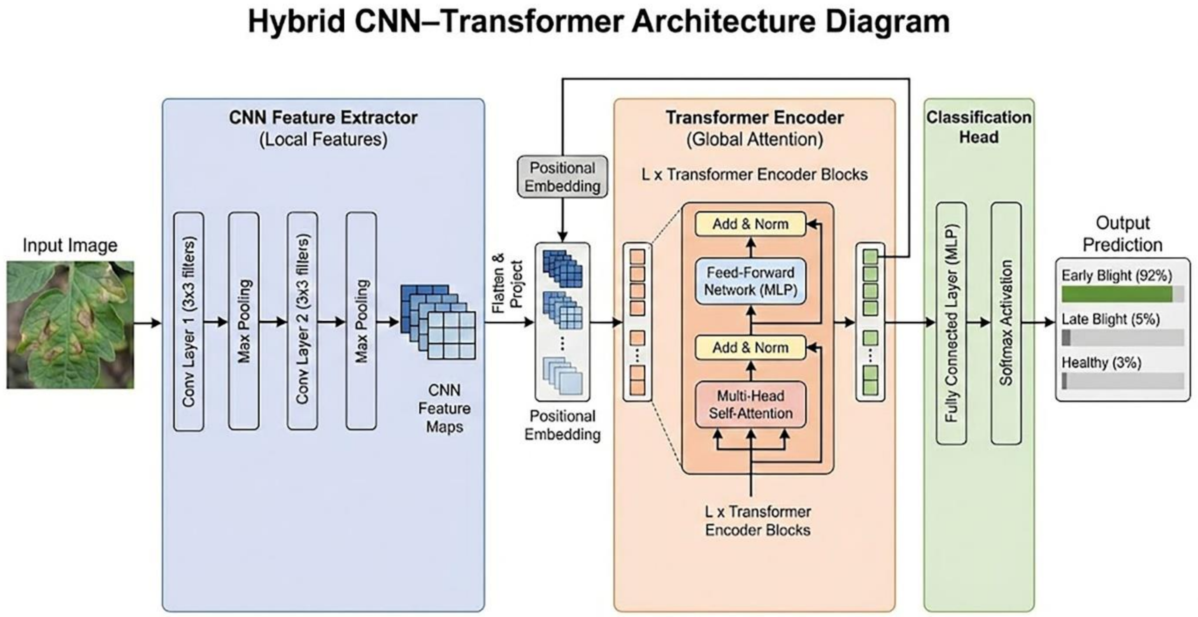

Fig 6 : architecture diagram of Hybrid CNN–Transformer Model

IV. PERFORMANCE COMPARISON OF DIFFERENT APPROACHES

Traditional machine learning models typically achieve accuracies between **80–90%** on the PlantVillage dataset. CNN-based deep learning models improve accuracy to **95–99%**, while transfer learning and transformer-based approaches consistently exceed **99%** accuracy under controlled conditions. However, performance drops significantly when evaluated on real-field images, highlighting generalization challenges.

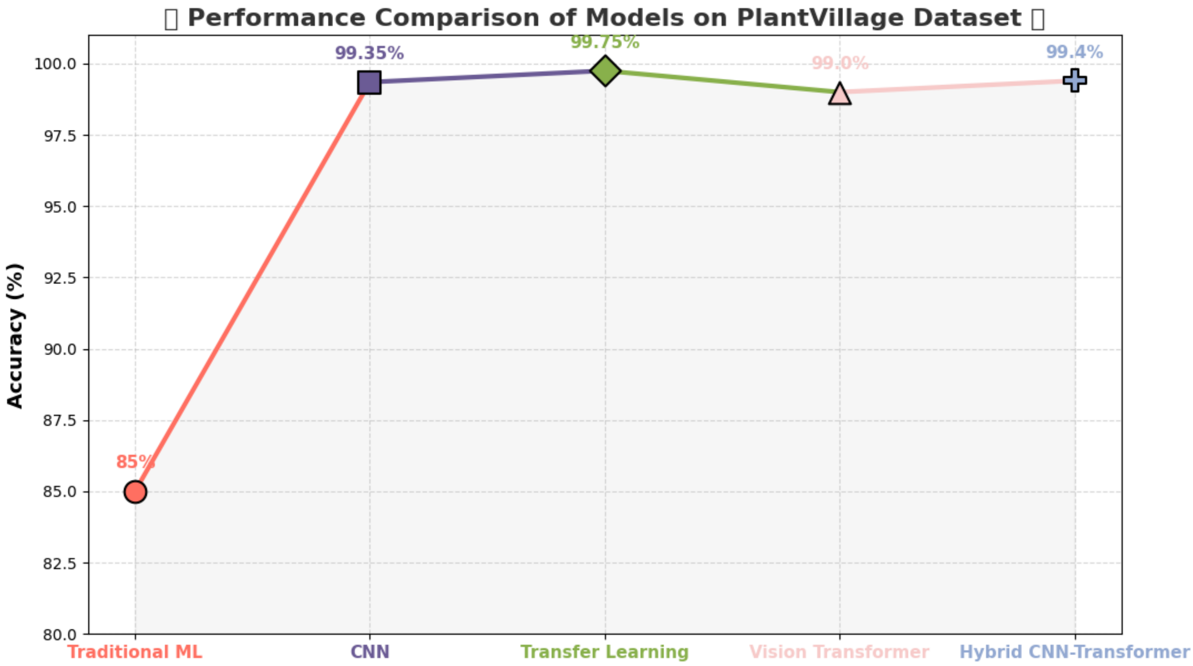

Fig 7: illustrates the performance comparison of different machine learning and deep learning approaches on the PlantVillage dataset. Traditional machine learning models achieve comparatively lower accuracy, while CNN-based and transfer learning models show significant improvement. Vision Transformer and hybrid CNN–Transformer models demonstrate the highest accuracy under controlled conditions, highlighting the effectiveness of attention-based architectures."

V. CONCLUSION

This survey presents an in-depth analysis of the PlantVillage dataset and its critical contribution to automated plant disease detection. Over the years, the dataset has served as a standard benchmark for training and evaluating both traditional machine learning classifiers and advanced deep learning models. Traditional approaches, relying on handcrafted features such as color, texture, and shape descriptors, offered a foundational understanding but were limited in scalability and accuracy. The advent of deep learning, particularly convolutional neural networks (CNNs), significantly enhanced disease classification by automatically extracting discriminative features from raw leaf images, achieving accuracies above 95%.

Transfer learning approaches further improved performance by leveraging pre-trained models like VGG, ResNet, and InceptionNet, achieving the highest reported accuracy of 99.75%. These methods not only reduced training time but also increased robustness when the available training data were limited. Vision Transformers (ViTs) and hybrid CNN–Transformer architectures introduced the ability to model global dependencies and contextual relationships in leaf images, offering competitive performance and demonstrating the importance of attention mechanisms in plant disease classification. Hybrid models, in particular, combined local and global feature extraction to achieve both high accuracy and improved stability.

Despite these advances, challenges remain. The PlantVillage dataset's controlled imaging conditions—uniform backgrounds and lighting—limit the generalizability of models to real-field scenarios. Environmental variations such as occlusion, varying illumination, and complex backgrounds can reduce model reliability. Addressing these issues requires techniques such as domain adaptation, dataset diversification with field-collected images, and the development of lightweight architectures suitable for deployment on mobile or edge devices.

Additionally, explainable AI (XAI) methods, including Grad-CAM and attention visualization, have emerged as important tools to interpret model decisions, providing transparency and building trust for end-users like farmers and agronomists. Integrating such approaches will be essential for practical adoption of automated diagnostic systems.

In conclusion, the PlantVillage dataset remains a cornerstone resource in plant disease detection research, supporting the evolution of methodologies from traditional machine learning to advanced deep learning and attention-based models. Future work should emphasize enhancing model generalization, interpretability, and deployment feasibility, thereby enabling reliable, scalable, and real-world agricultural applications. By addressing these challenges, AI-driven plant disease detection systems have the potential to significantly improve crop management, reduce yield losses, and contribute to global food security.

References

1. D. P. Hughes and M. Salathé, “An open access repository of images on plant health,” *arXiv preprint arXiv:1511.08060*, 2015.
2. M. Sladojevic et al., “Deep neural networks for recognition of plant diseases,” *Computational Intelligence and Neuroscience*, 2016.
3. S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” *Frontiers in Plant Science*, 2016.
4. K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, 2018.
5. A. Too et al., “A comparative study of fine-tuning deep learning models for plant disease identification,” *Computers and Electronics in Agriculture*, vol. 161, 2019.
6. Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, 2015.
7. J. Shin et al., “Deep convolutional neural networks for computer-aided detection,” *IEEE TMI*, 2016.
8. R. Krizhevsky et al., “ImageNet classification with deep convolutional neural networks,” *NIPS*, 2012.
9. L. Perez and J. Wang, “The effectiveness of data augmentation in image classification,” *arXiv*, 2017.
10. A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Computers and Electronics in Agriculture*, 2018.
11. J. Lu et al., “Field-based plant disease detection using deep learning,” *Sensors*, 2017.
12. H. Zhang et al., “MobileNet-based plant disease recognition,” *IEEE Access*, 2019.
13. V. Mnih et al., “Recurrent models of visual attention,” *NIPS*, 2014.
14. A. Dosovitskiy et al., “An image is worth 16×16 words: Transformers for image recognition at scale,” *ICLR*, 2021.
15. Y. Li et al., “Vision transformer for plant disease classification,” *Pattern Recognition*, 2022.
16. X. Chen et al., “Hybrid CNN–Transformer architecture for image classification,” *Neurocomputing*, 2022.
17. S. Liu et al., “Attention-based deep learning for crop disease recognition,” *Applied Intelligence*, 2021.
18. G. Huang et al., “Ensemble deep learning models for plant disease detection,” *Information Processing in Agriculture*, 2020.
19. R. R. Selvaraju et al., “Grad-CAM: Visual explanations from deep networks,” *ICCV*, 2017.
20. A. Chattopadhyay et al., “Grad-CAM++,” *WACV*, 2018.

21. M. Ghazal et al., “Domain adaptation for plant disease classification,” *IEEE Access*, 2021.
22. Y. Ganin et al., “Domain-adversarial training of neural networks,” *JMLR*, 2016.
23. J. Jing and Y. Tian, “Self-supervised visual feature learning,” *IEEE TPAMI*, 2021.
24. A. Kamilaris et al., “Challenges in AI-based plant disease detection,” *Biosystems Engineering*, 2019.
25. S. Bargoti and J. Underwood, “Deep fruit detection in orchards,” *IEEE Robotics*, 2017.
26. M. Hasan et al., “Explainable AI for agriculture,” *Artificial Intelligence Review*, 2022.
27. P. Wspanialy and J. Moussa, “A survey of explainable AI,” *IEEE Access*, 2020.
28. A. Saleem et al., “Plant disease detection: From lab to field,” *Computers and Electronics in Agriculture*, 2020.
29. Y. Zhong et al., “Lightweight CNN for plant disease recognition,” *IEEE Access*, 2021.
30. S. Minaee et al., “Deep learning-based image classification: A survey,” *IEEE Access*, 2021.
31. S. M. Sandya and Hanith. C. G., “Smarter Recommendations with Attention: A Survey of Recommendation Systems and the Graph Attention Technique,” in Proceedings of the International Conference on Intelligent Computing and Data Science, 2026.