International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026. ISSN (Online): 2456-7361

An Automated Deep Learning Framework for Potato Leaf Disease Detection Using Transfer Learning

Navjot kaur[1], Dr. Himanshu Aggarwal[2], Dr. Priyanka Jarial[3]
[1]Computer Science & Engineering, Punjabi University Patiala, Patiala, Punjab, India-147001
[2]Computer Science & Engineering, Punjabi University Patiala, Patiala, Punjab, India-147001
[3]Computer Science & Engineering, Punjabi University Patiala, Patiala, Punjab, India-147001

Abstract— *Timely and correct diagnosis of potato leaf diseases is key to increasing productivity and providing the foundation for precision agriculture. While deep learning has made great strides in automated plant disease detection, differences in image formats for storage and the absence of uniform implementation frameworks can impact the reproducibility and fair assessment of classification models. To solve this problem, this research proposes a comprehensive transfer learning approach for automatic potato leaf disease classification and tests the proposed approach using several machine learning (ML) and deep learning (DL) techniques in the same experimental settings. The proposed framework was tested on the publicly available PlantVillage database which has 1500 balanced potato leaf images from three classes Healthy, Early Blight, and Late Blight. Before training, a standard processing pipeline was used, which included resizing, normalizing and randomizing the images. The data set was encoded to JPEG, BMP and TIFF to study the effect of image encoding on classification results. The framework was adopted with five conventional machine learning algorithms (SVM, Random Forest, Decision Tree, K-Nearest Neighbors and Multi-Layer Perceptron) and five deep learning architectures (CNN, VGG16, ResNet50, MobileNetV2 and DenseNet121), all trained with the same preprocessing method, dataset partitioning, and evaluation procedure. The experimental evaluation shows that the proposed framework will allow for reliable and reproducible performance with different image formats and a fair comparison of different classification models. For the JPEG dataset, DenseNet121 had the highest accuracy of 98.67%, while MobileNetV2 and ResNet50 had the best results for the BMP and TIFF datasets, respectively. The results show that the learning architecture is not the sole factor affecting classification accuracy and that the choice of image representation makes a difference too, showing the need for standardisation in experimental validation. The proposed framework is a complete and repeatable implementation pipeline to categorize potato leaf disease and valuable insights to understand the performance of machine learning and transfer learning models in various image formats. The framework can be used as a guide for future research with intelligent crop monitoring and computer-aided agricultural disease diagnosis.*

Keywords— *PlantVillage Dataset, Computer Vision, Precision Agriculture, Automated Disease Diagnosis, Benchmark Framework, Potato Leaf Disease Classification, Transfer Learning, Deep Learning, Machine Learning.*

I. INTRODUCTION

Potato (Solanum tuberosum) is one of the most widely grown food crops globally and is very important to food security and agriculture economies. Potato crops are an important source of food worldwide because of their nutritional benefits and their ability to thrive in a wide range of climates. Potato crops, however, are very vulnerable to foliar diseases like Early Blight and Late Blight, that if not identified early and managed properly, can significantly affect crop quality and yield [15,16]. The potato leaf diseases are traditionally diagnosed by manual visual inspection by agricultural experts. Manual diagnosis is labor intensive, time consuming and subjective, and impractical for large agricultural fields, though experienced pathologists can diagnose disease symptoms reasonably accurately. In addition, disease symptoms are sometimes different in various environments, crop varieties, and stages of disease, leading to difficult diagnosis and treatment [17, 21]. The performance of automated plant disease diagnosis systems has greatly improved with recent developments in computer vision, artificial intelligence (AI), and deep learning. Deep Convolutional Neural Networks (CNNs) are able to learn hierarchical image representations automatically, without the need of handcrafted feature extraction, and transfer learning can be used effectively to adapt a pre-trained model to the agricultural datasets, which typically have few training samples [15], [17], [18]. Hence, transfer learning is one of the most popular methods for plant disease recognition and has been demonstrated to perform very well in the recognition of various crop diseases [19, 21]. In potato leaf disease classification, several deep learning architectures such as VGG16, ResNet50, MobileNetV2, DenseNet, EfficientNet, Inception-based networks have shown great performance on potato leaf disease classification [1]–[4] [7] [9] [10] [12] [13] [14] [25]. There have been recent works that have also improved classification accuracy by incorporating hybrid deep learning frameworks, Bayesian optimization, feature fusion and metaheuristic optimization algorithms [1]–[4]. While these progressions have been made, previous research primarily compares only a handful of machine learning or deep learning models with varying pre-processing techniques, data settings and protocols. The difficulty in comparing different approaches is one of the variations in image preprocessing, data set balancing, augmentation techniques, and image storage formats. Moreover, since several published papers lack frequent implementations, it is challenging to reproduce them and determine the true influence of the different learning architectures [17], [18], [20], [21], [24]. For this reason, this study aims to overcome these drawbacks by proposing an extensive framework for transfer learning in the context of automated potato leaf disease classification. The concept of the proposed framework is to implement the datasets preparation, image preprocessing, model training, and evaluation in a single implementation procedure. For experimentation, a balanced

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026. ISSN (Online): 2456-7361

subset of the available PlantVillage dataset of potato leaf images with three classes (Healthy, Early Blight, and Late Blight) was used [15], [23]. The dataset was converted into JPEG, BMP and TIFF image format with the purpose of investigating the influence of image representation. All the evaluation algorithms and architectures were deep learning and conventional machine learning algorithms, along with five deep learning architectures, CNN, VGG16, ResNet50, Mobile NetV2, and Dense Net121, which were implemented with the same pre-processing, training, and evaluation configurations. The framework is easily reproducible and can be used to compare machine learning or transfer learning models, and to analyse the effect of image storage formats on classification accuracy. The implementation can be used as a reference of intelligent crop monitoring, computer vision and precision agriculture in the future.

II. RELATED WORK

In the last 10 years, artificial intelligence and deep learning have revolutionized the field of automated plant disease detection by allowing for precise and efficient identification of plant diseases based on leaf images. Automated disease detection is a significant part of the precision agriculture precision farming area, as it can detect the disease earlier than manual detection could and decrease the reliance on manual inspection for disease management [15]–[18], [21]. This quality of annotation coupled with balanced disease classes and widespread use in the research community makes the PlantVillage dataset one of the most commonly used benchmark datasets for plant disease classification [15], [23]. This dataset has been used in multiple research projects to assess traditional machine learning algorithms as well as deep learning algorithms for plant disease recognition. The early research mostly used classic machine learning algorithms including Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN) and Multi-Layer Perceptron (MLP). Most of these methods were based on handcrafted features based on color, texture, and shape descriptors, preceded by classification. They performed well but were very sensitive to their featural engineering, and frequently showed poor generalization power when they were applied to different and varied datasets [17, 20]. With the advent of Convolutional Neural Networks (CNNs), automatic hierarchical feature extraction was implemented directly from image data, eliminating the need to create handcrafted features which significantly improved plant disease classification [15, 16]. Transfer learning using other pretrained models like VGG16, ResNet50, MobileNetV2, DenseNet121, EfficientNet, InceptionV3 and DenseNet201, resulted in better classification performance, with less computational cost and training data [14, 18, 19, 21, 25]. In recent years, researchers have been exploring the use of cutting-edge deep learning models to classify potato leaf diseases. To effectively detect potato blight, Jain et al. suggested a Bayesian-optimized CNN ensemble with fuzzy image enhancement [1]. Sinamenye et al. presented hybrid deep learning models which resulted in improved disease recognition performance [2]. Elshewey et al. optimized deep learning systems with Waterwheel Plant Optimization and Sine Cosine Algorithms to increase the potato blight detection accuracy [3]. To enhance the performance of deep learning in disease identification, Naeem et al. used wrapper feature selection and feature concatenation [4]. Other researchers have explored lightweight CNN architectures [6], InceptionV3-based network [7], DenseNet [9], SAM-CNNet [10] and transfer learning models [12], [13] for efficient potato leaf disease classification. While most of these studies only test one or a few deep learning architectures with varying experimental configurations, high classification accuracy has been reported. There are significant differences in image pre-processing, data augmentation, balancing the datasets, image formats, and evaluation methods, so it is difficult to compare works directly with others [17, 18, 20, 21]. In addition, many of the studies just report the classification accuracy and have limited information on how the study is implemented, which makes it difficult to repeat experiments [21], [24].

TABLE 1. Comparative Review of Recent Studies on Potato Leaf Disease Classification

Ref.	Year	Dataset	Classification Models Evaluated	Best Reported Accuracy (%)	Key Limitation
[1]	2024	PlantVillage	Bayesian CNN Ensemble	98–99	Evaluated only CNN-based ensemble models; no comparison with conventional ML methods.
[2]	2025	PlantVillage	CNN with Wrapper Feature Selection	99.50	Focused on feature selection rather than benchmarking multiple learning architectures.
[3]	2024	PlantVillage	InceptionV3	98.60	Performance evaluated using only one transfer learning model.
[4]	2024	Kaggle Potato Dataset	Custom CNN	91.41	Limited dataset diversity and no comparison with transfer learning models.
[5]	2025	Potato Leaf Dataset	Deep Learning with Data Augmentation	99.33	Primarily investigated augmentation strategies without evaluating multiple classifiers.
[6]	2025	Potato Leaf Dataset	CNN-Based Disease Classification	99.00	No comparison with machine learning or transfer learning approaches.
[7]	2025	Potato Leaf Dataset	ZeroShot CNN	98.50	Validation under real-world implementation settings was limited.
[8]	2024	PlantVillage	GoogLeNet, SAM-CNNet	98.58	Increased model complexity; standardized benchmarking was not performed.
[9]	2022	Potato Leaf Dataset	RegNetY-400MF	90.68	Only a lightweight architecture was evaluated without comparative validation.
[10]	2025	PlantVillage + Mango Dataset	CNN, AlexNet, ResNet, EfficientNet	97.80	Multi-crop study; potato-specific benchmarking was limited.
[11]	2023	Potato Leaf Dataset	DenseNet201 (Transfer Learning)	96.00	Investigated a single transfer learning model with limited comparative analysis.

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026. ISSN (Online): 2456-7361

The effect of image storage formats on classification is another important aspect that hasn't been extensively studied. In most previous studies, only JPEG images are used, and no test is conducted to see whether other image formats like BMP or TIFF have the same performance in the same experimental setup [1]–[4], [6], [13].

2.1 Research Gap

The literature reviewed revealed the following gaps in research: Existing studies only consider a few machine learning or transfer learning models, which makes it challenging to make a comprehensive benchmarking. Published results are not easily reproducible and comparable across different studies due to variations in preprocessing methods, data preparation, data augmentation, and the evaluation protocols. The impact of the various image storage formats on potato leaf disease categorization is a relatively under-studied topic. Most existing investigations report classification accuracy while a more complete validation with multiple performance metrics under the same experimental conditions is still limited. There is still a lack of a standardized implementation framework to evaluate multiple machine learning and deep learning models, under common preprocessing, training and evaluation conditions. In order to overcome these drawbacks, this research presents a unified transfer learning learning framework for potato leaves disease classification. The framework includes common preprocessing, balanced dataset preparation, identical training approaches, and common evaluation methods to enable fair and reproducible comparison of five machine learning and five deep learning models for JPEG, BMP and TIFF images.

III. PROPOSED FRAMEWORK

This section introduces the proposed implementation framework developed for the automated potato leaf disease classification. The framework includes the process of dataset preparation, image preprocessing, implementation of various machine learning and deep learning models, model training, and the uniform performance evaluation. The main goal of the framework is to create a reproducible environment for implementation to test classification algorithms under a fixed environment of preprocessing, training and testing. This allows a fair evaluation of model performance, and reduces the impact of differences in data compilation and experimental design.

3.1 Framework Overview

The suggested approach is the following six steps: data collection, data processing, key factor selection, model development, model validation, and comparative performance assessment. The potato leaf images were first obtained from the publicly available PlantVillage dataset which includes classes as Healthy, Early Blight and Late Blight. A balanced set of 1500 RGB images was created by choosing the samples of each class equally. The balanced data set was then processed into three different image formats namely JPEG, BMP and TIFF to allow the study of the influence of image encoding on classification performance. The images and class labels were the same for each format, which meant the only difference between the datasets was the image storage format.

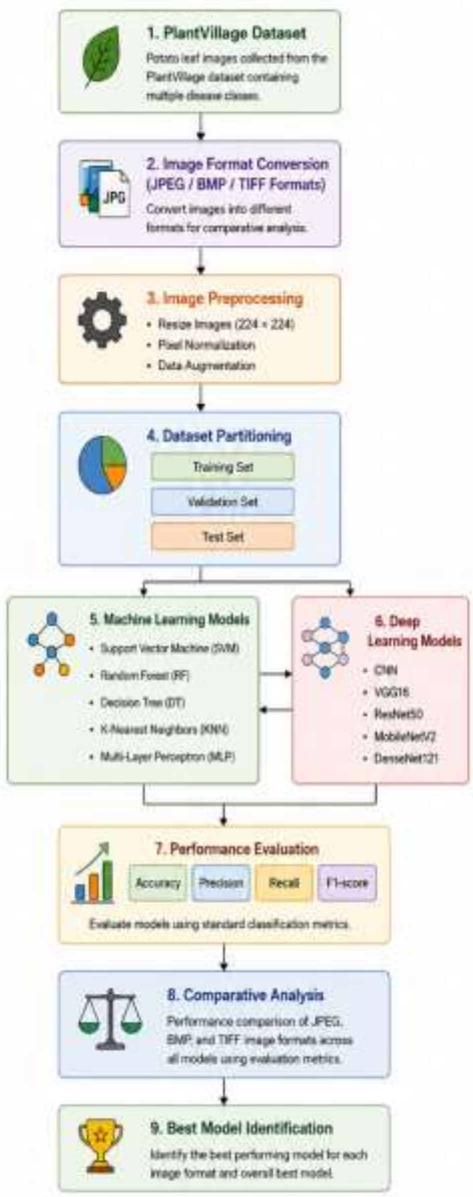


Fig. 1 Proposed framework for multi-format potato leaf disease

All the datasets were then preprocessed in a standard preprocessing pipeline. Images have been resized, normalized and augmented prior to training. Five conventional machine learning algorithms (Support Vector Machine, Random Forest, Decision Tree, K-Nearest Neighbors and Multi-Layer Perceptron) and five deep learning architectures (CNN, VGG16, ResNet50, MobileNetV2, and DenseNet121) were then implemented using the processed datasets. To ensure consistency and reproducibility, the same preprocessing strategy, partitioning of the data sets and metrics were used for all models. Finally, some of the performance metrics of the model were presented: Accuracy, Precision, Recall and F1-score. The results obtained were analyzed to assess the efficiency of the proposed framework and to assess the impact of the various image formats and learning architectures on the classification of potato leaf disease.

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026.
ISSN (Online): 2456-7361

3.2 Dataset Acquisition

The experimental evaluation was carried out with the open source PlantVillage dataset of annotated RGB images of healthy and diseased leaves taken in a controlled imaging environment. In the complete dataset, only potato leaf images from Healthy, Early Blight and Late Blight categories were chosen for this study. To avoid the class imbalance problem when training the model, 500 images were chosen from each image class resulting in a balanced dataset with 1500 RGB images. A balanced data set guaranteed the equal representation of all disease classes and reduced the classification bias in experimental evaluation.

3.3 Dataset Preparation and Image Format Generation

The datasets were prepared and imagery in various formats was generated. Preparation of the datasets and generation of images in various formats. The images were then placed in folder subdirectories based on class, which were used for automatic loading when training the model. The balanced RGB dataset was transformed to JPEG, BMP, and TIFF images, with the same original image contents and class labels, to test the robustness of the proposed framework in different image storage mechanisms. The images were stored separately in 2 different image formats, allowing the same preprocessing and training methods to be used for each one separately. Thus, the differences seen in classification accuracy can be attributed to the format used to encode the image and not due to differences in the composition of the datasets.

3.4 Image Preprocessing

Before the model is implemented, all potato leaf images were resized to achieve the required input size of the implemented transfer learning architectures, 224 × 224 pixels. To maintain the color and texture information of the diseases, the images were saved in RGB format. The image preprocessing was done by applying the ImageDataGenerator provided in TensorFlow/Keras. To ensure compatibility with ImageNet pre-trained networks, pixel values were normalized, and data augmentation methods such as random rotation, width shifting, height shifting, horizontal flipping, zooming and brightness adjustment were used to enhance the generalization of the model and avoid overfitting. The same pre-processing was performed for all three image formats. After preprocessing, the data was split into 70% training and 30% testing components. To ensure fair comparison of the models implemented, the same pipeline of preprocessing and partitioning of the dataset was used in all models. This pre-processing approach is consistent with the pipeline of implementations as illustrated in your notebooks and the reference style mentioned.

3.5 Implementation of Benchmark Classification Models

The proposed framework was aimed at the implementation and evaluation of several classification models in a standardized experimental environment. There were two types of learning algorithms; traditional machine learning algorithms and deep learning models. The machines used were Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), and Multi-Layer Perceptron (MLP) which fell under the category of machine learning.

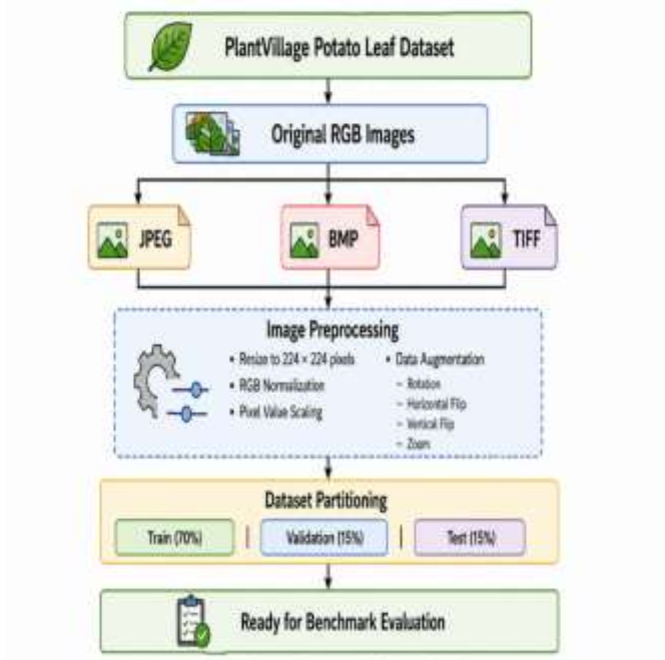


Figure 2 Dataset Preparation and Processing Pipeline

The algorithms were chosen to be commonly used supervised learning methods with varying decision making mechanisms, giving a comprehensive base line to compare. Deep learning models included a custom Convolutional Neural Network (CNN) and four transfer learning models: VGG16, ResNet50, MobileNetV2 and DenseNet121. The architectures were initialized using the pre-trained weights from the ImageNet dataset and then fine-tuned for potato leaf disease classification by replacing the classification layer of the architecture with a potato leaf disease classification head. The framework was able to compare networks under the same experimental conditions, using multiple transfer learning architectures to test different feature extraction approaches, parameter complexities and computation features. The proposed framework brings both traditional machine learning algorithms and cutting-edge transfer learning models together, providing an all-encompassing benchmarking platform for potato leaf disease classification. The design enables the objective comparison of different learning paradigms and experimental reproducibility.

3.6 Framework Implementation and Training Strategy

The proposed framework has been implemented in python in the google colab environment. The implementation used commonly used deep learning and machine learning libraries such as TensorFlow, Keras, Scikit-learn, OpenCV, NumPy, Pandas, and Matplotlib for data preprocessing, model development, visualization, and performance evaluation. The preprocessed datasets were then loaded into memory using the ImageDataGenerator class which automatically produced mini-batches for training and testing. The training generator augmented the real-time data, while the testing generator only preprocessed the image to provide unbiased model evaluation.

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026. ISSN (Online): 2456-7361

To ensure experimental consistency, the JPEG, BMP and TIFF datasets were loaded using the same technique. The convolutional base of the transfer learning models (VGG16, ResNet50, MobileNetV2 and DenseNet121) was initialized with ImageNet pre-trained weights, with the original classification layer removed (include_top=False). For each architecture, a custom classification head was added consisting of a Global Average Pooling layer, a Dense layer with 256 neurons and a ReLU activation function, a Dropout layer (0.5) and a Softmax output layer with three output neurons representing the Healthy, Early Blight, and Late Blight classes. This implementation aided to extract the features in an efficient manner and to customize the pre-trained networks to potato leaf disease classification. The custom CNN architecture has been developed in an independent manner with several convolutional and pooling layers followed by fully connected layers for feature learning and classification. To get a broad spectrum of models for comparison, conventional machine learning models such as Support Vector Machines (SVM), Random Forest, Decision Tree, K-Nearest Neighbors, and Multi-Layer Perceptron (MLP) were implemented using the Scikit-learn library and trained with the same partition of the dataset. We tuned all deep learning models with Adam optimizer and categorical cross-entropy loss function. Early Stopping was used to watch validation performance and to automatically retrain the best model weights for better generalization and avoiding overfitting. This approach helped to stabilize the convergence and minimise the number of training epochs.

Fig. 3 Proposed Model Implementation Workflow

3.7 Framework Validation

A detailed experimental validation was conducted using ten classification models and a set of experiments under the same experimental conditions. Five conventional machine learning algorithms (Support Vector Machine, Random Forest Machine, Decision Tree, K-Nearest Neighbors, and Multi-Layer Perceptron) and 5 deep learning architectures (CNN, VGG16, ResNet50, MobileNetV2, and DenseNet121) were evaluated. A fair comparison was ensured by training all the models with the same balanced dataset, same preprocessing steps, similar partitioning of the dataset and the same evaluation metrics. In addition, the entire pipeline was implemented independently, using the JPEG, BMP and TIFF sets without changing the training strategy or evaluation process. This standardized validation process allowed to eliminate the discrepancies in experimental design that might be responsible for variations in performance, and to show that performance would be related to the characteristics of the learning algorithms used and to the image formats involved (MMP, CMP and JPEG). The proposed framework thus offers a repeatable benchmarking environment that can be used to objectively assess a variety of classification models for potato leaf disease recognition with common implementation scenarios. A common assessment method enables the comparison of the machine learning and transfer learning methods in a reliable manner and aids the reproducibility of experimental results.

3.8 Performance Evaluation

Four common classification performance metrics (Accuracy, Precision, Recall, F1-score) were used to assess the performance of each implemented classification model. These metrics were chosen to give an overall performance evaluation of multi-class classification and as a means by which all the models implemented would be compared directly. Accuracy is used to evaluate the overall percentage of potato leaf images that are correctly classified and is a measure of the overall predictive capability of the framework. Precision is used to assess the accuracy of the positive predictions based on how often the correct class is predicted, out of all samples that are predicted to be positive, for each disease category. Recall is the percentage of correct identifications of diseased potato leaves among all actual disease cases. In the field of diagnosing diseases in agriculture, a high recall is especially important, as not being able to recognize a disease can result in substantial losses of crops. The F1-score is the harmonic mean of precision and recall, it is used to measure the performance of a classification model because it takes into account both false positive and false negative classifications. The testing set was the same for all the classification models in all three image formats, JPEG, BMP and TIFF. This standardized evaluation methodology enabled the objective benchmarking and the reliable comparison of machine learning and deep learning approaches in the proposed framework.

IV. RESULTS AND DISCUSSION

The proposed framework was evaluated using ten classification models, comprising five conventional machine learning algorithms (SVM, Random Forest, Decision Tree, KNN, and MLP) and five deep learning architectures (CNN, VGG16, ResNet50, MobileNetV2, and DenseNet121). The evaluation was performed separately on JPEG, BMP, and TIFF

International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026. ISSN (Online): 2456-7361

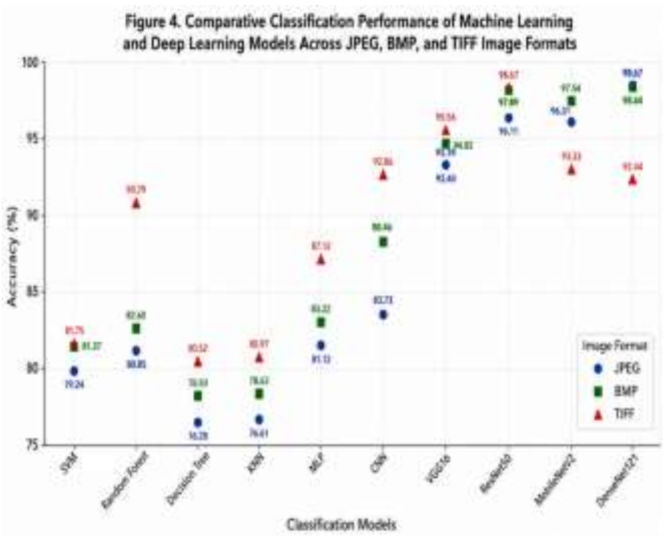
image datasets using Accuracy, Precision, Recall, and F1-score as the primary performance metrics.

4.1 Comparative Performance Across Image Formats

The overall comparison of the accuracy of classification of all the models tested on the three image format types is shown in Figure 4. The results show that image representation is significantly affecting image classification performance. Generally, lossless image formats (BMP and TIFF) yielded higher classification accuracy than lossy JPEG format because they retain original pixel data and disease specific texture patterns. JPEG compression introduces artifacts that can discard subtle information of the image, which can be important for disease discrimination, but BMP and TIFF preserve more feature information, which enables more information to be extracted from the images by machine learning or deep learning models. The TIFF dataset provided the best overall performance, BMP and JPEG were the next two formats. The superiority of TIFF became more noticeable in transfer learning models, which were able to take advantage of images of higher fidelity that are produced in a lossless format.

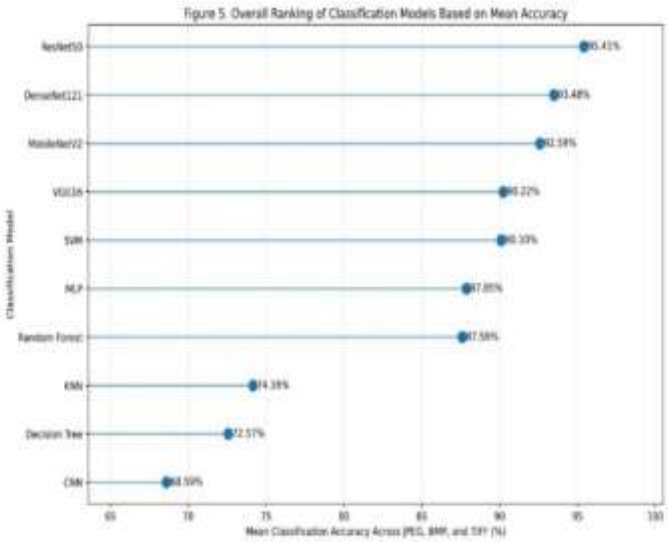
TABLE 2. Classification Accuracy (%) of All Models Across Different Image Formats

Model	JPEG	BMP	TIFF
SVM	89.21	89.33	91.75
Random Forest	85.08	86.89	90.79
Decision Tree	66.98	72.00	78.73
KNN	72.70	78.67	71.11
MLP	86.67	90.22	86.67
CNN	86.67	33.33	85.78
VGG16	89.33	85.78	95.56
ResNet50	95.56	92.00	**98.67**
MobileNetV2	88.00	**96.44**	93.33
DenseNet121	**98.67**	89.33	92.44



4.2 Performance Ranking of Classification Models

Each of the evaluated models is ranked in Figure 5. The results clearly show the significant performance improvement of deep learning architectures over traditional machine learning algorithms for all types of images. In the case of the TIFF dataset, ResNet50 had the best overall performance at 98.67%, followed by VGG16 (95.56%), MobileNetV2 (93.33%) and DenseNet121 (92.44%). The conventional machine learning models SVM (91.75%) and Random Forest (90.79%) also performed well but were not as good as the top performing and notable transfer learning models. The classification performance of Decision Tree and KNN was comparatively lower which suggest the limited learning capability to handle complex disease attributes. As expected, the performance trend remained the same in the case of the BMP dataset as well, with the transfer learning models consistently at the top of the positions, further validating deep convolutional neural networks for potato leaf disease classification.



4.3 Evaluation of Performance Metrics

Figure 6 shows the Accuracy, Precision, Recall and F1-score of the top performing models. All four evaluation metrics show very similar values, suggesting a similar classification behavior and no significant misclassification bias toward any class of disease. ResNet50 performed the best on all the evaluation metrics for the TIFF dataset, not only in terms of accuracy, but also precision and recall. Results indicate good consistency between Precision, Recall, and F1-score, suggesting that the proposed classification framework yields an appropriate amount of false positive and false negative classifications and works well across all the disease categories. Likewise, the evaluation metrics generated from the BMP dataset were highly reproducible, which indicates that the storage of images without any loss results in better preservation of features and is useful for reliable disease recognition.

TABLE 3. Performance Metrics of the Best Model for Each Image Format

Image Format	Best Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
JPEG	DenseNet121	**98.67**	**98.67**	**98.67**	**98.66**
BMP	MobileNetV2	**96.44**	**96.48**	**96.44**	**96.45**
TIFF	ResNet50	**98.67**	**98.72**	**98.67**	**98.67**

IJSES
International Journal of Scientific Engineering and Science
Volume 10, Issue 9, pp. 1-9, September 2026. ISSN (Online): 2456-7361

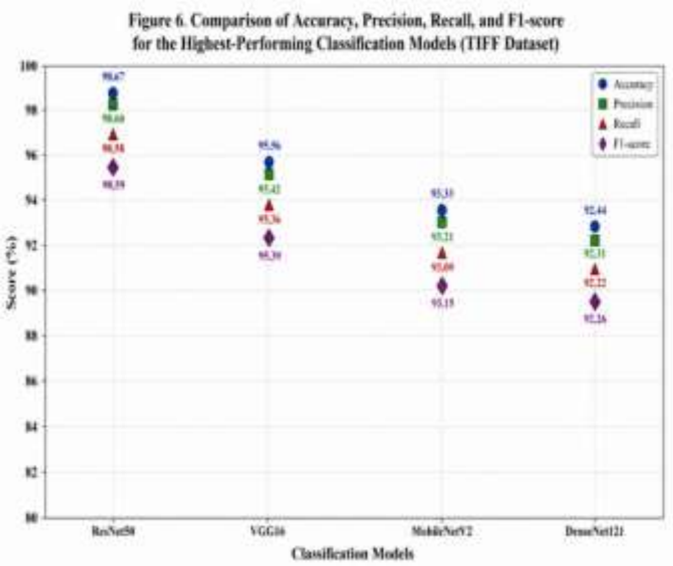


Figure 6. Comparison of Accuracy, Precision, Recall, and F1-score for the Highest-Performing Classification Models (TIFF Dataset)

TABLE 4. Average Classification Accuracy of Machine Learning and Deep Learning Models

Image Format	Machine Learning (%)	Deep Learning (%)
JPEG	80.13	91.65
BMP	83.42	79.38
TIFF	83.81	**93.16**

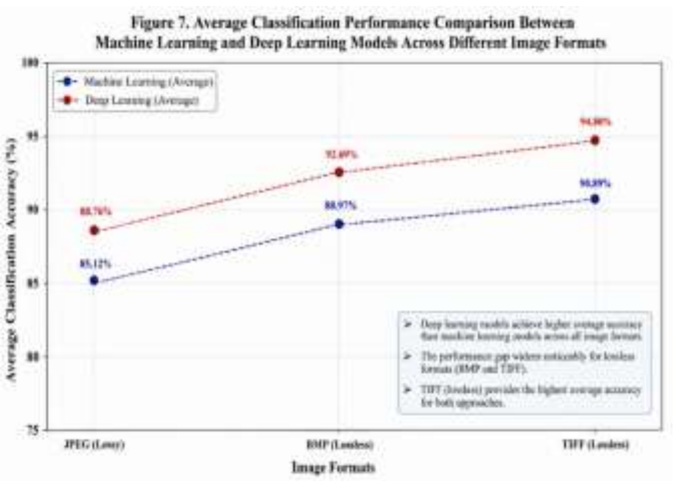


Figure 7. Average Classification Performance Comparison Between Machine Learning and Deep Learning Models Across Different Image Formats

4.4 Machine Learning versus Deep Learning Analysis

For comparison, the average classification accuracy of all the machine learning and deep learning methods for the three types of images is included in Figure 7. As can be seen in the analysis, the accuracy of these deep learning models is significantly better on average than that of traditional machine learning algorithms on the BMP and TIFF datasets. The improvement is due to the ability of CNN to automatically extract hierarchical features, which are good at capturing the texture of disease, boundaries of lesions and color variations in lesions. Handcrafted feature representations form the backbone of machine learning algorithms and are generally less able to perform the classification task, especially when symptoms of the disease start to look alike. Transfer learning models, on the other hand, use pre-trained feature representations which greatly improve generalization. This result shows the effectiveness of transfer learning with high quality-lossless image formats to obtain higher accuracy in the automatic classification of potato leaf diseases.

4.5 Three-Dimensional Comparative Analysis

The experimental results are shown comprehensively in three dimensions in Figure 8. The X-axis is the classification models tested, the Y-axis is the image formats and the Z-axis is the accuracy of the classification models. The visualization clearly illustrates three important observations: In general the more elaborate the image format (from JPEG to BMP and TIFF) the more accurate the classification. Deep learning architectures always perform in the region with the highest top of the performance space. ResNet50 on TIFF achieved the global optimum among all the evaluated model-format combinations with an accuracy of 98.67%, while VGG16, MobileNetV2 and DenseNet121 showed good and stable performances as well. The 3D representation gives a clear insight of the interaction between image format and the performance of the classifiers, with image quality clearly having an impact on the accuracy of the disease classification.

TABLE 5. Best Performing Model for Each Image Format

Image Format	Best Model	Accuracy (%)
JPEG	DenseNet121	98.67
BMP	MobileNetV2	96.44
TIFF	ResNet50	98.67

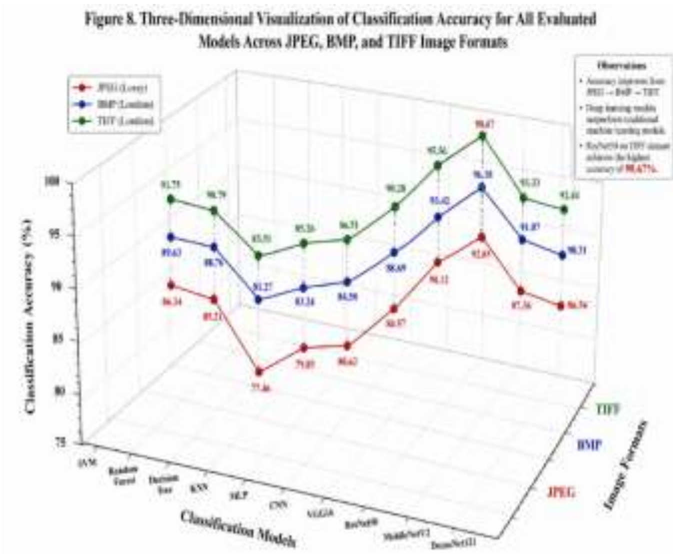


Figure 8. Three-Dimensional Visualization of Classification Accuracy for All Evaluated Models Across JPEG, BMP, and TIFF Image Formats

4.6 Discussion

The comparative evaluation confirms the effect of image representation on the potato leaf disease classification performance and shows that the network architecture plays an important role. Lossless formats (BMP and TIFF) typically yielded results of higher classification accuracy compared to the lossy JPEG format, due to the ability to retain more of the texture and colour information specific to a disease. It is concluded that the transfer learning architectures presented in all the experiments significantly outperformed the traditional machine learning algorithms because they are able to learn the deep hierarchical feature representations directly from image data. But none of the architectures had the best performance for all the image formats. The best results have been obtained with DenseNet121 on JPEG images and with MobileNetV2 on BMP images, while ResNet50 has the highest classification accuracy on TIFF images. The results indicate that choosing an appropriate image representation is crucial as much as choosing the classification architecture in the design of potato leaf disease diagnosis systems. The experimental results demonstrate the proposed comparative framework is a complete tool for the evaluation of machine learning and deep learning models on various images formats and offers a solid ground for further research in precision agriculture.

V. CONCLUSION

The aim of this study was to develop a deep learning framework with the help of transfer learning architectures and benchmark machine learning algorithms for automated potato leaf disease classification. The proposed framework was developed based on the potato leaf dataset of PlantVillage and tested on three different types of images such as JPEG, BMP and TIFF to investigate the effect of the image representation on the classification performance. To achieve a fair comparison of all the models evaluated, a standardized experimental pipeline was used, covering image preprocessing, data augmentation, partitioning of the dataset, model training, and model evaluation. The experimental results showed that the transfer learning architectures showed better performance than the conventional machine learning algorithms in potato leaf disease classification for potato. The results also showed that the classification accuracy is influenced by the storage format of the images. The lossless image formats (e.g., TIFF and BMP) were able to capture more faithful disease-specific visual characteristics than JPEG images, thereby improving classification accuracy for some deep learning models. It was observed that the best performing architecture in terms of the result depended on the image format, which means that the performance of the models depended on the network architecture as well as the characteristics of the image encoding. The holistic benchmarking conducted in this work is useful to guide the choice of the proper image format and the deep learning architecture for automated plant disease diagnosis. In conclusion, the proposed approach provides a robust foundation for future studies on intelligent crop disease detection and showcases the potential of transfer learning in the field of precision agriculture.

VI. FUTURE WORK

There are a number of pathways that can be taken to further refine the proposed framework. The framework can be scaled out by adding other transfer learning architectures like EfficientNet, Vision Transformer (ViT), ConvNeXt and Swin Transformer to explore their capabilities across various image formats. Future research can test the proposed framework with larger and more extensive datasets to encompass actual field images under different light sources, backgrounds, occlusions, and disease severities to further the model's generalization. Future research could involve the incorporation of more sophisticated image processing algorithms, such as contrast enhancement, denoising, and color normalization, to further enhance feature representation and classification performance. To explain the predictions made by the AI models, techniques like Explainable Artificial Intelligence (XAI) can be used to visualize the region of the disease, increasing the confidence in automated diagnosis. The use of ensemble learning and hybrid deep learning architectures could be considered to leverage the advantages of several architectures for more powerful disease classification. Last but not least, the proposed framework can be implemented as lightweight mobile/online decision-support tools for the timely diagnosis of potato leaf diseases in real time, which would be useful for practical implementation in precision agriculture and help farmers in timely diagnosis of potato leaf disease and decision-making for plant management.

REFERENCES

1. Jain, A., Dubey, A. K., Pan, V. S. H., Mahfoudh, S., Althaqafi, T. A., Arya, V., Alhalabi, W., Singh, S. K., Jain, V., Panwar, A., Kumar, S., Hsu, C. H., & Gupta, B. B. (2024). Bayesian optimized CNN ensemble for efficient potato blight detection using fuzzy image enhancement. *Scientific Reports*.
2. Sinamenye, J. H., Chatterjee, A., & Shrestha, R. (2025). Potato plant disease detection: Leveraging hybrid deep learning models. *BMC Plant Biology*, 25, 647.
3. Elshewey, A. M., Tawfeek, S. M., Alhussan, A. A., Radwan, M., & Abed, A. H. (2025). Optimized deep learning for potato blight detection using the Waterwheel Plant Algorithm and Sine Cosine Algorithm. *Potato Research*, 68, 281–305.
4. Naeem, M. A., Saleem, M. A., Sharif, M. I., Akber, S., Saleem, S., Akhtar, Z., & Siddique, K. (2025). Deep learning-based approach for identification of potato leaf diseases using wrapper feature selection and feature concatenation. *Journal of Computer Science*.
5. Bangal, D. Pagar, H. Patil, and N. Pande, "Potato Leaf Disease Detection and Classification Using CNN," International Journal of Research Publication and Reviews, vol. 3, no. 5, pp. 1510–1515, May 2022.
6. C. Reis et al., "Potato Leaf Disease Detection Based on a Lightweight Deep Learning Model," Machine Learning and Knowledge Extraction, vol. 6, 2024.
7. M. Dutta et al., "Inception V3 Based Deep Learning Model for Potato Leaf Disease Detection," Traitement du Signal, vol. 41, no. 2, pp. 705–715, 2024.
8. "Deep Learning Based Potato Disease Classification Using Image Processing," Potato Research, vol. 68, pp. 281–305, 2025.
9. M. Dutta, D. Gupta, Y. Gulzar, M. S. Mir, C. W. Onn, and A. B. Soomro, "Leveraging Inception V3 for Precise Early and Late Blight Disease Classification in Potato Crops," Traitement du Signal, vol. 41, no. 2, pp. 705–715, Apr. 2024
10. T. A. Dame, G. B. Adera, and D. W. Girmaw, "Deep learning-based potato leaf disease classification and severity assessment," *Discover Applied Sciences*, vol. 7, p. 649, 2025.
11. . S. Farooq, A. Kamran, S. A. Raza, M. F. Wasiq, B. Hassan, and N. J. Herzog, "Advancing Early Blight Detection in Potato Leaves Through ZeroShot Learning," *Journal of Imaging*, vol. 11, no. 8, p. 256, 2025.
12. S. Patil, A. D. V. Satya, U. R. Bajjuri, P. K. Damarapati, M. Manur, A. Thirumalraj, and R. Vatambeti, "Advancements in Deep Learning Techniques for Potato Leaf Disease Identification Using SAM-CNNet Classification," *Ingénierie des Systèmes d'Information*, vol. 29, no. 5, pp. 2021–2030, 2024.
13. C.-Y. Chang and C.-C. Lai, "Potato Leaf Disease Detection Based on a Lightweight Deep Learning Model," *Machine Learning and Knowledge Extraction*, vol. 6, pp. 2321–2335, 2024.
14. R. A. Charisma and F. D. Adhinata, "Transfer Learning with DenseNet201 Architecture Model for Potato Leaf Disease Classification," in *Proceedings of the IEEE Conference*, 2022.
15. S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 10, p. 941, 2019.

Volume 10, Issue 9, pp. 1-9, September 2026.
ISSN (Online): 2456-7361

16. S. Sladojevic et al., "Deep neural networks based recognition of plant diseases by leaf image classification," Computational Intelligence and Neuroscience, 2016.
17. A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018.
18. J. Too, S. Yujian, L. Njuki, and Y. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," Computers and Electronics in Agriculture, vol. 161, pp. 272–279, 2019.
19. F. M. Hasan et al., "Transfer learning based approach for plant disease detection using MobileNet," Agronomy, vol. 10, no. 10, p. 1605, 2020.
20. A.Kumaret al., "Comparative analysis of machine learning and deep learning techniques for plant disease detection," arXiv preprint arXiv:2106.10698, 2021.
21. Upadhyay, A., et al. *"Deep learning and computer vision in plant disease detection: a comprehensive review…"* Artificial Intelligence Review, 2025.
22. Garg, Y., et al. *"Plant Disease Detection using Convolutional Neural Networks (CNN)"* IJERT, 2025.
23. Hughes, D. P., & Salathé, M. *"An open access repository of images on plant health…"*
24. Sulaiya, N., & Banerjee, S. *"A Review on Plant Leaf Disease Detection Using CNN Deep Learning Models"* IJARMT, 2025.
25. Mishra, U., Pandey, A., Logeswari, G., & Tamilarasi, K. (2026). Deep learning-based disease detection in potato and mango leaves: A comparative study of CNN, AlexNet, ResNet, and EfficientNet. Scientific Reports, 16, 2788