# Improving Potato Disease Classification with Convolutional Neural Networks and Feature Fusion

Sakshi Deshmukh
*School of E&TC Engineering*
*MIT Academy of Engineering*
Pune, India
deshmukh.sakshi@mitaoe.ac.in

Anushka Parkhi
*School of E&TC Engineering*
*MIT Academy of Engineering*
Pune, India
anushka.parkhi@mitaoe.ac.in

Shravani Gandhi
*School of E&TC Engineering*
*MIT Academy of Engineering*
Pune, India
shravani.gandhi@mitaoe.ac.in

Diptee Ghusse
*School of Computer Engineering*
*MIT Academy of Engineering*
Pune, India
diptee.chikmurge@mitaoe.ac.in

***Abstract*—Accurate prediction of potato diseases plays a vital role in maintaining crop health and maximizing agricultural yield. This study presents a deep learning approach utilizing Convolutional Neural Networks (CNNs) integrated with advanced feature fusion strategies to enhance the precision of disease classification. The methodology employs three feature fusion strategies—Early, Intermediate, and Late Fusion—on a benchmark dataset of 2,152 potato leaf images categorized into Early Blight, Late Blight, and Healthy. The workflow includes data preprocessing, feature extraction using CNNs, and classification through fully connected layers.**

**Results indicate that Late Fusion outperforms other strategies, achieving the highest classification accuracy and robust generalization despite class imbalance. It effectively distinguishes between Early and Late Blight, while Early Fusion shows better performance for Healthy leaves. The study highlights the importance of feature fusion and data augmentation in enhancing disease diagnosis, paving the way for scalable and efficient agricultural solutions.**

***Index Terms*—Potato disease prediction, CNN, feature fusion, deep learning**

## I. Introduction

Potatoes are a vital crop globally, but they are vulnerable to diseases such as Early Blight and Late Blight, which can significantly impact both yield and quality. Timely and accurate identification of these diseases is essential for effective management. However, traditional diagnostic methods are often slow and rely on expert input. This study focuses on utilizing deep learning techniques, specifically Convolutional Neural Networks (CNNs), to classify potato leaf images into three categories: Early Blight, Late Blight, and Healthy. To enhance classification performance, three different feature fusion methods are explored and evaluated. The goal is to identify the most effective fusion strategy for accurate disease detection.

The structure of the paper is as follows: Section 2 provides an overview of existing techniques and identifies gaps in the current research. Section 3 describes the dataset used and the augmentation strategies employed to ensure class balance. Section 4 outlines the methodology, detailing preprocessing steps and the implementation of Early, Intermediate, and Late Fusion approaches for combining features. Section 5 presents the results, including an analysis of the model's accuracy and reliability. Finally, Section 6 concludes with insights into how this approach can contribute to better agricultural disease management practices. The paper is structured as follows: Section II reviews related work, Section III explains the methodology, Section IV presents the experimental setup, Section V discusses the results, and Section VI concludes with future work.

## II. Literature Survey

A literature review on potato disease classification using deep learning shows that traditional methods are slow and inaccurate, causing significant crop damage. CNN models like VGG16 and VGG19 have been successfully used for classifying potato leaf diseases, achieving high accuracy (around 91%) and automating disease detection, which aids in timely intervention and better crop management [1].One study utilizing K-means clustering along with data augmentation achieved 97% accuracy using the VGG16 model, demonstrating its potential for enhancing agricultural disease management [2]. Another research focused on fine-tuning a pre-trained CNN model, which achieved an accuracy of 97.4% in identifying diseases such as early and late blight, contributing to automated dis- ease detection for sustainable farming [3]. In Bangladesh, a machine learning model trained on 2,034 images reached an impressive 99.23% accuracy, providing an automated system for swift disease diagnosis [4]. A CNN model trained with 2,000 images in India reported 91.41% accuracy, offering valuable support to farmers in managing potato crops [5]. Additionally, an optimized deep learning model achieved 99.22% accuracy, significantly enhancing disease detection and helping reduce crop losses [6]. A study comparing various models, including MLP, achieved 98.3% accuracy for potato disease prediction using weather data and machine learning

[7]. In another study, the combination of pre-trained models and optimization algorithms resulted in an impressive 99.5% accuracy, highlighting the effectiveness of deep learning in agricultural disease diagnosis [8]. Additionally, Convolutional Neural Networks (CNNs) have been successfully applied to identify 13 different plant diseases, with reported accuracies ranging from 91% to 98% [9], [10].. A study on potato disease detection with a dataset of 5,000 images achieved 100% accuracy in some classes, demonstrating deep learning's effectiveness in improving crop management and food security [11], [12]. Using a CNN model on 2,034 images, the model achieved 99% accuracy, addressing various potato diseases and providing an automated solution for early detection [13]. VGG16, VGG19, and ResNet50 were also tested for disease classification, with VGG16 performing at 97% accuracy, improving crop production and early disease detection [14]. A dataset of 3,076 images in uncontrolled conditions achieved 73.63% accuracy with EfficientNetV2B3, showing challenges in real-world disease detection [15]. Lastly, MDSCIRNet, a transformer-based deep learning model, achieved 99.33% accuracy, further improving disease diagnosis and crop production [16]. The work demonstrates how edge detection can enhance the identification of key features in medical images for further analysis and diagnosis [17]. The use of deep learn- ing for feature extraction in agricultural image classification. Their approach inspires similar techniques for processing and interpreting visual data [18].

## III. About Dataset

The dataset [19] utilized in this research includes 2,152 images of potato leaves, categorized into three classes: Early Blight, Late Blight, and Healthy. It contains 1,000 images for both Early Blight and Late Blight, with approximately 152 images representing Healthy leaves, indicating an inherent class imbalance. These images vary in lighting conditions, angles, and leaf states, providing a realistic representation of agricultural environments. This diversity aids in enhancing the robustness of the model in real-world applications.

Early Blight and Late Blight are two of the most devastating diseases affecting potato crops globally. Early Blight, caused by Alternaria solani, typically appears early in the growing season, leading to dark, concentric spots on the leaves. Late Blight, caused by Phytophthora infestans, tends to manifest later but spreads rapidly, causing water-soaked lesions that turn brown and necrotic. Despite their timing differences, both diseases can coexist in the field, complicating manual diagnosis.

The dataset is structured to enable comprehensive analysis and model evaluation. Preprocessing methods, including resizing and normalization, are used to standardize the images. Furthermore, techniques like data augmentation and class balancing are implemented to address the underrepresentation of Healthy leaves, ensuring the model learns effectively from all classes. This dataset serves as the basis for the proposed CNN-based framework with feature fusion, aiming to classify potato leaf conditions with high accuracy and efficiency.

## IV. Workflow for Potato Disease Prediction

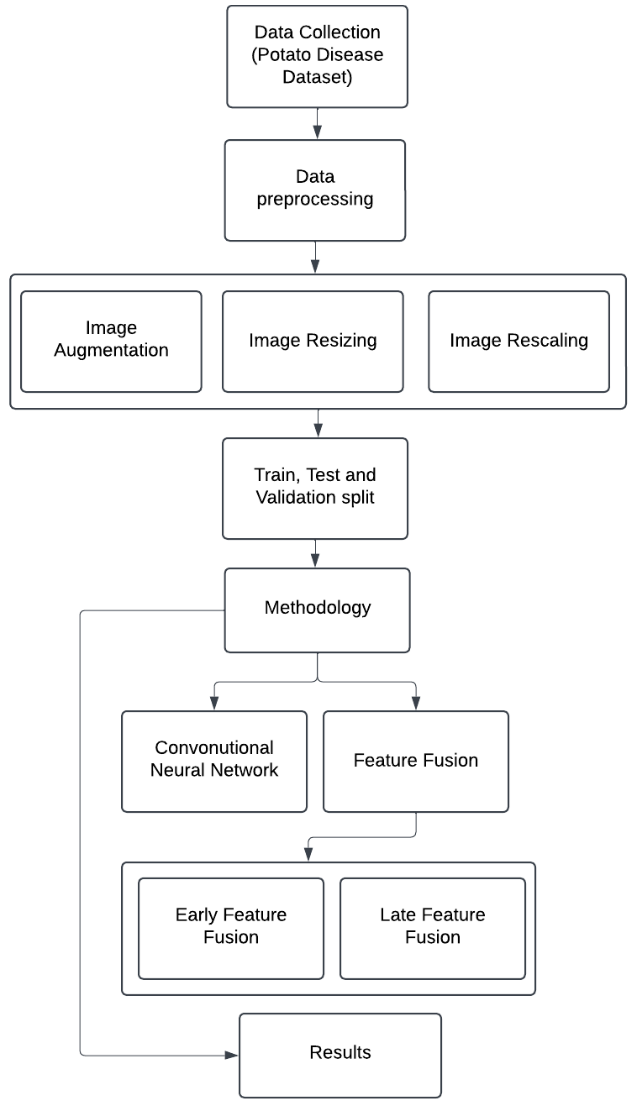


Fig. 1. Workflow Diagram for Potato Disease Classification

Fig.1 outlines the workflow for a potato disease classification model. The process begins with data collection (potato disease dataset), followed by data preprocessing steps, including image augmentation, resizing, and rescaling. The dataset is then split into training, testing, and validation sets.

The methodology section includes the use of Convolutional Neural Networks (CNN) for feature extraction, coupled with feature fusion techniques. These techniques are divided into early feature fusion and late feature fusion. The results are then evaluated to determine the model's performance.

The workflow for predicting potato diseases follows these key steps:

1) **Data Collection and Preprocessing:** The dataset consists of 2,152 potato leaf images, labeled into three

categories: Early Blight, Late Blight, and Healthy. To ensure consistency and mitigate class imbalances, the images are subjected to preprocessing steps includ- ing resizing, normalization, and augmentation. These techniques enhance the model's ability to effectively classify diverse and varied leaf conditions. The images are resized to 224x224 pixels, normalized to the range [0, 1], and augmented using techniques like rotation, zooming, flipping, and shifting.

1) **Feature Extraction:** A pre-trained CNN model (e.g., ResNet or VGG16) is employed to extract features from the images. The CNN layers consist of convolutional operations, max-pooling, and ReLU activation functions.
2) **Feature Fusion:** Three fusion strategies are applied:
   - **Early Fusion:** Features from multiple convolutional layers are combined early in the network.
   - **Intermediate Fusion:** Features from intermediate layers are merged to capture both low-level and high-level information.
   - **Late Fusion:** Features from the final layers are concatenated before classification.
3) **Fully Connected Layer:** The fused feature map is passed through fully connected layers for classification, outputting probabilities for each class (Early Blight, Late Blight, Healthy).
4) **Model Training:** The model is trained using categorical crossentropy as the loss function, and Adam optimizer is applied. Early stopping is implemented to prevent overfitting.
5) **Model Evaluation:** The dataset is split into training (70%), validation (15%), and test (15%) sets. Performance is evaluated using accuracy, precision, recall, F1-score, and confusion matrix.
6) **Results Comparison:** The performance of the three fusion strategies (Early, Intermediate, Late Fusion) is compared based on classification accuracy and other metrics, considering class imbalance.

## V. Methodology

### A. Feature Fusion

Feature fusion is a technique used to combine multiple sources of information, enabling the model to learn more complex patterns and improve performance. In this study, three distinct feature fusion strategies are examined: early fusion, intermediate fusion, and late fusion.

*1) Early Fusion:* Early fusion involves the combination of feature maps from different layers of the convolutional network at the initial stages. This method allows the model to form joint representations by incorporating both low-level and high-level features from the start.

*2) Intermediate Fusion:* Intermediate fusion merges feature maps at one or more intermediate layers of the network. This technique enables the model to capture and integrate both low- and high-level patterns from various stages of the network, contributing to more robust feature extraction.

*3) Late Fusion:* Late fusion brings together feature maps from various layers at the final stage of the network, just prior to the classification layer. This approach ensures that the model makes its final prediction based on a comprehensive set of features extracted across different levels of the network. This approach allows the network to learn separate representations of each feature map before combining them, which can be useful in cases where the feature maps represent different characteristics of the input data.

After fusion, the resulting feature map $\mathbf{F}$ (whether $\mathbf{F}_{\text{early}}$, $\mathbf{F}_{\text{intermediate}}$, or $\mathbf{F}_{\text{late}}$) is passed to the fully connected (dense) layers for final classification.

### B. Fully Connected Layer and Classification

The output of a fully connected layer $\mathbf{z}$ is computed as:

$$\mathbf{z} = \mathbf{W} \cdot \mathbf{F} + \mathbf{b} \tag{1}$$

In equation (1):

- $\mathbf{W}$ is the weight matrix of the fully connected layer.
- $\mathbf{F}$ is the fused feature map (from early, intermediate, or late fusion).
- $\mathbf{b}$ is the bias vector.

The final classification probabilities for each class (Early Blight, Late Blight, and Healthy) are computed using a Softmax activation function shown in equation (2):

$$p(y = c \mid \mathbf{F}) = \frac{e^{z_c}}{\sum_{k=1}^{3} e^{z_k}} \tag{2}$$

Where:

- $p(y = c \mid \mathbf{F})$ is the probability of class $c$ given the feature map $\mathbf{F}$.
- $z_c$ is the output score for class $c$ from the fully connected layer.
- 3 is the total number of classes: Early Blight, Late Blight, and Healthy.
- The denominator sums over all class scores to normalize the values to a probability distribution.

### C. Loss Function

The model is trained using Categorical Cross-Entropy loss, which is commonly used for multi-class classification tasks. The loss function for a single example is given in equation (3) as:

$$L(y, \hat{y}) = -\sum_{c=1} y_c \log(\hat{y}_c) \tag{3}$$

Where:

- $y$ is the true one-hot encoded label vector.
- $\hat{y}$ is the predicted probability vector from the Softmax output.
- $y_c$ is the true label for class $c$, and $\hat{y}_c$ is the predicted probability for class $c$.

The total loss over all training examples is averaged to guide the training process.

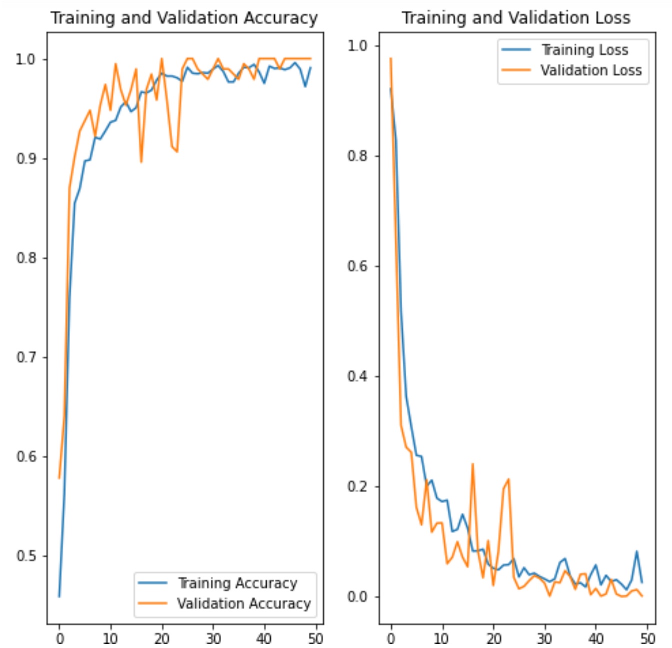


Fig. 2. Training and Validation (Accuracy and Loss)

### *D. Optimization*

To minimize the loss, an optimizer like Adam is used, which adapts the learning rate based on the gradients of the loss function. The weight update rule for Adam is:

$$\mathbf{W}_{t+1} = \mathbf{W}_t - \eta \cdot \frac{\mathbf{m}_t}{\sqrt{\mathbf{v}_t} + \epsilon} \tag{4}$$

In equation (4):

- $\mathbf{W}_t$ is the weight at time step $t$.
- $\mathbf{m}_t$ and $\mathbf{v}_t$ are the first and second moment estimates of the gradients.
- $\eta$ is the learning rate.
- $\epsilon$ is a small constant to prevent division by zero.

The Adam optimizer adapts the learning rate based on both the first moment estimate (mean) and second moment estimate (variance) of the gradients, which helps improve convergence and training stability.

## VI. Results and Discussion

The performance of the three feature fusion strate- gies—Early Fusion, Intermediate Fusion, and Late Fu- sion—was evaluated based on classification accuracy, preci- sion, recall, F1-score, and model generalization. Late Fusion achieved the highest classification accuracy, outperforming both Early and Intermediate Fusion methods. To enhance the comprehensiveness and robustness of this study, several key areas could be addressed. A more detailed explanation of why the Late Fusion strategy outperforms other strategies should be included. This could involve incorporating visualizations of feature maps to better support the findings and provide deeper insights into the model's performance. Additionally, insights into the computational requirements of the proposed approach would significantly improve its practical applicability by helping to assess its feasibility for real-world scenarios.

A comparative analysis with non-CNN methods or alternative techniques, such as attention mechanisms, could further validate the results and highlight the advantages or limitations of the proposed model. Expanding the dataset by incorporating additional data sources, such as environmental factors, would enhance the model's generalizability and make the findings more robust across various conditions.

Lastly, refining the discussion on preprocessing steps and their impact on the overall performance would provide a clearer understanding of how data preparation influences model outcomes. Addressing these recommendations would not only strengthen the paper but also provide a more holistic perspective on the proposed approach.

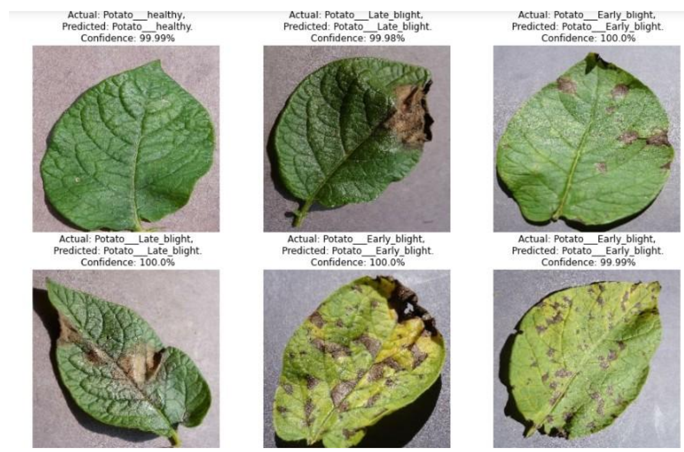


Fig. 3. Training Scores of Various ML Algorithms for RUL Prediction

Fig. 3 shows superior performance in detecting Early Blight and Late Blight, while Early Fusion was better at detecting Healthy leaves due to its focus on low-level features. The confusion matrix revealed that Late Fusion had fewer mis-classifications between Early Blight and Late Blight, whereas Early Fusion struggled with distinguishing these two classes. Despite the class imbalance in the dataset, the fusion strategies, combined with data augmentation, improved the model's ability to generalize, especially for the disease classes. In conclusion, Late Fusion yielded the best results, highlighting the effectiveness of combining features at later stages of the network, while Early and Intermediate Fusion provided valuable insights into the feature interactions at various levels.

## VII. Conclusion

This study highlighted the importance of feature fusion techniques in enhancing the performance of CNN models for potato disease prediction. By evaluating three fusion strategies—Early, Intermediate, and Late Fusion—the results demonstrated that Late Fusion provided superior results, achieving higher classification accuracy and better identification of Early and Late Blight diseases. Late Fusion was particularly effective as it combined features at later stages of the network, which allowed the model to generalize better, despite challenges such as class imbalance in the dataset.

Additionally, the incorporation of data augmentation techniques further strengthened the model's robustness, making it capable of handling variations in the data and improving overall prediction performance. The use of augmentation techniques allowed the model to generalize across a wider variety of disease symptoms and image conditions. This makes the approach especially promising for early disease detection in agriculture, where accurate, real-time monitoring is crucial for effective crop management and timely interventions.

Looking ahead, future research could explore the integration of more advanced fusion strategies, such as attention mechanisms or multi-modal fusion techniques, which could capture more complex relationships between features. Furthermore, incorporating additional data sources, such as environmental or weather data, could enhance model accuracy by providing a more comprehensive understanding of the factors influencing disease progression. These advancements would contribute to improving both the prediction accuracy and efficiency of agricultural disease detection systems, paving the way for more effective and scalable solutions.

## VIII. Future Scope

Addressing dataset imbalance remains a critical area for improvement. Techniques such as synthetic data generation or oversampling can be explored to ensure balanced representation across all classes, enhancing the model's generalization and reliability. Additionally, incorporating more diverse data sources, including environmental or contextual factors, could expand the applicability of the proposed approach to real-world scenarios. Further studies could also investigate alternative methods, such as attention mechanisms, to compare their performance against the current model. Visualizing feature maps and analyzing computational requirements in detail would provide deeper insights into the model's behavior and practical feasibility.

## Acknowledgment

Our heartfelt appreciation goes out to everyone who helped us finish this study report. We would especially want to thank Dr. Diptee Ghusse for all of her help and assistance during the research process.We also acknowledge the support of the MIT Academy of Engineering.

## References

[1] R. A. Sholihati, I. A. Sulistijono, A. Risnumawan and E. Kusumawati, "Potato Leaf Disease Classification Using Deep Learning Approach," *2020 International Electronics Symposium (IES)*, Surabaya, Indonesia, 2020, pp. 392-397, doi: 10.1109/IES50839.2020.9231784.

[2] Md. A. R. Nishad, M. A. Mitu, N. Jahan, "Predicting and Classifying Potato Leaf Disease using K-means Segmentation Techniques and Deep Learning Networks," *Procedia Computer Science*, vol. 212, 2022, pp. 220-229, ISSN: 1877-0509, doi: 10.1016/j.procs.2022.11.006.

[3] J. Pasalkar, G. Gorde, C. More, S. Memane and V. Gaikwad, "Potato Leaf Disease Detection Using Machine Learning," *Current Agriculture Research Journal*, vol. 11, 2024, pp. 949-954.

[4] M. I. Tarik, S. Akter, A. A. Mamun and A. Sattar, "Potato Disease Detection Using Machine Learning," *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, Tirunelveli, India, 2021, pp. 800-803, doi: 10.1109/ICICV50876.2021.9388606.

[5] Y. Khaparde, P. Mardikar, J. Pritam, and D. Nayak, "PLANT CHECK: POTATO LEAF DISEASE DETECTION USING CNN MODEL," *International Journal of Engineering Applied Sciences and Technology*, 2023.

[6] V. K. Shrivastava, C. J. Shelke, A. Shrivastava, S. N. Mohanty, and N. Sharma, "Optimized Deep Learning Model for Disease Prediction in Potato Leaves," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, Sep. 2023.

[7] M. Radwan, A. Alhussan, A. Ibrahim, and E.-S. El-kenawy, "Potato Leaf Disease Classification Using Optimized Machine Learning Models and Feature Selection Techniques," *Potato Research*, vol. 1-25, 2024, doi: 10.1007/s11540-024-09763-8.

[8] A. M. Elshewey, S. M. Tawfeek, A. A. Alhussan, et al., "Optimized Deep Learning for Potato Blight Detection Using the Waterwheel Plant Algorithm and Sine Cosine Algorithm," *Engineering Applications of Artificial Intelligence*, 2024.

[9] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1-11, 2016, doi: 10.1155/2016/3289801.

[10] K. Beals, "Potatoes, Nutrition and Health," *American Journal of Potato Research*, vol. 96, 2018, doi: 10.1007/s12230-018-09705-4.

[11] N. Shabrina, S. Indarti, R. Maharani, et al., "A Novel Dataset of Potato Leaf Disease in Uncontrolled Environment," *Data in Brief*, vol. 52, 2023, 109955, doi: 10.1016/j.dib.2023.109955.

[12] H. C¸ atal and V. Turk, "Potato leaf disease detection with a novel deep learning model based on depthwise separable convolution and transformer networks," *Engineering Applications of Artificial Intelligence*, vol. 133, 2024, 108307, doi: 10.1016/j.engappai.2024.108307.

[13] D. Chikmurge and S. Harnale, "Feature Extraction of DICOM Images Using Canny Edge Detection Algorithm," *Advances in Intelligent Systems and Computing*, vol. 632, Springer, Singapore, 2018, doi: 10.1007/978-981-10-5520-1_18.

[14] Y. Dahikar, D. Chikmurge, and S. Kharat, "Sketch Captioning Using LSTM and BiLSTM," *2023 International Conference on Network, Multimedia and Information Technology (NMITCON)*, Bengaluru, India, 2023, pp. 1-9, doi: 10.1109/NMITCON58196.2023.10275867.

[15] N. Shabrina, S. Indarti, R. Maharani, D. Kristiyanti, Irmawati, N. Prastomo, and T. Mutiara, "A Novel Dataset of Potato Leaf Disease in Uncontrolled Environment," *Data in Brief*, vol. 52, 2023, Art. no. 109955, doi: 10.1016/j.dib.2023.109955.

[16] H. C¸ atal and V. Turk, "Potato Leaf Disease Detection with a Novel Deep Learning Model Based on Depthwise Separable Convolution and Transformer Networks," *Engineering Applications of Artificial Intelligence*, vol. 133, pt. C, 2024, Art. no. 108307, doi: 10.1016/j.engappai.2024.108307.

[17] D. Chikmurge and S. Harnale, "Feature Extraction of DICOM Images Using Canny Edge Detection Algorithm," in *International Conference on Intelligent Computing and Applications*, S. Dash, S. Das, and B. Panigrahi, Eds. Singapore: Springer, 2018, vol. 632, Advances in Intelligent Systems and Computing, pp. 123–131, doi: 10.1007/978-981-10-5520-1_18.

[18] Y. Dahikar, D. Chikmurge, and S. Kharat, "Sketch Captioning Using LSTM and BiLSTM," in 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1–9, doi: 10.1109/NMITCON58196.2023.10275867.

[19] M. A. Putra, "Potato Leaf Disease Dataset," Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/muhammadardiputra/potato-leaf-disease-dataset. [Accessed: Dec. 13, 2024].