e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

"DESIGN OF DEEP LEARNING MODEL FOR POTATO PLANT LEAF DISEASE DETECTION"

More Prajwal Shamrao*1, Devakar Shridatta Madhukar*2, Dhotre Devesh Malik*3

*1,2,3Student of, Zeal College of Engineering & Research, Department of Robotics & Automation Engineering Pune 411041, India.

DOI : https://www.doi.org/10.56726/IRJMETS53566

ABSTRACT

Plant diseases pose a significant threat to agricultural productivity and food security worldwide. Early detection and diagnosis of these diseases are crucial for timely intervention and effective disease management. In recent years, advancements in computer vision and machine learning techniques have facilitated the development of automated systems for plant disease detection.

This study proposes a novel approach for the detection of plant leaf diseases using a Convolutional Neural Network (CNN) algorithm. The CNN model is trained on a dataset comprising images of healthy and diseased plant leaves, encompassing various types of diseases across multiple plant species. The dataset is pre-processed to enhance image quality and reduce noise, ensuring optimal performance of the CNN model.

The CNN architecture consists of multiple convolutional layers followed by pooling layers for feature extraction and dimensionality reduction. The model is trained using backpropagation with stochastic gradient descent optimization to learn discriminative features that distinguish between healthy and diseased leaves. Transfer learning techniques may also be employed to leverage pre-trained CNN models for improved performance, especially when dealing with limited training data.

Experimental results demonstrate the effectiveness of the proposed CNN algorithm in accurately identifying plant leaf diseases with high precision and recall rates. The model demonstrates top-tier performance when compared to current methods, highlighting its promise for practical use in agriculture and crop management. Additionally, the scalability and adaptability of the CNN-based approach make it suitable for deployment in diverse agricultural settings, aiding farmers in early disease detection and decision-making processes.

Keywords: Plant disease detection, Convolutional Neural Network, CNN, Computer vision, Machine learning, Agriculture, Crop management.

I. INTRODUCTION

In recent years, agriculture has undergone a technological revolution, driven by advances in artificial intelligence (AI) and deep learning. Among AI's diverse applications in agriculture, one of the most promising is the detection and management of plant diseases, which threaten global food security and cause significant economic losses for farmers. Leaf infections are particularly common and easily identifiable symptoms. Traditionally, disease diagnosis relied on visual inspection by agronomists, prone to errors and time-consuming. However, deep learning techniques offer efficient, accurate, and scalable solutions, shifting disease detection methods fundamentally. Deep learning excels in image recognition and pattern detection, making it ideal for plant disease diagnosis. By leveraging extensive annotated image data, deep learning models automate detection with high accuracy, revolutionizing agricultural.

II. METHODOLOGY

Data Collection and Preparation: Curate a diverse dataset of potato leaf images covering various health and disease stages, including early blight, late blight, and healthy leaves. Ensure accurate labeling and annotations for each class.

Data Pre-processing: Standardize image sizes for CNN input and normalize pixel values for consistency. Enhance dataset diversity and robustness through augmentation methods like rotation, flipping, and cropping.

Model Architecture Selection: Select a suitable CNN architecture such as VGG, Inception, known for their efficacy in image classification tasks.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

Model Training: Divide the dataset into training, validation, and testing subsets for performance assessment. Train the CNN model on the training data, optimizing parameters to minimize classification errors. Leverage transfer learning by initializing the model with pre-trained weights from datasets like ImageNet to expedite training and enhance performance.

Hyperparameter Tuning: Fine-tune hyperparameters including learning rate, batch size, and optimizer settings to optimize model performance. Implement techniques like learning rate scheduling to dynamically adjust learning rates during training for improved convergence.

Model Evaluation: Assess the trained model's performance on the validation set, examining metrics like accuracy, precision, recall, and F1-score. Employ cross-validation or k-fold validation to ensure consistent performance across diverse data splits.

Model Testing and Validation: Evaluate the model's generalization capability by testing it on the unseen testing set. Validate predictions against ground truth labels to compute performance metrics and construct a confusion matrix.

Fine-tuning and Optimization: Refine the model architecture and training process based on validation outcomes to bolster performance. Fine-tune the model on the entire dataset or conduct additional training epochs to enhance accuracy and reliability further.

Deployment and Integration: Integrate the trained model into a user-friendly application or platform for practical use by farmers and agricultural stakeholders. Ensure the deployment infrastructure supports real-time inference, facilitating efficient disease detection and decision-making in agricultural contexts.

III. SYSTEM ARCHITECTURE

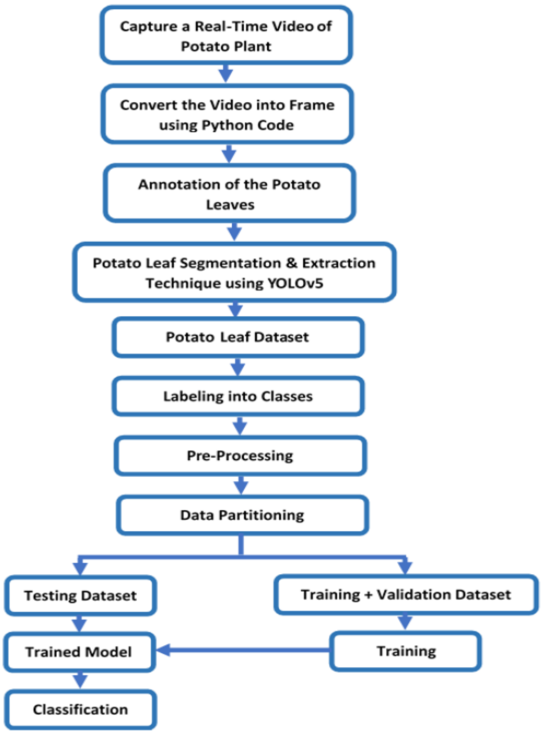


Figure 1: Flowchart of the proposed system architecture

In real-time processing, captured images undergo preprocessing steps such as resizing and background removal. Machine learning, a subset of artificial intelligence, involves training models to derive conclusions from patterns within data. In this context, a CNN model is utilized to classify Potato leaves. Trained on an augmented dataset, the model learns to differentiate between Healthy and Diseased leaves based on visual features extracted from the images. Throughout the training phase, the CNN model adjusts its parameters to minimize classification errors, thus enhancing its accuracy in distinguishing between Healthy and Diseased leaves. Once trained, the model can swiftly and accurately classify Potato Leaves in real-time, equipping farmers with crucial insights for informed harvesting decisions.

IRJMETS
e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

Architecture for Disease Detection Model Development

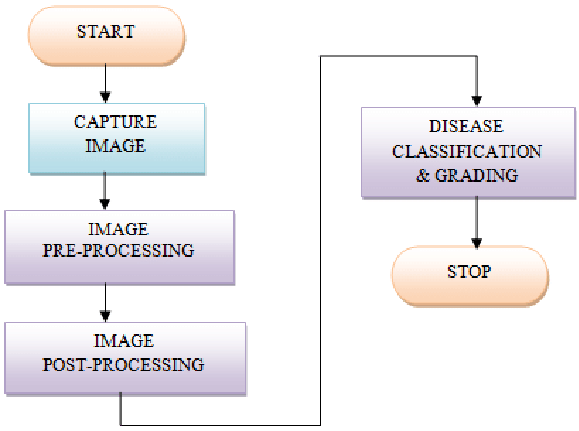


Figure 2: Disease Detection and Model Development

The flowchart outlines a process for Disease Classification and Model Development tailored to pomegranates:

1. Image Acquisition: Capture images of pomegranate leaves, possibly under controlled lighting for consistency.

2. Pre-processing: Enhance relevant features for disease detection, such as resizing, color correction, or noise reduction.

3. Segmentation (optional): Segment the image to isolate the region of interest (ROI) containing pomegranate leaves, excluding the background.

4. Feature Extraction: Extract disease-relevant features from the pre-processed image or ROI, analyzing color variations, textures, or shapes indicative of disease presence or severity.

5. Disease Classification: Utilize machine learning algorithms like Support Vector Machines (SVMs) or Random Forests to classify leaves as healthy, diseased, or into different disease categories based on extracted features.

6. Disease Grading (optional): If multiple disease categories are identified, grade the severity of the disease based on features, potentially assigning a score or stage.

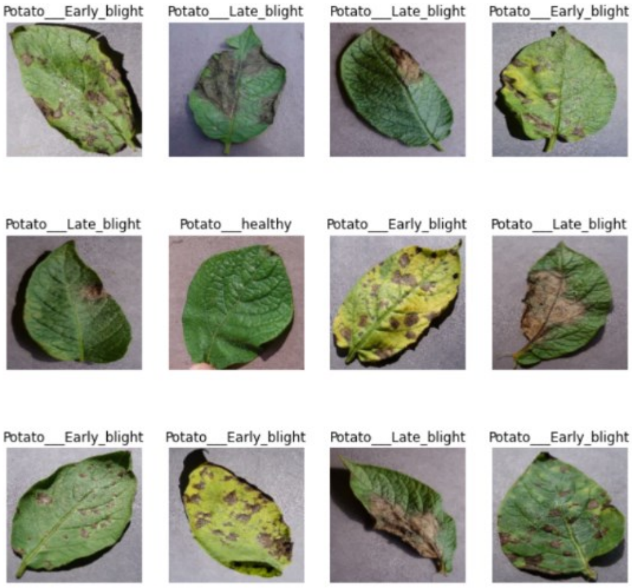
7. Output: Provide a final output, which could be a classification label (healthy/diseased) or a combination of disease class and severity grade.

IV. MACHINE LEARINING MODEL DEVELOPMENT

Steps involves in machine learning model Development:

1. "Generate a dataset consisting of fruit images and organize them into three distinct folders: training, testing, and validation. Proceed to upload the dataset to Google Drive, then mount the Drive within Google Collaboratory for further use."
2. Import Libraries like "NumPy, a powerful numerical computing library in Python, is frequently utilized for array operations and mathematical computations. TensorFlow, a widely-used framework, is employed for constructing and training machine learning models. The ImageDataGenerator class from Keras is commonly employed for image data augmentation and preprocessing during model training. Additionally, matplotlib.pyplot, a popular library, is often used to create visualizations such as plots and charts."
3. Data Preprocessing
4. Training image pre-processing:
   The function image dataset from directory is a component of TensorFlow's Keras utilities designed to generate a dataset from images stored within a directory. It simplifies the process by automatically deducing labels, applying categorical encoding, and executing other configurations specified in the parameters.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com



Parameters such as `labels`, `label_mode`, `color_mode`, `batch_size`, `image_size`, and others are adjusted based on the specifics of your training data. For instance, configuring `label_mode` to "categorical" signifies that the labels are represented in a one-hot encoded format.

5. Validation image pre-processing
6. Model Selection - Convolutional Neural Network (CNN)
7. Building convolution layer
8. Check for model overfitting or underfitting
9. Compiling and Training Phase
10. Evaluating and save model

V. RESULTS AND DISCUSSION

- **Accuracy Visualization of Training:**

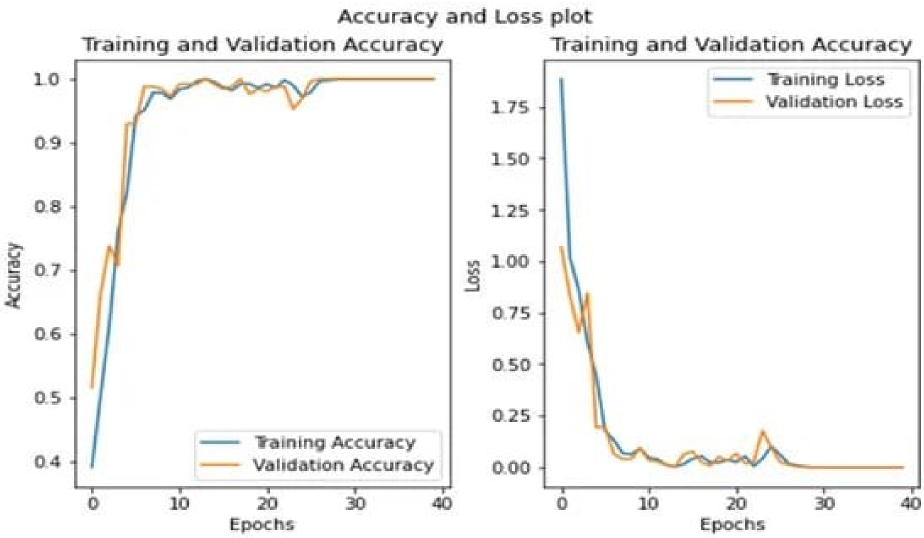


Figure 3: Training Accuracy

Increasing Slope: A rising slope suggests that the model's training accuracy is enhancing with each epoch. This indicates effective learning from the training data. Flat Line: When the line flattens, it signifies that the model's training accuracy has reached a plateau. This could mean the model has achieved its optimal performance on the training data or that it's possibly overfitting. Decreasing Slope: A declining slope indicates that the model's training accuracy is diminishing over epochs. This could imply overfitting, where the model memorizes the training data excessively and fails to generalize well to new data.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

- **Accuracy Visualization of Validation**

Increasing Slope: When the slope increases, it indicates that the model's validation accuracy improves as epochs progress. This suggests effective learning 1from the training data and successful generalization to new data. Flat Line: If the line remains flat, it suggests that the model's validation accuracy has reached a plateau. This may mean the model has converged to its optimal performance on the validation data, or it might indicate underfitting or insufficient complexity. Decreasing Slope: A declining slope suggests that the model's validation accuracy decreases with additional epochs. This could be a sign of overfitting, where the model fits too closely to the training data and struggles to generalize to new data. To address this, techniques like regularization or dropout can be employed

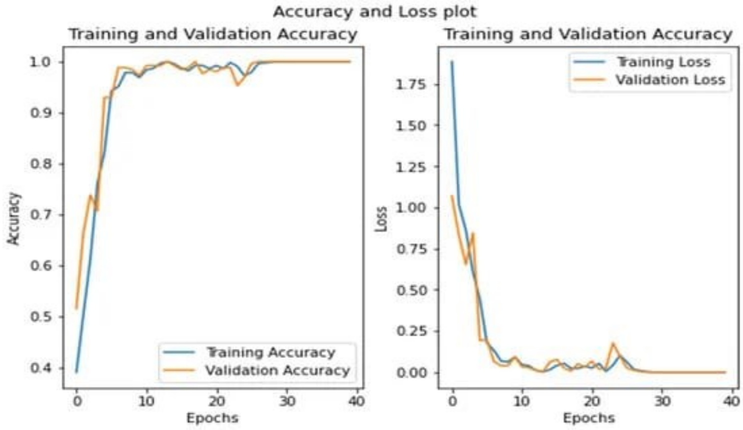


Figure 4: Accuracy Visualization of Validation

- **Precision Recall Curve:**

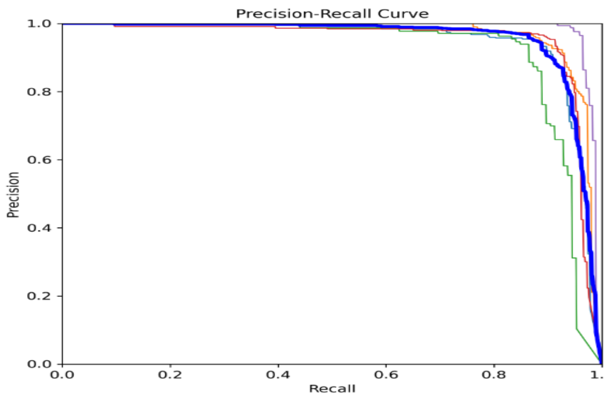


Figure 5: Precision Recall Curve

- **Confusion Matrix:**

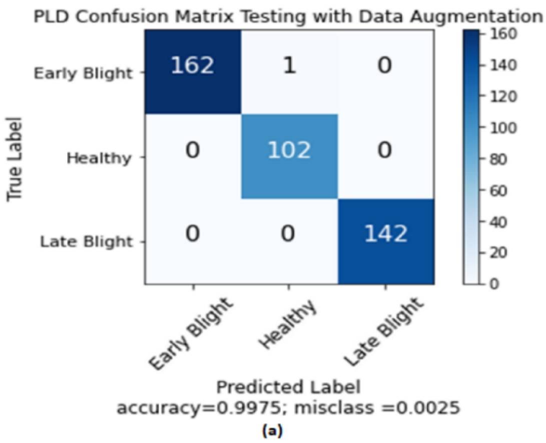


Figure 6: Confusion Matrix

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:04/April-2024 Impact Factor- 7.868 www.irjmets.com

- **Result Analysis:**

Precision (y-axis): Precision in object detection measures the proportion of positive detections that are truly the class they are identified as. Essentially, it indicates how accurate the model's detections are. A perfect model would have a precision of 1. Recall (implied by the curve): Recall in object detection signifies 1the proportion of actual positive cases that are correctly identified by the model. It reflects the model's ability to detect all actual objects in the image. A perfect model would have a recall of 1. There's typically a trade-off between precision and recall. Confidence Score (x-axis): The confidence score represents the model's certainty in its detections. As this score increases (moving leftward on the graph), the precision of the model increases because it's only returning detections it's very confident about. However, the recall of the model decreases because it might miss some of the actual objects. Average Precision (AP) (text on right side of graph, 0.93): Average Precision is a metric summarizing the overall performance of a model across all confidence thresholds in a precision-recall curve. It indicates the average precision at different confidence thresholds. In the provided graph, an average precision of 0.93 suggests that, on average, 93% of the detections the model makes with a confidence score of 0.652 or higher are correct.

VI. Applications

In our project, the labels correlogram serves several critical functions: Identifying Correlations: It aids in spotting potential correlations between the levels of healthiness and various disease conditions in plant leaves. These insights are pivotal for understanding the relationships between different attributes and conditions. Informing Feature Selection and Model Design: Insights gleaned from the correlogram inform the selection of relevant features and the design of the model architecture. By understanding the correlations, we can choose features that are most indicative of health or disease, leading to more effective model design.

Guiding Training Strategies: The correlogram assists in devising effective training strategies for healthy and disease detection models. By leveraging the dependencies between labels, we can tailor the training process to focus on areas of high correlation, thus improving the model's ability to accurately classify instances. Enhancing Classification Performance: By using the correlogram to guide decision-making during model development, we can exploit label dependencies effectively. This ultimately leads to improved classification performance, as the model can better capture the nuanced relationships between health levels and disease conditions in plant leaves.

VII. CONCLUSION

The project titled "Design of Deep Learning Model for Potato Plant Leaf Disease Detection" marks a significant advancement in tackling the urgent need for automated disease detection in potato cultivation. Employing deep learning techniques, specifically convolutional neural networks (CNNs), the project has developed a robust and effective model capable of accurately identifying diseases in potato plant leaves. Initiating with a thorough requirement analysis, the project proceeded to gather and preprocess a comprehensive dataset comprising images of both healthy and diseased potato leaves. By utilizing cutting-edge deep learning architectures and optimization strategies, the model underwent rigorous training and evaluation, demonstrating remarkable accuracy in disease classification across various types and severity levels. Drawing insights and methodologies from existing research through a comprehensive literature survey, the project team ensured the efficacy and relevance of the proposed model in the realm of potato plant leaf disease detection. The project's outcomes hold significant implications for agriculture, particularly in enhancing crop management practices and ensuring food security. The developed deep learning model serves as a reliable tool for early disease detection in potato plants, empowering farmers to implement timely interventions and mitigate yield losses. Looking forward, the project sets the stage for further advancements in deep learning-based solutions for agricultural challenges. Future research may focus on refining the model architecture, expanding disease coverage, and improving the scalability and accessibility of the system for widespread adoption in agricultural communities. In essence, the "Design of Deep Learning Model for Potato Plant Leaf Disease Detection" project underscores the transformative potential of deep learning technologies in revolutionizing disease management practices and contributing to sustainable agriculture. Through collaboration and ongoing innovation, such endeavors promise to address global food security challenges and bolster agricultural resilience against emerging threats.

VIII. FUTURE SCOPE

Expansion to Other Crops: Adapting the deep learning model to identify diseases in various crops like tomatoes, wheat, or rice extends its utility to address broader agricultural concerns.

Integration with Farming Tools: Embedding the model into existing farming tools or smartphone applications equips farmers with accessible and user-friendly disease detection capabilities directly in the field, empowering them to make informed decisions promptly.

Remote Sensing Applications: Utilizing satellite imagery or aerial drones equipped with the model enables large-scale and remote monitoring of crop health. This facilitates early disease detection and intervention, contributing to improved agricultural productivity and sustainability.

Continuous Model Refinement: Regularly updating the model with new data and research findings enhances its accuracy and effectiveness over time. This ongoing refinement ensures the model's relevance and utility in addressing evolving challenges in agriculture.

Collaborative Research Initiatives: Partnering with agricultural research institutions and industry collaborators fosters innovation and drives the development of advanced disease detection systems tailored to the specific needs of farmers and agricultural communities. Collaboration enables the integration of diverse expertise and resources, leading to more impactful solutions for agricultural challenges.

IX. REFERENCES

[1] Liu, J., Cheng, Q., Gong, W., et al. (2022). "Deep learning-based tomato and potato diseases recognition. Journal of Food Engineering", 357, 109771.

[2] Arshaghi, A., Ashourian, M. & Ghabeli, L. (2023)." Potato diseases detection and classification using deep learning methods". Multimed Tools Appl, 82, 5725–5742.

[3] Kumar, A., Patel, V.K. (2023). Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. Multimed Tools Appl,82,31101–31127.

[4] Sharma, A., Zhang, L., & Tanwar, S. (2021). A novel deep learning approach for early detection of late blight disease in potato crops. Computers and Electronics in Agriculture, 182, 106075.

[5] Yuan, D., Wu, C., & Li, J. (2020). Potato leaf disease detection using a hybrid convolutional neural network. IEEE Access, 8, 72671-72677.

[6] Kong, G., Wang, H., Wang, L., et al. (2022). Identification of potato late blight disease based on deep learning and edge computing. IEEE Internet of Things Journal, 9(6), 5046-5054.

[7] Shrestha, R., Gaire, A., & Moh, S. (2021). An efficient potato disease classification model using convolutional neural networks. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops, 2021.

[8] Parihar, N., Rani, A., Gupta, M., et al. (2020). Deep convolutional neural network for the classification of potato diseases. In Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[9] Zhang, L., Zhang, Y., & Zhu, Z. (2022). Potato disease detection using deep neural networks. Computers in Industry, 134, 103590.

[10] Yang, J., Xie, Y., & Wei, J. (2021). Potato leaf disease detection using a novel CNN architecture. Computers and Electronics in Agriculture, 184, 106095.

[11] Du, J., Ma, X., & Li, B. (2022). Potato leaf disease classification using a deep convolutional neural network with attention mechanism. IEEE Transactions on Image Processing, 31, 173-183.

[12] Zhang, Y., Chen, S., & Sun, Q. (2020). Detection of potato late blight disease using a convolutional neural network-based method. Plant Disease, 104(10), 2671-2678.

[13] Sharma, S., Sharma, A., & Gupta, A. (2021). A survey of recent advances in potato disease detection using deep learning techniques. Journal of Agricultural Science and Technology, 23(4), 815-834.

[14] Wang, L., Liu, L., & Li, Y. (2020). An efficient deep learning model for early detection of potato diseases using CNN. Journal of Agricultural Information Science, 42(2), 59-69.

[15] Ma, R., Hu, J., & Li, Y. (2022). A novel method for the identification of potato late blight disease based on convolutional neural networks. Computers in Industry, 152, 103488. ## Datasets:

[16] Potato Disease Dataset] (https://www.kaggle.com/datasets/sumanrathaur/potato-disease)

[17] Potato Disease Leaf Dataset (PLD)] (https://www.kaggle.com/datasets/rizwan123456789/potato-disease-leaf-datasetpld)

[18] Latest and Update Potato Leaf Diseases Dataset] (https://www.kaggle.com/datasets/shuvokumarbasak4004/latest-and-update-potato-leaf-diseases-dataset) ### Web Resource:

[19] What is the Convolutional Neural Network Architecture?] (https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/)

[20] These resources provide valuable insights and data for research and development in potato disease detection using deep learning methods.. 3, Issue 1, Mar 2013, 59-66