Ensuring Quality Potato Yields with Early Detection of Diseases

**Maaz Qureshi, Md Saqib, Mohd Sami Khan**

Guide : Mr. Harsh Pratap Singh

AssistantProfessor

*Department of Computer Science and Engineering*

*Medi-Caps University, Indore, India*

**Abstract**

In today's fast-paced world, technology has become an integral part of our lives, and almost everyone relies on digital tools to accomplish daily tasks. However, users and farmers still face several issues due to slow internet connectivity and the need to rely on others to manage their online tools. These challenges can significantly impact their ability to manage their crops and respond effectively to plant diseases that may occur.

One major issue that farmers face is the slow detection of diseases in their crops, which they may not know how to cure. Traditional methods of identifying plant diseases can be time-consuming and often require expert knowledge. This can be a significant challenge for farmers who may not have access to expert advice or have limited time to manage their crops effectively.

To address this issue, we have developed a mobile/web application that leverages the power of TensorFlow and a CNN model to help users detect diseases. Our application aims to provide farmers with faster and more efficient disease detection for their growing plants. The mobile/web application enables users to easily search for plant diseases and their cures by taking photos of their plants and accessing basic functions such as disease identification.

By using our application, users can quickly detect plant diseases. The application also allows users to store data on their crops and monitor their progress, helping farmers make more informed decisions about managing their crops.

Overall, our project aims to empower users and farmers with the tools they need to manage their crops effectively and efficiently. With our application, farmers can detect plant diseases early and treat them promptly, leading to healthier plants and better yields. Ultimately, we believe

that our application can make a significant contribution to sustainable agriculture and help farmers overcome the challenges they face in today's technology-driven world.

# Introduction

Agriculture plays a vital role in reducing poverty and promoting economic growth. It is important to understand the significance of agriculture in reducing poverty and promoting economic growth. Food production plays a vital role in ensuring food security and preventing malnutrition. Rural populations, who are largely involved in agriculture, are particularly vulnerable to food insecurity. According to the World Bank, approximately 80% of people in rural areas are involved in farming.[1]

The potato is a widely popular and successful vegetable crop that is grown in gardens around the world. It is typically cultivated as a winter crop between the months of October and March. In Bangladesh, it is the third most important crop for enhancing economic growth, following rice and wheat. Bangladesh is also the fourth-largest producer of potatoes in Asia and ranks seventh in the world for potato production.[2]

Potato plants are vulnerable to various diseases during the harvesting period. Thus, early detection of the conditions in potato fields and prompt treatment can be an effective solution to boost potato production, which was the primary objective of this study. Traditional machine learning algorithms have been widely used to classify potato leaf diseases. In this study, we aimed to improve classification results by combining segmentation techniques and deep learning algorithms. Image segmentation was employed to mask potato leaf images, resulting in a better image dataset. Image segmentation is a technique commonly used to identify objects and boundaries in pictures. It involves assigning a label to each pixel in an image, such that pixels with the same label share specific characteristics. Various algorithms can be used for image segmentation, such as Otsu's Binary threshold algorithm, Contour Detection, and K-means clustering Algorithm. K-means algorithm is one of the most popular segmentation algorithms, which clusters objects based on their similarity and proximity to the nearest neighbors. The Euclidean distance is used to calculate the similarity distance of K-means algorithm, and the value of K usually ranges from 2 to 10.[3]

Various conventional machine learning techniques were commonly utilized for both identifying plant diseases and executing a range of computer vision tasks. Md. Asif Iqbal and Kamrul Hasan Talukder put forward a method that employed seven widely-used conventional machine learning algorithms. In their research, they employed image segmentation techniques on 450 images gathered from the Plant Village dataset. Out of the seven algorithms, random forest was found to be the best model with an accuracy of 97%.[4]

Chaojun Hou et al. conducted a study that investigated different machine learning algorithms and utilized graph cut segmentation techniques to predict early and late blight diseases on potato leaves. Their research resulted in a 91% accuracy rate achieved by an SVM classifier after image segmentation was performed. [5]

In contrast, deep learning algorithms are now considered suitable for enhancing the accuracy of performance. A range of deep learning algorithms are available to conduct various experiments on diverse agricultural products, including rice, tomato, bell pepper, and potato. [6]

Image segmentation techniques are known to improve the accuracy of machine learning algorithms in various agricultural applications. This study aims to combine image segmentation techniques with deep learning algorithms to predict potato leaf disease with better performance results.

# Literature Review

Numerous studies have been conducted on the progress of agriculture, as it can not only contribute to economic growth but also promote a healthier environment for humans. The application of deep learning models and computer vision-based techniques has received significant attention in accelerating crop production. This section presents a comprehensive overview of previous research on this topic.

A research article entitled "Krishi Mitra: Using Machine Learning to Identify Diseases in Plants" utilized the CNN model methodology through the TensorFlow framework to execute their project. This model's benefit was that it could detect fungal diseases in sugarcane by only measuring the leaf area. However, the drawback was that it necessitated a high level of

computational complexity to implement.

In a research paper titled "Severity Identification of Potato Late Blight Disease from Crop Images Captured under Uncontrolled Environment," researchers utilized Fuzzy c-mean clustering and Neural Network to create the model. The model's main advantage was that it did not require special training for farmers, as it contained images captured from various angles. However, the disadvantage was that images captured by untrained farmers were not properly oriented and contained clusters of leaves with the background visible in several segments.

The research article titled "Potato Disease Detection Using Machine Learning" utilized image processing technology. The CNN model was used, which was a major advantage of the project as it achieved a validation accuracy of 90%. However, a significant drawback of this model was the requirement for a large training dataset.

# Problem Statement

Potato farming is a significant source of income for farmers worldwide. However, the crop is susceptible to various diseases, leading to significant financial losses each year. Early Blight and Late Blight are the most common diseases that affect potato plants, and they can cause devastating economic losses if not managed correctly.

Early Blight is a fungal infection that causes brown spots on leaves, which can then spread to the rest of the plant. The infection can also cause the leaves to yellow and die, leading to a significant reduction in potato yields. Late Blight, on the other hand, is caused by micro-organisms that infect the leaves and stems, causing them to turn brown and eventually collapse.

Timely detection and appropriate treatment are critical to prevent significant waste and economic losses in potato farming. However, identifying the type of disease affecting the potato plant can be challenging, and the treatments for Early Blight and Late Blight differ slightly. Therefore, accurate disease identification is crucial to avoid mismanagement and ensure effective treatments.

To address this issue, Convolutional Neural Network-Deep Learning technology is being utilized to detect potato plant diseases accurately. Convolutional Neural Networks (CNN) are a type of deep learning algorithm that can recognize patterns and features in digital images. These networks can be trained using large datasets of images to accurately classify images into different categories, such as healthy or diseased plants.

The CNN model is trained on a large dataset of potato plant images, both healthy and diseased. When a farmer submits an image of a diseased plant to the system, the CNN algorithm processes the image, extracts the relevant features, and compares them to the pre-existing dataset. This comparison allows the algorithm to determine the type of disease affecting the plant accurately.

By utilizing CNN-Deep Learning technology, farmers can detect potato plant diseases accurately and efficiently. This technology can help farmers make more informed decisions about the management of their crops, leading to better yields and significant economic gains. By preventing the spread of disease, farmers can reduce the need for costly treatments, which can be harmful to the environment and human health.

In conclusion, the use of CNN-Deep Learning technology in potato farming is a significant step forward in the management of plant diseases. This technology enables farmers to detect and treat diseases accurately, reducing economic losses and environmental harm. The use of advanced technology in agriculture can contribute to sustainable farming practices and improve the livelihoods of farmers worldwide.

# Methodology

# The project involves various research phases, as illustrated in Figure 1 within the research framework. The proposed research framework consists of distinct stages, which are outlined below.

# 

# Fig. 1 : Data Flow Diagram

**Data Collection**

Developing accurate leaf classification and disease detection algorithms is critical in the field of agriculture to help farmers identify and manage crop diseases quickly and effectively. In this research, two distinct datasets, namely Plant Village and Mendeley, were utilized to train and evaluate a model for classifying potato plant leaves into three categories: early blight, late blight, and healthy leaves.

The inclusion of a healthy leaf class in the dataset is essential to ensure that the model can accurately distinguish between healthy leaves and leaves affected by early blight or late blight. The dataset was divided into an 80:20 ratio for training and testing the models, with 80% of the dataset used for training and 20% used for testing.

The proposed network was evaluated on its ability to accurately classify potato plant leaves into the three classes. The results showed that the network was effective in distinguishing between the different classes, resulting in decent performance. This is a promising result as it indicates that the proposed network can potentially be used as a tool to assist farmers in identifying and managing potato plant diseases.

The use of multiple datasets in this research is a significant strength, as it helps to improve the accuracy of the model by providing a more comprehensive range of data. Additionally, the division of the dataset into training and testing sets ensures that the proposed network is evaluated objectively and can generalize to new, unseen data.

Overall, this research demonstrates the potential of using machine learning techniques to develop accurate leaf classification and disease detection algorithms for potato plant diseases. The proposed network has shown promising results in distinguishing between early blight, late blight, and healthy leaves, indicating that it could be a valuable tool for farmers in identifying and managing potato plant diseases.

The dataset was taken from Kaggle website under the name “PlantVillage Dataset”.

|  |  |
| --- | --- |
| **Samples** | **Number** |
| Healthy leaf | 152 |
| Early blight | 1000 |
| Late blight | 1000 |
| **Total** | 2152 |

# Augmentation

Data augmentation is a technique used to improve the accuracy of image classifiers, and in this study, several augmentation methods were used to enhance the model's performance and recognition capabilities. One method involved applying a shear range of 0.2 to training images with one axis fixed and the other stretched to a predetermined angle. This helped to correct perception angles and produce images from different perspectives. Zooming with a range of 0.2x and horizontal flipping were also used to further enhance the dataset. These methods generated a total of 10,320 images, with 8,256 used for training and 2,064 used for testing the model.

# Classification

The Convolutional Neural Network (CNN) architecture is a popular choice for image classification tasks, including identifying diseases in potato leaves. This architecture is a supervised learning technique that utilizes an existing dataset to train a model to recognize images based on their attributes. In this project, we will use CNN to classify potato leaf images into healthy leaves, early blight, and late blight.

The CNN architecture is particularly useful for recognizing potato leaves based on their attributes. The convolutional layer in CNN applies a filter to the leaf image to identify features and patterns that can be used to distinguish between healthy leaves and leaves affected by early blight or late blight. The output image is then passed through a pooling layer, which reduces the image's resolution while preserving its quality.

In this project, the leaf images will be resized to 150x150x3 pixels, which means that each image will have three channels (red, green, and blue). The output image resulting from pooling will then undergo MaxPooling, which reduces the spatial size of the output while retaining the most significant information.

To enable the CNN to classify the images, the output of the MaxPooling layer will be flattened into vector form. The proposed model for identifying diseases in potato leaves will use four convolutional layers and four MaxPooling layers. The CNN architecture will be trained on the dataset of potato leaf images to learn the features that distinguish healthy leaves from those affected by early blight or late blight.

Overall, the proposed CNN architecture is a promising approach to identifying potato leaf diseases, as it utilizes the powerful capabilities of CNN to recognize patterns and features in potato leaf images. By training the model on a dataset of potato leaf images, the CNN can learn to accurately classify potato leaves as healthy or affected by early blight or late blight.

Fig. 2 : Healthy Leaves



Fig. 3 : Early Blight





Fig. 4 : Late Blight





# Experimental Setup and Result Analysis

# Plant leaf diseases are a major concern for farmers worldwide as they can cause significant losses in crop productivity and quality. In recent years, deep learning techniques have emerged as a promising approach for detecting and diagnosing plant leaf diseases. These techniques can help farmers to control biotic variables that cause severe crop yield losses, thereby improving overall crop productivity and quality.

# Our project focuses on the detection of potato leaf diseases using a multi-level deep learning model. A Convolutional Neural Network (CNN) is a deep learning algorithm used for analyzing visual data. It learns and extracts features directly from raw input data by applying convolutional filters, using activation functions for non-linearity, down sampling with pooling layers, utilizing fully connected layers, and producing output probabilities for classification. CNNs are trained using backpropagation and gradient descent to optimize the network's weights. By employing these operations, CNNs can learn hierarchical representations of visual data, making them effective for tasks like image classification, object detection, and image segmentation.

# The proposed model is fast and straightforward and can classify different types of potato leaf diseases. At the first level, the model takes potato leaf images as input and extracts potato leaves from the images. At the second level, a convolutional neural network is developed for potato leaf disease detection, which can classify early blight and late blight potato diseases based on the potato leaf images.

# In addition, the proposed model considers the effects of environmental factors such as temperature, humidity, and light on potato leaf diseases. By taking into account these factors, the model can better predict and diagnose potato leaf diseases. This can help farmers to take appropriate measures to control the spread of diseases and improve crop productivity and quality. Overall, our project demonstrates the effectiveness of deep learning techniques for plant leaf disease detection and diagnosis, particularly in the context of potato crops.

# In this statement, it is being conveyed that the CNN techniques proposed for detecting potato leaf disease were tested on a dataset that was different from the one it was initially trained on. The results showed that the proposed approach outperformed other methods that were being used for the same purpose.

# The technique was trained on two versions of the PLD dataset - one with data augmentation techniques and the other without. Despite this, it achieved an accuracy of 96.41% and showed high precision, recall, F1-score, and ROC curve metrics.

# What's interesting is that despite being simpler and having fewer parameters than existing state-of-the-art methods, the proposed CNN technique demonstrated superior performance, which resulted in significant savings in computational costs and speed.

# Lastly, the results obtained from this study were compared with existing studies on potato leaf disease detection for comparison purposes. Overall, the statement highlights the effectiveness of the proposed CNN technique for detecting potato leaf disease and its potential to contribute to the development of more efficient and cost-effective methods for crop disease detection.

# Fig. 5 : Training and Validation

# Actual : Potato\_\_\_Early\_blight,

# Predicted : Potato\_\_\_Early\_blight,

# Confidence : 97.79%

# 

# 

# (a)

# Actual : Potato\_\_\_Late\_blight,

# Predicted : Potato\_\_\_Late\_blight,

# Confidence : 96.52%

# 

# (b)

# Actual : Potato\_\_\_Late\_blight,

# Predicted : Potato\_\_\_Late\_blight,

# Confidence : 89.05%

# 

# (c)

# Fig. 6 : Outputs

# Comparison

|  |  |  |
| --- | --- | --- |
| **Model** | **Dataset** | **ACcuracy** |
| VGG16 (K-Means Clustering) | PlantVillage | 97% |
| SVM (K-Means Clustering) | PlantVillage | 95.99% |
| NOVEL CNN | PlantVillage | 99.75% |
| CNN (Proposed) | PlantVillage | 96.41% |

# Discussion

The results of this project demonstrate the potential of using deep learning algorithms for the detection of potato diseases. The high accuracy of the system suggests that it could be a valuable tool for farmers and researchers who need to quickly and accurately identify diseases in potato plants.

One of the advantages of using deep learning algorithms for disease detection is that they can be trained on large datasets and can learn complex patterns in the images. This allows them to be more accurate than traditional methods of disease detection, which rely on human expertise and visual inspection.

However, there are also some limitations to using deep learning algorithms for disease detection. One of the main challenges is the need for large and diverse datasets to train the models. Collecting and labeling large datasets can be time-consuming and expensive, which may limit the scalability of these systems. In addition, deep learning algorithms can be susceptible to overfitting, where they become too specialized to the training dataset and do not generalize well to new data.

In conclusion, the results of this project suggest that deep learning algorithms have the potential to be an effective tool for the detection of potato diseases. Further research is needed to address the limitations of these systems and to explore their applicability in different agricultural settings.

# Conclusion

# Potato farming is a critical source of income for many farmers in India. However, potato plants are susceptible to various diseases, leading to significant economic losses each year. Late blight and early blight are two of the most common potato plant diseases that affect potato yield and quality. Unfortunately, many farmers lack knowledge about potato diseases and have limited access to the latest technology to manage them.

# To address this issue, a deep learning-based approach using a convolutional neural network (CNN) was proposed to detect late blight, early blight, and healthy leaf images of potato plants. This method can help farmers identify the type of disease affecting their plants quickly and accurately, allowing for timely treatment and reducing economic losses.

# The proposed approach utilizes CNN technology to classify potato plant images into three categories: late blight, early blight, and healthy leaves. The dataset used for training the model was prepared using three different augmentation techniques, which helped increase the number of training samples and improve the accuracy of the model. Additionally, k-means clustering segmentation was applied with three different values during the experiment to improve the accuracy of disease detection. The technique was trained on two versions of the PLD dataset - one with data augmentation techniques and the other without. Despite this, it achieved an accuracy of 91.41% and showed high precision, recall, F1-score, and ROC curve metrics.

# The proposed approach has the potential to revolutionize the way potato diseases are detected and managed in India. By providing farmers with an accurate and efficient way of detecting potato diseases, the proposed method can help reduce economic losses and improve crop yields. This technology can also be used to educate farmers about potato diseases, helping them make more informed decisions about their crops.

# Furthermore, the use of deep learning-based approaches in agriculture has significant potential to improve crop management and increase crop yields globally. This project is an excellent example of how technology can be used to address critical issues facing the agriculture sector, particularly in developing countries such as India.

# In conclusion, the proposed deep learning-based approach using a CNN to detect late blight, early blight, and healthy leaf images of potato plants could have a significant impact on potato growers in India. This technology can help farmers detect and manage potato diseases efficiently, reducing economic losses and improving crop yields. The use of advanced technology in agriculture has the potential to revolutionize the sector and contribute to sustainable farming practices, ultimately improving the livelihoods of farmers worldwide.

# References

1. Nugroho N. Sunjoyo (2022) “Agriculture and Food.” https:/[/www.worldbank.org/en/topi](http://www.worldbank.org/en/topic/agric)c[/agric](http://www.worldbank.org/en/topic/agric) ulture. Accessed 11 Aug. 2022.
2. Bangladesh - Potatoes Production. https:/[/www.nationmaster.com/nmx/tim](http://www.nationmaster.com/nmx/time)e series/bangladesh-potatoes-production-fao. Accessed 24 Jun. 2022.
3. Zheng, Xin, et al. (2018) “Image segmentation based on adaptive K-means algorithm.” EURASIP Journal on Image and Video Processing 2018.1: 1-10.
4. Iqbal, Md Asif, and Kamrul Hasan Talukder. (2020) “Detection of potato disease using image segmentation and machine learning.” International Conference on Wireless Communications Signal Processing
5. Hou, Chaojun, et al. (2021) “Recognition of early blight and late blight diseases on potato leaves based on graph cut segmentation.” Journal of Agriculture and Food Research 5: 100154.
6. Lamba, Monika, Gigras, Yogita and Dhull, Anuradha. (2021) “Classification of plant diseases using machine and deep learning.” Open Computer Science, vol. 11, no. 1, pp. 491-508. https://doi.org/10.1515/comp-2020-0122.
7. Resistance Evaluation for Native Potato Accessions against Late Blight Disease and Potato Cyst Nematodes by Molecular Markers and Phenotypic Screening in India.
8. Automated recognition of optical image based potato leaf blight diseases using deep learning by Kulendu Kashyap Chakraborty , Rashmi Mukherjee , Chandan Chakroborty , Kangkana Bora.
9. An effective CNN and Transformer complementary network for medical image segmentation
10. VGG-ICNN: A Lightweight CNN model for crop disease identification
11. A survey on disease detection of a potato leaf using cnn 2022
12. Recent Advances in Deep Learning CNN Models for Plant
13. https://[www.tensorflow.org/](http://www.tensorflow.org/)
14. https://[www.tensorflow.org/tfx/guide/serving](http://www.tensorflow.org/tfx/guide/serving)
15. https://[www.youtube.com/results?search\_query=tf+serving](http://www.youtube.com/results?search_query=tf%2Bserving)