**Artificial Intelligence Project**

*Potato Plant Leaf Disease Detection Using Deep Learning Method*



|  |  |
| --- | --- |
| Name | Hooria Attas, Uzair Rauf |
| Email | [hooria.gaba129@gmail.com](mailto:hooria.gaba129@gmail.com), [raufuzair87@gmail.com](mailto:raufuzair87@gmail.com) |
| Batch & Section | BSCS-02 (A) |
| Instructor’s Name | Tayyaba Arshad |
| Instructor’s Email | [tayyaba.arshad@ist.edu.pk](mailto:tayyaba.arshad@ist.edu.pk) |

# **Abstract:**

Agriculture is one of the essential sectors for the survival of humankind. Farmers face economic losses as plants are affected by diseases, especially those farmers who harvest potatoes. In order to avoid financial losses, it’s essential to detect the disease in its early stages. Technology should be implemented in the agriculture field to assist both farmers and consumers as well. This research focuses on implementing technology to increase crop yields and minimize losses.

We offer a methodology based on the Plant Village Dataset, a publicly accessible, standardized, and trustworthy dataset, for the identification and categorization of diseases affecting potato plants. Two main diseases are in focus: late blight (produced by a tiny bacterium) and early blight (caused by fungus).We used TensorFlow for data cleaning, preprocessing, and augmentation before using Convolutional Neural Networks (CNNs) for image processing and model training. The objective is to accurately detect diseases by building a deep learning model, allowing prompt and suitable actions to save crops and enhance agricultural results in Pakistan.

**Table of Contents**

[Abstract: 2](#_Toc168271034)

[Introduction: 4](#_Toc168271035)

[Literature Review: 4](#_Toc168271036)

[Data Understanding and Preprocessing: 5](#_Toc168271037)

[Modeling and Analysis: 5](#_Toc168271038)

[Results: 6](#_Toc168271039)

[Conclusion: 8](#_Toc168271040)

[References 9](#_Toc168271041)

# **Introduction:**

Pakistan's economy is mostly dependenton the agricultural industry. Farmers face economic losses as plants are affected by diseases, especially those who harvest potatoes. Early diagnosis of these diseases is necessary to prevent financial loss.One of the most consuming crops in Pakistan is potato, making it country’s significant crop.The production of potato crops is at risk due to these diseases like early and late blight and by detecting these diseases early would help farmers to protect crops. While weather patterns are out of human control and animal damage can be mitigated by field protection, managing infections is a crucial issue that has an immediate impact on crop development and output.

The principal aim of this study is to investigate the problem of disease detection in potato plants by utilizing a deep learning model. We want to create a reliable system that can detect potato plant diseases early on by utilizing Convolutional Neural Networks (CNNs). This would help farmers protect their crops by enabling them to take prompt and appropriate action, including applying targeted fertilizers.To accomplish this, we used a test dataset to assess the CNN model's performance after it had been trained on a sizable dataset of photos of potato plants. Throughout the project, TensorFlow's sophisticated preprocessing and data handling features were used to guarantee effective data management. The goal of this research is to give farmers a useful tool to improve crop protection and lower financial losses brought on by plant diseases.

# **Literature Review:**

Monzurul Islam and colleagues attempted to diagnose root related problems in potato plants. To address this, Hall et al. (2017) proposed an automated approach utilizing machine learning and image processing. Their code uses multiclass Support Vector Machine (SVM) for segmenting an classifying the images in the Plant Village data to almost 95% accuracy. Identifying diseases early using this methodology provides an effective and scalable solution enabling farmers to prevent losses from diseases. They explain very well the use of multiclass SVM based on its combination of the power to solve important classification tasks with the strength and fast processing to handle large datasets. Another work for diagnosing grape leaf diseases was done by Harshal Waghmare, Radha Kokare (2016). Additionally, their method implemented multiclass SVM for classification and image segments were obtained using high-pass filters. This is a study which basically concentrates on studying the texture patterns of the leaves, so that the diseases can be detected accurately and categorized correctly. This resulted in a 96.6% high accuracy rate, showing that similar methods can be applied to various agriculture crops like the potatoes. Another study by Shima Ramesh and Mr. Ramachandra Hebbar (2018) was conducted to classify healthy and diseased papaya leaves using a combination of Random Forest classifiers. While the accuracy of about 70% was lower for their model, the team demonstrated the significance feature extraction techniques such as Histogram of Oriented Gradients[10] (HOG). This method helps in improving classification by providing additional features which help to better distinguish in between classes. They proposed that using more local and global features can increase the accuracy even more. Furthermore, Shruthi U et al. In an extensive assessment of different Machine Learning algorithms for plant disease classification, Barbedo et al. (2019) reported higher accuracy of the classification of Convolutional Neural Networks (CNNs) The study showed that CNNs could handle tough image classification challenges with some flexibility and ease. CNNs have been found to be very successful knowledge bases due to their capacity to gather appropriate features from imagery data automatically, — making them an excellent algorithm to apply to agriculture applications because identification of diseases from crop imagery needs to be best done accurate. The literature highlights the imperative nature of making use of machine learning as well as image processing techniques in agricultural applications. As an example, Monzurul Islam et al. which may help in achieving high precision rates in disease detection [20, 21] as well as demonstrated by Harshal Waghmare et al. and Radha Kokare et al. These methods build a structure for advanced automated systems that can rapidly help farmers to identify diseases on time and minimize their crop loss as well as increase their yield. Furthermore, the inclusion of the HOG feature extraction technique, that has been previously utilized by Shima Ramesh and Mr. Ramachandra Hebbar, indicated the requirement for additional preprocessing steps to improve classification results. We could improve the basic features+cossim approach by including more complicated features, thus achieve a higher accuracy. As presented by Shruthi U et al, the CNN has indeed been a major breakthrough in this field. High accuracy of CNNs, as well as their ability to diagnose many diseases have allowed to use this algorithm as universal for the plant disease diagnosis. Their capacity to handle large datasets and learn pertinent features without requiring too much manual intervention makes them great candidates for creating reliable agricultural applications. Discussion Overall, the literature exposes a rising deployment of higher-order association methods for plant disease classification. Finally, the combined use of image processing and machine learning, especially multiclass SVM in conjugation with CNNs, has shown a good potential to automatically enhance the precision and speed of disease detection systems. The results of these studies could be used as a foundation upon which future research could be conducted aimed at developing practical tools for farmers that would increase the sustainability and productivity of practices in agriculture.

# 

# **Data UnderstandingandPreprocessing:**

* **Data Collection:**

The dataset“Plant Village” [5] used in this study is taken from Kaggle. This dataset contains the images of plant leaves that has three categorized as healthy, light blight, or early blight. The tf.data.Dataset API is used to load the 2152 images in the dataset into a TensorFlow dataset. Data loading becomes easier by this training process.

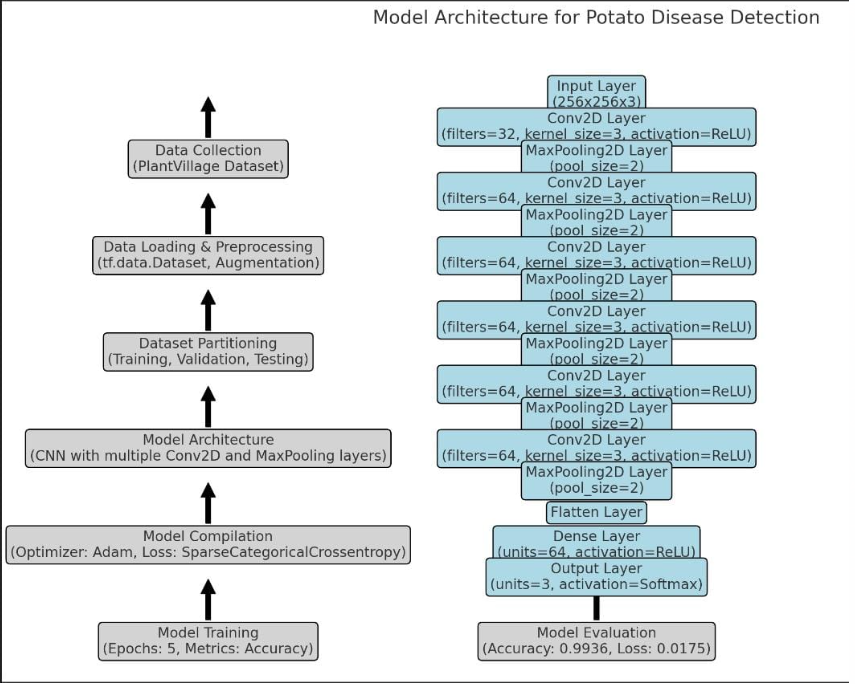
* **Data Preprocessing:**

For model training, the crucial step is data preprocessing in getting the dataset ready for model training. The dataset were divided into three categories of training, validation, and test sets 80%, 10%, and 10%, respectively. This guarantees a solid assessment of the model's functionality. Shrinking and rescaling are preprocessing techniques that were used to improve the dataset. By resizing we are ensuring that every image has the same size, resizing makes it easier for the neural network to receive the images. By normalizing the pixel values to a range of 0 to 1, rescaling enhances the model's performance and training efficiency. Data augmentation techniques were also implemented to the training set to boost the dataset’s diversity. The training images were varied using techniques like random flipping horizontally and vertically and rotation, which improved the model's ability to generalize to new data.

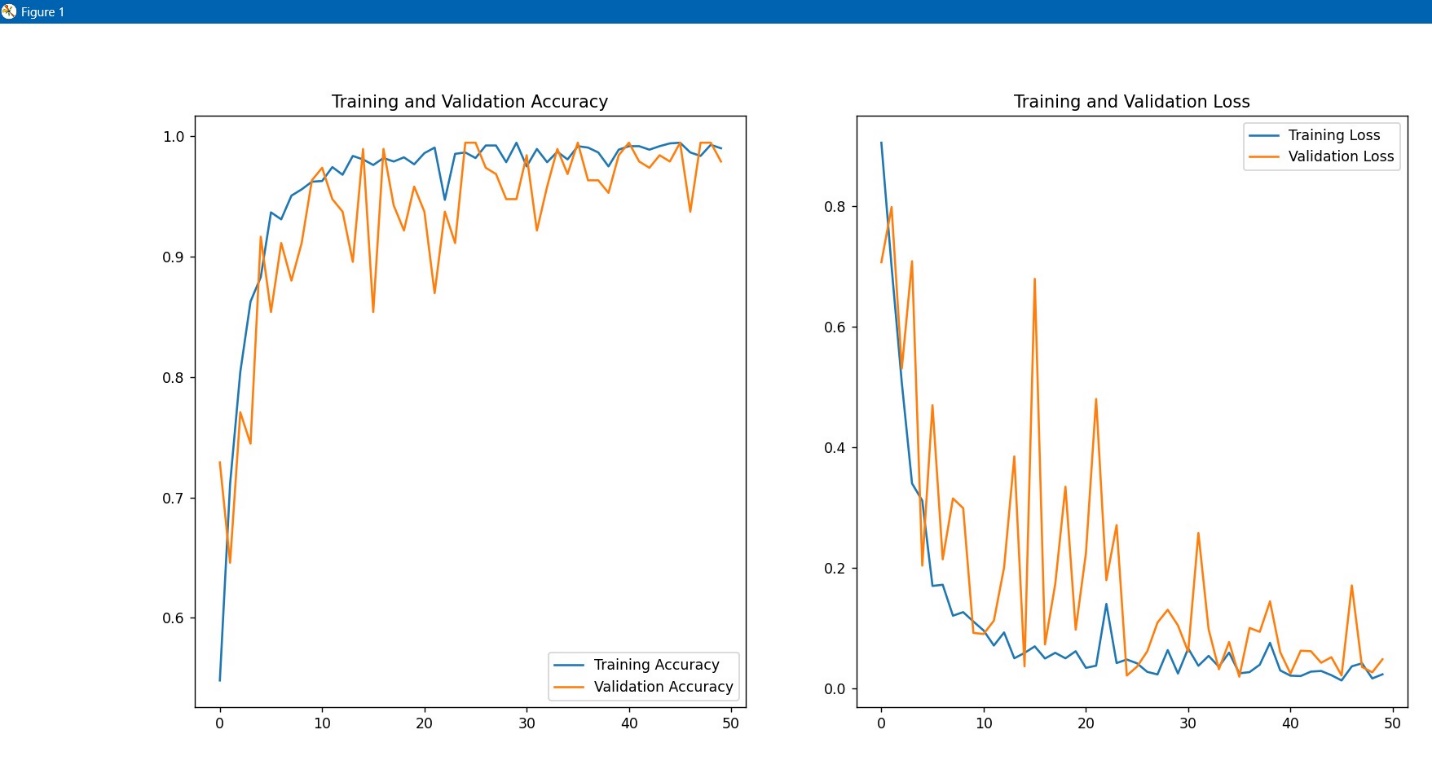
# **Modeling and Analysis:**

To categorize the leaves of the plants as healthy or with disease, a Convolutional Neural Network (CNN) was built. Multiple convolutional layers preceded each max-pooling layer in the CNN architecture. These layers facilitate the extraction and down sampling of significant image characteristics. The finallayers of the networkconsists of fully linked layers through which probabilities for each class were produced. For multi-class classification tasks, the Adam optimizer and the Sparse Categorical Cross entropy loss function were used in the model's compilation. The main performance indicator used to monitor the model's performance was accuracy. Fitting the model to the training dataset and assessing its performance on the validation set comprised the training phase. To guarantee appropriate learning and identify any indications of overfitting, the training and validation accuracies and losses were tracked over epochs. The test dataset is used for testing the model to determine its performance after training. Important parameters such as loss and precision were computed. The training and validation accuracies and losses over the epochs have been plotted to visualize the model’s performance. On the test images, predictions were produced, and the anticipated labels were compared to the actual labels to further assess the model’s performance. After completion of training, the model was saved for further deployment and for later usage as required. For tracking and management of various model iterations, the model versioning was implemented. The approach of model saving makes it easier to retrieve and reproduce for use in future studies or applications.

# **Modeling and Analysis:**



# **Results:**



A black screen with white text

Description automatically generated

A close-up of several leaves

Description automatically generated

# **Conclusion:**

The implementation of digitization is expected to yield considerable benefits for the agricultural business as it continues to permeate numerous industries. The research illustrates the importance of implanting digital technologies into agriculture to improve crop growth, and toavoid the financial losses. The main objective of model is to distinguish between affected potato leave and unaffected potato leaves. Our model shows encouraging results with a low loss of 0.0175 and an obtained accuracy of 99.36%. Still, there's potential for more accuracy development.Subsequent investigations may examine the application of artificial neural networks, namely convolutional neural networks (CNNs), as a means of improving precision. CNN approaches have the ability to provide more accurate and dependable results for image-related research. CNN architectures are optimized for maximum accuracy through the use of key architectural elements such fully connected layers, convolutional layers, batch normalizations, and activation functions. Through the utilization of deep learning methodology and digital technology breakthroughs, the agricultural industry may achieve increased production, sustainability, and efficiency in crop management practices. This will ultimately contribute to both economic success and food security.

.

# **References**

|  |  |
| --- | --- |
| [1] | V. N. a. B. K. R. U. Shruthi, " Review on Machine Learning Classification Techniques for Plant Disease Detection," *5th Int. Conf. Adv. Comput. Commun. Syst.,* 2019. |
| [2] | A. D. K. W. a. P. B. M. Islam, "Detection of Potato Disease using Image segmentation and multiclass support vector machine," *Can.Conf.Electr.Comput,* 2017. |
| [3] | R. K. a. Y. D. H. Waghmare, "Detection and classification of diseases of Grape plant using opposite color Local Binary Pattern feature and machine learning for automated Decision Support System," *Conf. Signal Process. Integ,* 2016. |
| [4] | S. R. M. e. al, "Plant disease detection using machine learning," 2018. |