Predicting Potato Leaf Diseases with Convolutional Neural Networks

BY

Md. Arafat Islam
ID: 192-15-2875

This Report Presented in Partial Fulfillment of the Requirements for
The Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Zakia Sultana

Sr. Lecturer

Department of CSE

Daffodil International University

Co- Supervised By

Nadira Anjum Nipa

Lecturer

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

July 2023

APPROVAL

This Project titled **"Predicting Potato Leaf Diseases with Convolutional Neural Networks"**, submitted by Md. Arafat Islam to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering (BSc) and approved as to its style and contents. The presentation has been held on **02/08/2023**.

BOARD OF EXAMINERS

______________________________ Chairman

Dr. Sheak Rashed Haider Noori (SRH)
Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

______________________________ Internal Examiner

Dr. Md. Zahid Hasan (ZH)
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

______________________________ Internal Examiner

Ms. Lamia Rukhsara (LR)
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

______________________________ External Examiner

Ahmed Wasif Reza (DWR)
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

We hereby declare that, this thesis has been done by us under the supervision of **Zakia Sultana, Sr. Lecturer, Department of CSE** Daffodil International University.

We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Zakia Sultana
Sr. Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Nadira Anjum Nipa
Lecturer
Department of CSE
Daffodil International University

Submitted by:

Md. Arafat Islam
ID: 192-15-2875
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final thesis successfully.

We really grateful and wish our profound our indebtedness to **Touhid Bhuiyan,** Professor and Head, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "image processing" to carry out this thesis. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this thesis.

We would like to express our heartiest gratitude to **Dr. Touhid Bhuiyan, Professor and Head, Department of CSE**, for his kind help to finish our thesis and also to other faculty members and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and passion of our parents.

ABSTRACT

Early disease diagnosis in potato leaves is complicated by the wide range of crop types, agricultural disease signs, and environmental factors involved. These problems make it difficult to detect diseases in potato leaves at an early stage. For the purpose of identifying diseases in potato leaves, a number of machine-learning techniques have been developed. The models used to detect crop species and agricultural illnesses are only tested on photographs of plant leaves from a specific geographical area, limiting the effectiveness of existing methodologies. The farmer can prevent severe financial losses by promptly detecting and controlling such outbreaks. The results of this study contribute to a unique approach that makes use of image processing to accurately detect illnesses in potato leaf populations. There are several machine learning methods for spotting symptoms of disease in potato leaves pictures; here, we employ the Convolutional Neural Network (CNN) model. The goal of this research is to develop a convolutional neural network (CNN)-based sequential model for disease prediction in potato leaves. This study's model accuracy was 92.58%. The presented model was tested on both typical and deformed potato leaves, with mixed results. Next, the algorithm is applied to the images, and the potato tree's leaves are classified as healthy or unhealthy.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Effective disease control and the prevention of significant financial losses in potato farming depend on the early identification of infections in potato leaves. However, early disease detection in potato leaves is difficult due to the diversity of crop kinds, agricultural disease symptoms, and environmental conditions [1]. To this end, a number of machine-learning methods have been developed for diagnosing potato leaf diseases. However, many current approaches have constraints that limit their usefulness, such as only being tested on pictures of plant leaves from particular geographical regions. In this study, we present a new method for reliable disease identification in potato leaf populations by combining image processing with a Convolutional Neural Network (CNN) model [2]. Our objective is to create a sequential model for disease prediction in potato leaves using convolutional neural networks (CNNs) so that farmers may receive accurate and timely diagnoses.

It is impossible to stress the importance of diagnosing diseases in potato leaves as soon as possible. Without proper disease control, potatoes, a crucial food staple, can suffer severe output losses [3]. In addition, the destructive effects of disease epidemics may spread fast over agricultural areas. To lessen the influence of illnesses on potato output, early identification and efficient disease control measures are essential. Potato leaf disease diagnosis has always relied on the visual inspection of leaves by professionals. This method, however, is tedious, laborious, and error-prone. There has been a rise in interest in creating automated plant disease detection systems thanks to the development of machine learning methods [4].

Recently, image processing and analysis have become useful instruments for identifying plant diseases in leaves. It is feasible to distinguish between healthy and unhealthy plant leaves by analyzing digital photographs of the leaves and extracting significant traits [5]. In order to understand the intricate patterns and traits associated with various illnesses, machine learning techniques, such as convolutional neural networks (CNNs), can be trained on a huge collection of tagged pictures.

In this study, we utilize CNNs to forecast diseases that may affect potato leaves. Convolutional neural networks (CNNs) are a sort of deep learning model optimized for visual data processing. They've had a lot of success in a number of computer vision tasks, such as picture categorization and object identification. The model automatically learns and extracts significant information from the input photos thanks to the hierarchical structure of CNNs, allowing for precise illness prediction.

We ran trials using a collection of potato leaf pictures to gauge the efficacy of our proposed CNN-based sequential model. The dataset includes both healthy and diseased potato leaves to capture morphological shifts in the leaves as the illness progresses. Our model's high rate of accuracy (92.58%) suggests it may be a useful tool for detecting diseases in potato leaves. However, when tested on misshapen potato leaves, the model's performance was inconsistent. This suggests that further work is required to strengthen the model's capability of identifying illnesses with unusual symptoms. However, our study paves the way for more research into resolving this difficulty and enhancing the precision of illness identification in potato leaves. In order to help farmers take preventative and control actions in a timely manner, we present a CNN-based sequential model for properly recognizing and categorizing good and diseased potato leaves. Farmers may reduce the spread and severity of illnesses by adopting disease-prevention measures include early identification and the use of targeted pesticide treatments, crop rotation, and cultural practices.

This study introduces an original method for predicting diseases in potato leaves through the use of image processing and a sequential model built on the convolutional neural network. The outcomes prove the viability of the suggested methodology for disease detection in potato leaf populations. Our work adds to the creation of automated systems that can help farmers avoid catastrophic financial losses and maintain sustainable potato production, while more study is needed to improve the model's performance on damaged leaves.

1.2 Motivation

The motivation of this study stems from the challenges and implications associated with early disease diagnosis in potato leaves. The researchers recognize the complex nature of disease detection in potato crops, which is hindered by various factors such as diverse crop types, agricultural disease signs, and environmental influences. The inability to promptly and accurately identify diseases in potato leaves can result in severe financial losses for farmers. The researchers aim to address this pressing issue by leveraging the potential of machine learning techniques, specifically image processing and Convolutional Neural Networks (CNNs), to develop an automated system for disease prediction in potato leaves. By utilizing these advanced technologies, they seek to enhance the efficiency and accuracy of disease detection, enabling farmers to take proactive measures to prevent disease outbreaks and mitigate their impact.

1.3 Research Objective

- To find the best accuracy
- To contribute in the agriculture field
- Make an innovative way to detect the leaf disease
- To get the best machine-learning model

1.4 Research Questions

- How well does the CNN-based sequential model identify potato leaf diseases compared to visual inspection?
- What is the accuracy?
- How researchers pre-process the data?
- What research have performed by others?

1.5 Expected Output

The expected output of this research is a CNN-based sequential model for disease prediction in potato leaves that demonstrates high accuracy in detecting diseases. The model is anticipated to outperform traditional visual inspection methods, providing a more efficient and reliable approach for early disease diagnosis. It is expected that the model will showcase its effectiveness in detecting a wide range of diseases in potato leaves, contributing to timely disease management strategies and preventing severe financial losses for farmers. Furthermore, the research aims to assess the model's performance on both typical and deformed potato leaves, with the anticipation of gaining insights into its ability to handle variations in leaf morphology and atypical disease symptoms. The study also aims to evaluate the generalizability of the model across different geographical regions, ensuring its applicability in diverse agricultural settings. Overall, the expected output of this research is a robust and adaptable CNN-based sequential model that can significantly enhance disease detection and aid in preserving the health and productivity of potato crops.

1.6 Report Layout

The report comprises a total of six chapters and will be accompanied by instructions to be followed.

The discourse on the prognosis of potato leaf disease, which was extensively covered in Chapter 1, has been further expounded upon in this section. This section provides an overview of the research question, the motivation behind conducting the investigation, the rationale, and the expected outcomes.

The second chapter of the study involved a comprehensive review of existing literature on the prediction of foliage diseases. This section encompasses a discussion of the methodologies, discoveries, and viewpoints of various scholars. The manuscript additionally discusses the prospective scope of the investigation and potential challenges that may be encountered

Chapter 3 has centered on the discourse of study methodologies. This chapter presents information pertaining to data gathering, data statistics, classification algorithms, and various statistical measures. This chapter has also addressed the fundamental aspects of implementing this into practice.

Chapter 4 contains the majority of the results obtained from this research. This section presents the results obtained from all the algorithms that participated in the study.

The fifth chapter of the research showcases the societal ramifications of the investigation. This section primarily illustrates the importance of this study and its potential for sustained relevance in the respective field.

Chapter 6 provides a discussion and conclusion regarding the findings of the study. Furthermore, the present section encompasses an analysis of the extent of forthcoming research and the limitations of the current study.

CHAPTER 2

BACKGROUND

2.1 Introduction

Conducting background research is a crucial component of any research endeavor, including investigations pertaining to the identification of potato leaf disease. It facilitates researchers in acquiring a more profound comprehension of the ailment, encompassing its etiology, clinical manifestations, and incidence. This data has the potential to assist researchers in devising more effective detection techniques and suitable strategies for management. The utilization of background investigation outcomes can serve as a means to direct subsequent research endeavors and guarantee the relevance and applicability of the findings.

2.2 Related Works

Recent research has made significant progress in the categorization of leaf diseases. The present study scrutinizes the diverse methodologies for disease detection that have been explored by scholars in the domain. The present investigation ascertained the proportion of the solar spectrum that is assimilated by the leaf architecture employed for the identification of potato mold. Automated segmentation of images utilizing deep neural networks. Following the development of models capable of distinguishing between bays and spectra, greenhouse studies employed spectra to independently assess the severity of diseases. The precision of foliage is augmented to 92% through the incorporation of an analysis of three discrete patterns of wound growth that are induced by blight. The algorithm's accuracy was evaluated on datasets comprising of pre-annotated artificial leaves, simulated healthy trees, and simulated black-crowned trees, yielding a precision rate of 74.6% [6].

Ferentinos et al. conducted a study in which they evaluated the performance of five distinct CNN architectures. Their findings indicated that the VGG architecture exhibited the highest level of accuracy, achieving a rate of 99.53%, in the classification of 58 distinct categories. The study evaluated the models' performances by utilizing a subset of the dataset that was employed for training purposes. According to a study, the model's performance significantly deteriorates when trained on simulated data and subsequently assessed on real-world data [7].

Kamal et al. [8] developed two models, namely Modified MobileNet and Simplified MobileNet, for the identification of plant foliage diseases by replacing the convolution layer of MobileNet with depth-wise separable convolution. The proposed model was trained utilizing data from the Plant Village dataset, which comprises images of plant leaves obtained from a limited geographical range of fields.

This study outlines the systematic approach employed in the identification and classification of potato leaf diseases through the utilization of a combination of traditional methods and machine learning-based systems for the purpose of automation. The present study presents the results of a comprehensive picture categorization task that involved collecting 450 photographs of Potato leaves, both healthy and diseased, from publicly accessible plant repositories. The images were subjected to recognition and classification procedures utilizing databases and seven distinct techniques. The Irregular Forest Classifier yielded the highest level of accuracy, reaching 97% [9].

The application of various machine learning techniques such as support vector machines, RGB image analysis, and other related methods have been extensively employed in agricultural disease forecasting research. Islam et al. employed a Support Vector Machine-based multi-class model to detect tuber diseases. The subject of the photographs amounted to a total of 300 images. The performance of the model investigator has been evaluated based on practical qualities such as sensitivity, accuracy, F1-score, and recall. A deep Convolutional Neural Networks model was developed by Sladojevic et al. with the ability to differentiate between apple and tomato plants. The method employed by scientists involved the analysis of healthy and diseased leaves to distinguish among 13 distinct illnesses [10]. Ferentinos et al. successfully differentiated 25 unique disorders by utilizing AlexNetOWTBn and VGG, which are deep learning systems. By employing a Region-based fully convolutional network, Regions integrated with deep neural networks, and a Solid-State Drive, we successfully achieved real-time identification of diseases in tomato plants.

According to Khalifa et al., Convolutional Neural Networks (CNNs) exhibit the ability to differentiate between healthy plants and those that are affected by early and late blight. The algorithm's predictive capabilities are limited to the success of products manufactured in particular regions, as it was exclusively trained on the Plant Village dataset. Rozaqi and Sunyoto proposed a convolutional neural network (CNN) architecture for the purpose of diagnosing illnesses in potatoes. The Plant Village dataset was utilized to train models for the purpose of detecting prevalent diseases in the region [11]. Sanjeev et al. have presented a Feed-Forward Neural Network (FFNN) that is capable of distinguishing between healthy and diseased foliage. The Plant Village dataset was employed for both the training and testing phases of the proposed method. Barman and colleagues developed a self-constructed convolutional neural network (SBCNN) model to classify images of potato leaves into three categories: early blight, late blight, or healthy, as reported in their study [12].

2.3 Scope of the problem

The focus of this study pertains to the identification of diseases in potato leaves at an early stage. The intricacy of the matter stems from the diverse array of crop varieties, manifestations of agricultural diseases, and ecological elements that can impact the identification of diseases. The issue pertains to the difficulties encountered by farmers in precisely detecting diseases during their nascent stages. This is because a delayed or erroneous diagnosis can lead to significant economic setbacks. The study centers on the utilization of machine learning methodologies, particularly image processing and convolutional neural network (CNN) models, to create an effective approach for identifying diseases in potato leaf populations. The scope of this study involves the evaluation of the model's performance on potato leaves that are both typical and deformed. Additionally, the study aims to assess the generalizability of the model across various geographical regions. The research endeavors to offer a comprehensive solution that can enhance disease management practices in potato farming by addressing these aspects.

2.4 Challenges

- To get the model accuracy above 90%
- Implementation with good accuracy
- Making literature review
- Finding other works research gap

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

This section provides a succinct overview of our research aims and the methodology utilized to achieve these objectives. The provided information and strategies are comprehensive and applicable. From our perspective, this can be interpreted as a focused execution strategy. The reliability of our data collection methods will undergo a similar outcome. The forthcoming discussion will elucidate the methodology we intend to employ in the application of said criteria towards the categorization of maladies affecting potato foliage. This section will expound on the methodologies and materials that can be employed in a research investigation.

3.2 Research Subject and Instrumentation

We are currently investigating a more reliable approach for classifying diseases in potato foliage. We opted to conduct a more thorough analysis of certain images to ascertain whether any modifications had been made to the vegetation. There has been an increase in the quantity of potatoes harvested in Bangladesh. The majority of farms in the United States lack individuals with college degrees. The identification of diseases that afflict potatoes is widely recognized as a challenging task. Due to the similarity of symptoms among different diseases, visual inspection alone is insufficient to differentiate a diseased potato leaf from a healthy one. The efficacy of any treatment modality is contingent upon the availability of an accurate prognosis. The implementation of an alternative therapeutic approach for a particular ailment may potentially result in adverse impacts on the surrounding environment, including the agricultural produce and the land. The decline in soil fertility may result in crop failure, leading to substantial economic losses for producers. These factors led us to conclude that this location would be the most suitable area to concentrate our research efforts. Convolutional neural networks (CNNs) can be utilized to construct a neural model that eliminates superfluous information from images and subsequently converts the residual data into correlations with quality scores. A specialized environment is provided for the construction of diverse neural networks that can effectively address the critical challenges that emerge in the request process. Prioritizing their resolution before comparing them to conventional techniques can facilitate systematic dispatch and enhance the effectiveness of their implementation. In order to prevent potato disease, we have experimented with several rigorous reading formats.

3.3 Data Description

In order to properly carry out research, it is required to collect a significant amount of data. In this particular circumstance, a legitimate data gathering would be of great use. We went out and shot some pictures of the potato fields in the hamlet so that we could supplement the information that we were able to gain from the publicly available statistics. The database had around 2158 photographs that were divided into three groups. Every image you've ever acquired has been saved as a JPG after being compressed, and its colors have been modified so that they conform to the RGB standard. The percentage of late blight, early blight, and healthy data that is present in potato samples is shown in Fig. 3.1.

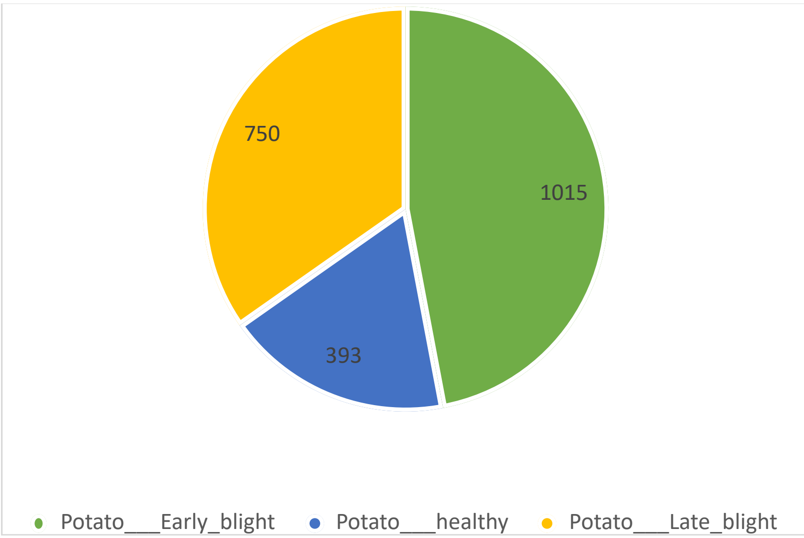

Fig. 3.1. Data Ratio

The quality of these pictures isn't very high across the board. This indicated that the audio was not altered in any manner prior to being sent over the air. As a result of the photos in the database being scaled down to a resolution of 60 pixels by 60 pixels, the training exercise may be finished in a shorter amount of time. You may be able to acquire new information more quickly by using this method to evaluate relevant or crucial metrics. The formal standing of the growth problem is enhanced in order to reach this goal. Confirmation was given to several inaccurate hypotheses regarding the beginning and the finish of the narrative. The fundamental purpose of this investigation is to determine all pixel respects that have equal access by employing conventional repercussions and modifications to image ratings. In the field of artificial intelligence, this metric is most typically referred to by its more familiar name, the Z-score. Figures 3.2–3.4 show an example image of the obtained dataset, which may be found further down in this section.

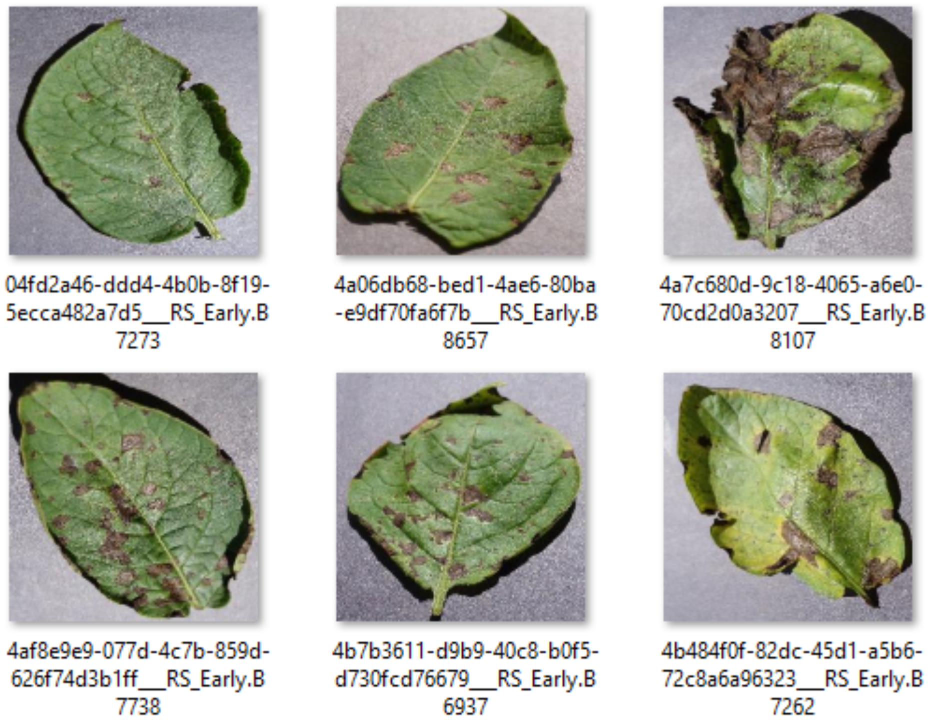

Fig. 3.2. Leaf of Potato Early Blight

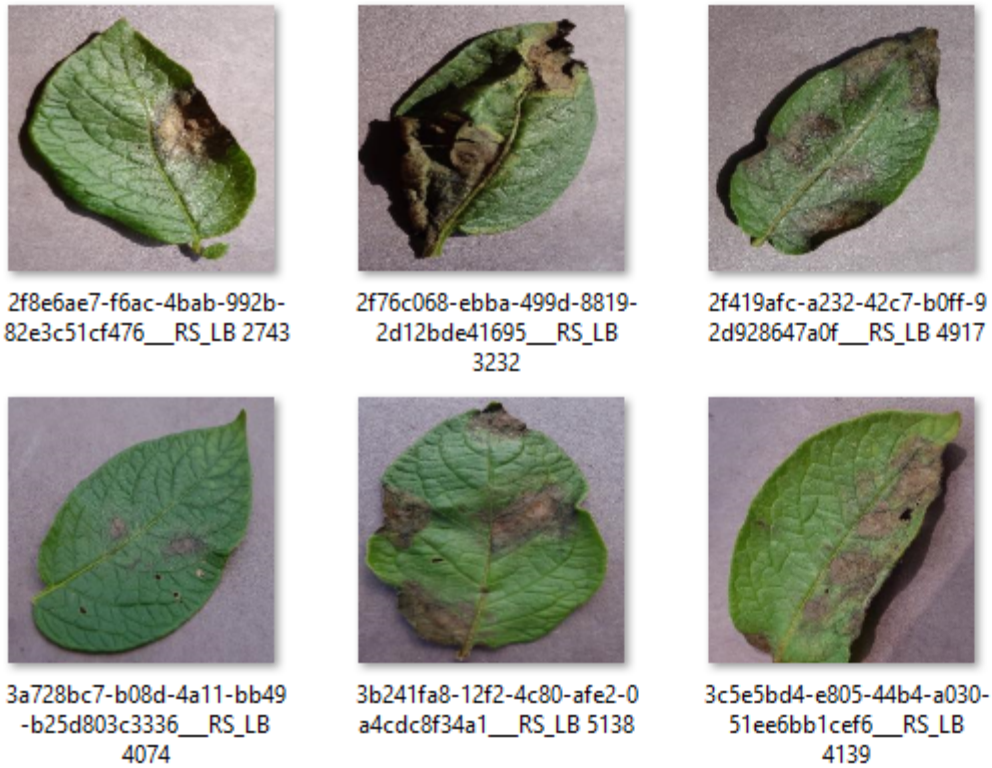

Fig. 3.3. Leaf of Potato Late Blight

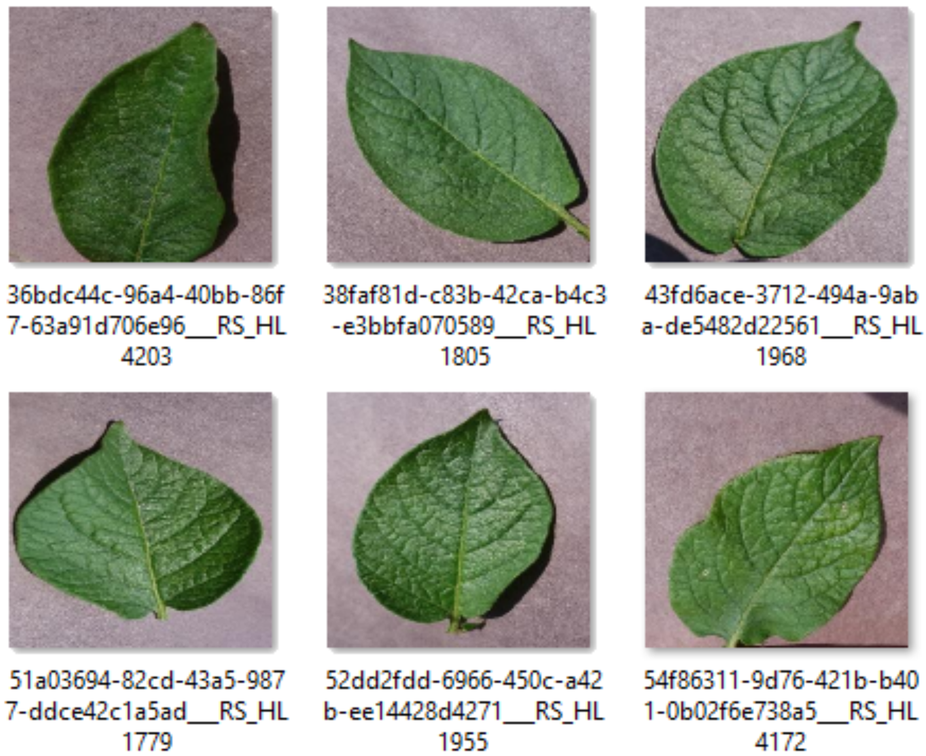

Fig. 3.4. Healthy Potato Leaf

3.4 Data Pre-Processing

As part of our data preparation, we employed a standardization approach. When image pixels are projected to a new area, usually (0,1) or (-1, 1), this is known as rescaling. Because of its widespread application, we standardize several data types to ensure uniform processing. This guarantees that every image is given the same weight. If the images are all scaled to the same range, say [0,1] or [-1,1], then the loss is shared equally among them. Rescaling guarantees that all photos improve at the same rate, as lower learning rates are necessary for higher-quality photographs and vice versa.

After that, a grayscaling process was applied, which turned the color photos into black and white. In order to reduce computational complexity, it is extensively used in machine learning processes (like SVM in the current work) [17]. Grayscale images are preferred over color ones when shade identification is not needed since they have fewer pixels and hence need fewer calculations. When images are standardized, the height and width are scaled to the same proportions rather than just the height. The data are normalized with a mean of zero and a standard deviation of one. variation per unit. This approach has the potential to increase data consistency and trustworthiness.

3.5 Proposed Methodology

A computer model can be developed through the utilization of a convolutional neural network (CNN) to filter out irrelevant visual data and establish associations with output scores. Each of the numerous brain structures possesses a distinct compartment, facilitating personalized focus on the particular concerns involved in a given inquiry. Prioritizing the establishment of control over the subject matter is crucial in order to facilitate a systematic delivery that can produce superior outcomes compared to traditional methods. Several methods have been attempted to gain a comprehensive understanding of the tomato environment, utilizing bottom-up learning approaches to achieve clarity in the image obtained. The process of categorizing images is facilitated by a CNN sequential model, which employs a sequence of convolutional, pooling, and dense layers to extract relevant features from incoming images. The training of the model involves the utilization of a set of annotated images of potato leaves that exhibit both healthy and diseased states, specifically for the purpose of detecting diseases in the leaves. During the course of its training, the model acquires discernible signals from the images that signify the presence of a healthy or diseased subject. Upon input of a novel image for prognostication, the model undergoes processing via identical layers utilized during its training phase, albeit with pre-learned weights. The final softmax activation layer produces the probabilities of the image aligning with each class. An image is deemed to be afflicted if the probability of it belonging to the diseased category exceeds a pre-established threshold, usually set at 0.5. The CNN sequential model demonstrates efficacy in detecting intricate patterns in image data, rendering it a valuable tool for detecting potato leaf diseases in agricultural settings.

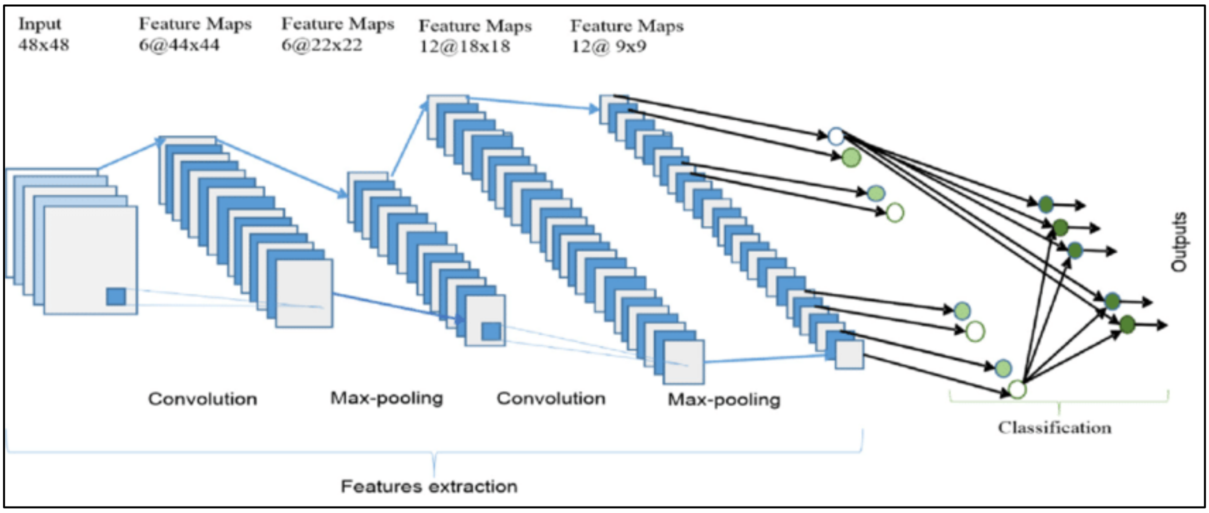

Fig. 3.5. CNN architecture

The utilization of the convolutional layer, which is capable of removing highlights, is employed in this context. As the ascent progresses from initiation to culmination, the exceptional caliber of the segments intensifies. Although there has been a significant surge in Long-Term Ecological Research (LTER) conducted across various regions, the size of the channel area has remained constant at 2 by 2. The maximum Long-Term Ecological Research (LTER) identifier within the vicinity is 66, while the remaining numerical values have remained constant. The utilization of aggregate levels across all squares results in an extensive array of entities, however, augmenting the quantity of Long-Term Ecological Research sites can potentially mitigate this issue [13]. Upon the application of the convolution technique to an image, the resultant object mappings are superimposed with zero values to preserve the initial dimensions of the image. By utilizing the uppermost integration layer, it is possible to reduce the dimensions of object representations, expedite preparedness, and enhance the precision of data modifications. The pieces utilized for maximum-pooling are of a spherical shape and possess a dimension of 22. The ReLU activation function is utilized as the execution layer for all blocks within a detached introduction. Furthermore, a Dropout production methodology was devised, which has the capability to conserve 0.5, while retaining an equivalent tactical separation from the railway. Engaging in a random act of dumping can alleviate concerns within an organization amidst a period of heightened focus on mitigating model disparities and streamlining the structure to prevent overcrowding. The amalgamation of a dual structure block has been fully integrated by the 500-nerve and 10-nerve cerebral formations.

3.6 Proposed Model Workflow

The identification of a disease that affects potato leaves needs many approaches. Following the presentation of Fig. 3.6. comes the presentation of the block diagram that depicts the proposed model method for this investigation.

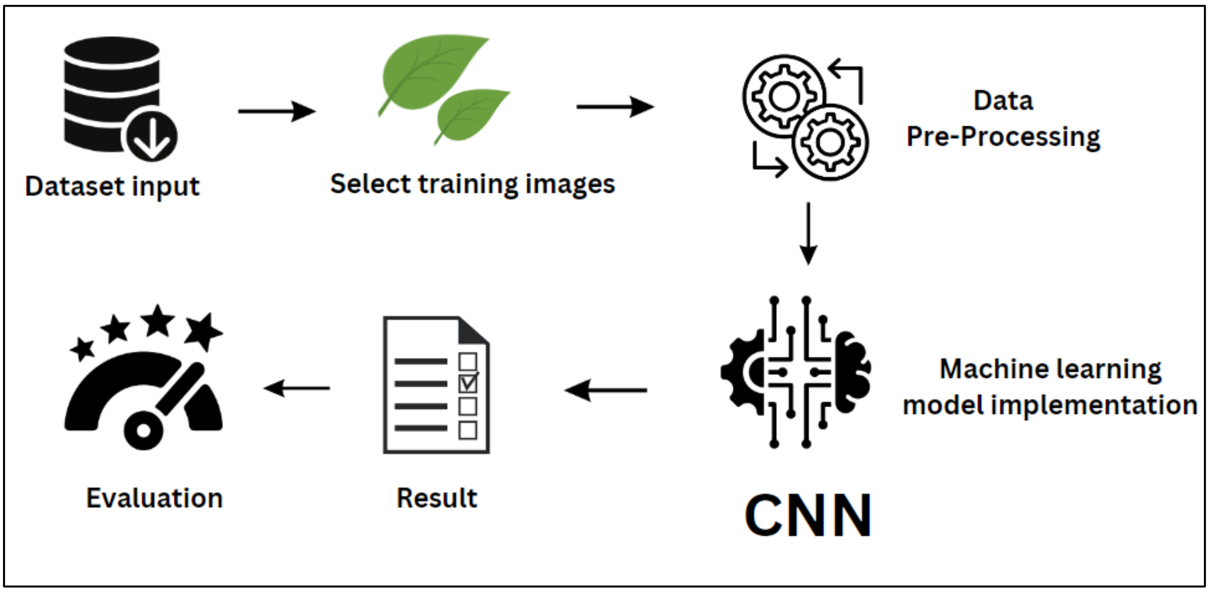

Fig. 3.6. Proposed Model Workflow

3.7 CNN

The full name of this structure is the Convolutional Neural Network. In the field of computer vision, this type of neural network is widely used for tasks including object detection, segmentation, and the classification of images. The idea upon which CNNs are built is that they can execute visual tasks as well as the human brain. There are several tiers of connected nodes in this network, and they all handle the data in their own special way. The first layer of a convolutional neural network (CNN) is often used to extract features (such as edges, angles, and patterns) from an incoming image. The layer's output is fed into a non-linear activation function, such as ReLU (Rectified Linear Unit), which introduces non-linearity to the model. Following the first convolutional and activation layers, pooling layers are commonly used to lower the feature maps' dimensionality. The result is then sent into one or more fully connected layers, which provide the predictions. During training, a CNN's weights are adjusted so as to minimize a loss function that measures the amount of error between the predicted values and the ground truth labels. For this aim, gradient-based optimization techniques like Adam and Stochastic Gradient Descent are frequently employed.

3.8 Implementation Requirements

- Windows 10
- Internet Connection
- Google Colab
- Photoshop
- Python

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

The primary focus of the fourth chapter centers on a descriptive analysis of the research data and testing outcomes obtained from our project. It is imperative to provide equal opportunity to every strategy and idea. Demonstrating the validity of one's reasoning through evidence is an effective means of enhancing its quality. The efficacy of our vision could be evaluated through the implementation of the dynamic framework outlined in my suggested approach. This section addresses the significance of the report, both internally and externally. This section additionally demonstrates the external and internal limits of the construction of the casing.

4.2 Experimental Results

A diverse range of metrics were employed by us to evaluate the efficacy of their models. The majority of research endeavors employed a set of multiple indicators to assess efficacy, while a minority relied on a singular indicator. This study's quality is evaluated based on the precision of the data, the graphical representation of the findings, and the degree of uncertainty reflected in the index. In this study, a CNN successor model consisting of seven steps was developed. By utilizing the Adam optimizer and implementing error monitoring techniques, along with strategic adjustments to the cross-entropy function, optimal performance can be achieved. The utilization of transfer learning is employed to accomplish the task of model completion. The Keras library [14], along with its associated applications, offers pre-trained weights that are essential for this objective. Fig. 4.1 shows the training accuracy vs validation accuracy and Fig. 4.2 shows the training loss vs validation loss.

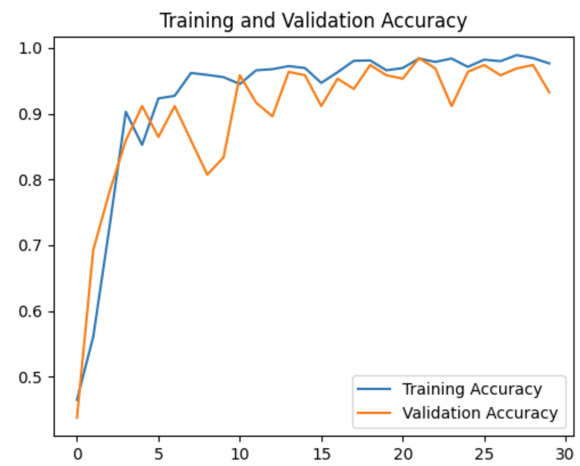

Fig. 4.1. Training AUC vs Validation AUC

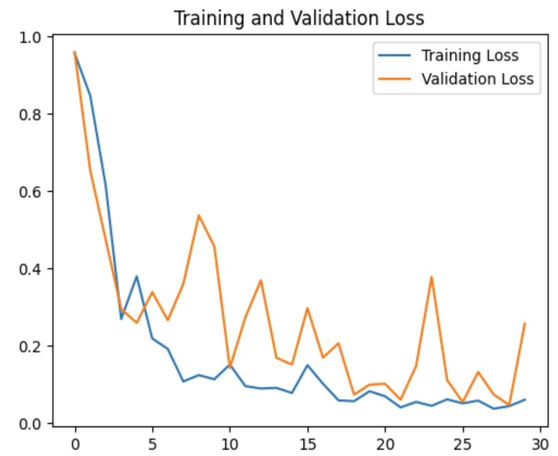

Fig. 4.2. Training Loss vs Validation Loss

According to the results of the evaluation of the model, this effort has been successful. CNN has managed to acquire a model accuracy of 92.58%. The results of this study are summarized in Table 2. Fig 4.3 shows the accuracy result of this work.

```
[ ]  scores = model.evaluate(test_ds)

     8/8 [==============================] - 30s 2s/step - loss: 0.2264 - accuracy: 0.9258
```

Fig. 4.3. Accuracy vs Data Loss

4.3 Discussion

The use of confusion matrices to characterize a categorization method's performance has grown in recent years. Classification precision alone may be misleading if your dataset contains more than two classes or if there is a large disparity between the number of samples in each class. The benefits and flaws of your classification model can be better understood by computing a confusion matrix [15]. The confusion matrix serves as a visual depiction of an algorithm's results. We compare and contrast the right and incorrect beliefs. Model equation accuracy and classification rate equation precision are illustrated by the accompanying instances.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

Here, TP= True Positive, TN= True Negative, FN= False Negative, FP= False Positive.

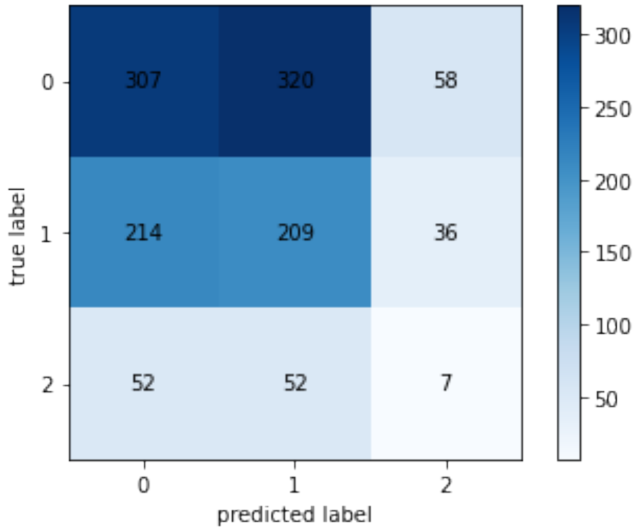

Fig. 4.4. Confusion Matrix

A confusion matrix depicting our categorization model can be seen in Fig. 4.5. This is due to Row 1 having the highest cross-gain. This means that the technique we present will be as accurate as feasible.

The Google collaborative notepad acts as a central archive for all phases of the work. The first stage involved computing several images of infected potato leaves for classification. The photos are selected from the dataset's collection. Once the images have been submitted, the prediction can be viewed.

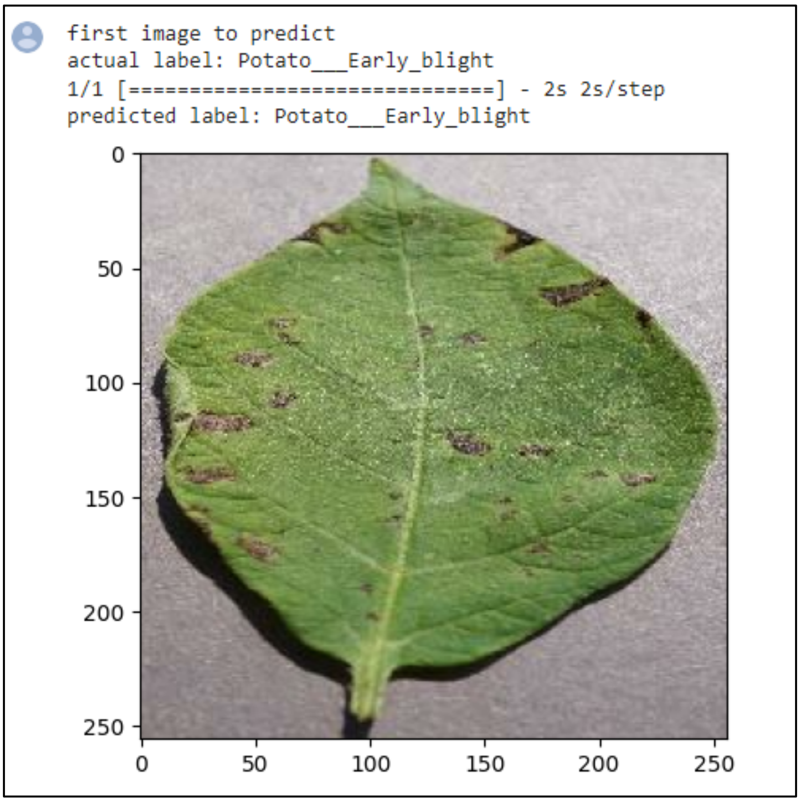

Fig. 4.5. Predicting leaf disease

As we can see in figure 4.5 our model can predict the disease perfectly. Here actual label is potato_early_blight, and our model have predicted potato_early_blight also.

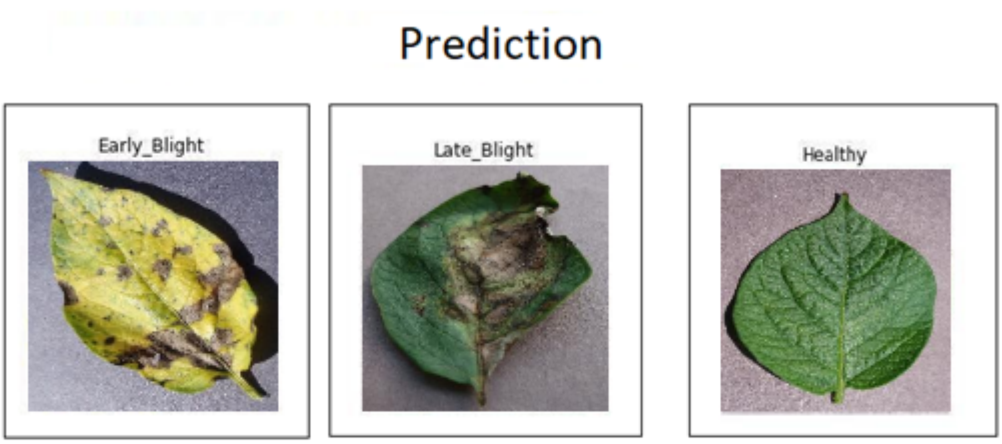

Fig. 4.6. (a) Early Blight, (b) Late Blight, and (c) Healthy

Similar observations can be made as depicted in Figure 4.6 (a). The manifestation of Early Blight on the leaf is evident. Frequently, the dots will enlarge and transform into a circular shape that exhibits a golden or brilliant green-yellow hue. The occurrence of extensive spotting may lead to the yellowing and subsequent death of entire leaves. Once the scores reach a certain threshold, this phenomenon may manifest. The impact of late blight on foliage is evident in Figure 4.6. (b). The presence of diminutive and asymmetrical verdant markings exhibiting varying degrees of luminosity indicates the incipient phases of foliage fragments. During cold and wet weather conditions, the small specks exhibit a rapid expansion and transform into extensive, polychromatic regions spanning a spectrum of colors from brown to deep purple. The complete ingestion of leaflets or the gradual spread of the disease from branchlets to the stalk can result in the fatality of the plant. Figure 4.6 (c) depicts a sturdy leaf. Thus, this approach can be utilized to precisely differentiate between images portraying health and those portraying sickness.

The objective of this research was to identify the particular disease that affects the leaves of potatoes. The incorporation of CNN techniques into the sequence model resulted in a commendable performance of the model on the given dataset. The present study exhibits a higher level of efficacy compared to certain works referenced in the literature review.

CHAPTER 5

IMPACT ON SOCIETY, ENVIRONMENT, AND SUSTAINABILITY

5.1 Impact on Society

Potatoes are a widely consumed dietary staple across various regions of the world. While this particular cuisine may be perceived as a supplementary dish in Western culture, it holds an essential status in Bangladesh. Historically, the cost of Bangladeshi rice was relatively high. The economically disadvantaged individuals were unable to avert their predicament due to their inability to procure rice at the escalated market rate. During that period, potatoes were priced at such a low rate that they were deemed irresistible to the impoverished population. In the event that such a circumstance arises, potatoes may serve as a means of sustenance for survival. Potatoes were utilized as a cost-saving measure in lieu of rice. The abundant harvest of potatoes that year had a direct correlation with the low market price of this commodity.

Potatoes are utilized for a diverse range of applications. Potatoes are utilized in the production of various types of alcoholic beverages. Potato starch is a widely utilized agent for thickening and binding in the food sector, particularly in the preparation of stews and condiments. Potatoes are a frequently utilized component in the culinary traditions of numerous Asian nations, and are extensively relished by their respective populations. The capacity to predict the occurrence of potato leaf disease, as demonstrated in this study, holds significant implications for the field of agriculture.

5.2 Impact on Environment

The investigation of potato leaf disease detection has the potential to yield favorable environmental outcomes by mitigating the employment of pesticides and fungicides, which are detrimental to both ecological and human health. The timely detection of diseases in potato plants enables producers to implement specific measures to contain the spread of the disease, including the removal of infected plants or the use of natural predators to manage pests. Furthermore, timely identification can mitigate crop damage caused by pathogenic agents, leading to increased agricultural output and reduced food spoilage. The reduction of resources required for sustenance production can have a positive impact on the environment.

On the other hand, in cases where the research entails the utilization of dangerous substances or methodologies that have an impact on the soil or water systems, it may result in adverse environmental consequences.

5.3 Ethical Aspects

This study has taken into account the safety and privacy of all users. The material has been manually collected by our team of researchers. The present investigation did not utilize any pre-existing materials. The presented data is precise and reliable. The study has been purged of inaccuracies. The research conducted did not entail the utilization of any non-human organisms. All members of the group contributed equally, and our leader took charge as we walked.

5.4 Sustainability Plan

The detection of leaf disease is a crucial aspect of national security, as we are acutely aware. This may potentially offer utility in combatting illnesses within the agricultural sector. The process has the potential to perpetuate indefinitely. The potato is a widely-consumed culinary item with global popularity. This study is expected to be beneficial in aiding growers in their efforts to prevent potato leaf disease. The enduring impact of this research on the domain of technology is expected to persist for numerous years. It is anticipated that forthcoming iterations will enhance both precision and coverage. In the future, we intend to employ supplementary techniques to juxtapose and differentiate them in order to determine the most precise one.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Summary of the study

Upon the completion of the research, it was determined that the production of sufficient goods is of paramount importance for individuals. Access to nutritious food is crucial for maintaining a healthy lifestyle, as sustenance is a basic human need and adequate nourishment is essential for optimal health. Potatoes provide a plethora of beneficial minerals. Furthermore, it has the potential to generate a significant amount of energy. This implies that there exists a noteworthy demand for potatoes among individuals. Additionally, it is utilized in the preparation of various dehydrated meals that possess an extended shelf life without succumbing to spoilage. Moreover, it has gained immense popularity among the younger generation. Consequently, an increase in potato production could potentially lead to a significant alteration in our dietary patterns. The CNN algorithm was utilized to forecast the occurrence of diseased leaves. The research yielded a model accuracy of 92.58% upon running the model.

6.2 Conclusions

The primary aim of this study is to develop a machine learning methodology for the identification of diseases in potatoes that is not dependent on the surface characteristics of the potato. (CNN). The results of this study indicate that CNN is the most efficacious method for object recognition of this nature. Nevertheless, this particular model attains a higher level of confirmation accuracy, specifically 92.58%. This study employs both contemporary and archival datasets. A project of this nature holds significant potential for the agricultural industries worldwide. The majority of farmers residing in this rural settlement of Bangladesh lack formal education, rendering them vulnerable to the affliction. It is imperative that one refrains from disclosing their capacity to diagnose ailments to others. The potato crop is experiencing significant reduction due to insect infestation, with the agricultural sector being the most affected. The present study holds promise in improving the standard of living for potato cultivators in Bangladesh. The primary issues are inaccuracies present in the model and the information utilized for the development of the product. The improvement of accuracy can be achieved through the acquisition of additional data. The ultimate objective is to develop an Android application that can effectively identify problems in a given product and suggest suitable solutions. Enhancement of precision is a prospective outcome as the expansion of this collection ensues. The present study involves the creation of an Android application that will expand upon prior endeavors. Furthermore, a system will be established to facilitate convenient accessibility of disease diagnosis, advice, and assistance to affected farmers in Bangladesh.

6.3 Implication for further study

- To do this, we plan to: Apply further machine learning methods to enhance this model's accuracy.
- In the near future, we plan to use a deep learning technique to implement this model.
- In the future, we plan to: collect even more data in bulk; develop real-time detection methods.

REFERENCE

1. Wani, J. A., Sharma, S., Muzamil, M., Ahmed, S., Sharma, S., & Singh, S. (2022). Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges. *Archives of Computational Methods in Engineering*, *29*(1), 641-677.

2. Islam, M. T., Ahmed, T., Rashid, A. R., Islam, T., Rahman, M. S., & Habib, M. T. (2022, April). Convolutional Neural Network Based Partial Face Detection. In *2022 IEEE 7th International conference for Convergence in Technology (I2CT)* (pp. 1-6). IEEE.

3. Shi, Y., Han, L., Kleerekoper, A., Chang, S., & Hu, T. (2022). Novel cropdocnet model for automated potato late blight disease detection from unmanned aerial vehicle-based hyperspectral imagery. *Remote Sensing*, *14*(02), 396.

4. Bangari, S., Rachana, P., Gupta, N., Sudi, P. S., & Baniya, K. K. (2022, February). A survey on disease detection of a potato leaf using cnn. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (pp. 144-149). IEEE.

5. Suganyadevi, S., Seethalakshmi, V., & Balasamy, K. (2022). A review on deep learning in medical image analysis. *International Journal of Multimedia Information Retrieval*, *11*(1), 19-38.

6. Tarik, M. I., Akter, S., Al Mamun, A., & Sattar, A. (2021, February). Potato disease detection using machine learning. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 800-803). IEEE.

7. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, *145*, 311-318.

8. Kamal, K. C., Yin, Z., Wu, M., & Wu, Z. (2019). Depthwise separable convolution architectures for plant disease classification. *Computers and Electronics in Agriculture*, *165*, 104948.

9. Frost, K. E., Groves, R. L., & Charkowski, A. O. (2013). Integrated control of potato pathogens through seed potato certification and provision of clean seed potatoes. *Plant Disease*, *97*(10), 1268-1280.

10. Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and electronics in agriculture*, *74*(1), 91-99.

11. Hammouda, K., Khalifa, F., Soliman, A., Ghazal, M., Abou El-Ghar, M., Haddad, A., ... & El-Baz, A. (2019, October). A CNN-based framework for bladder wall segmentation using MRI. In *2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME)* (pp. 1-4). IEEE.

12. Rashid, J., Khan, I., Ali, G., Almotiri, S. H., AlGhamdi, M. A., & Masood, K. (2021). Multi-level deep learning model for potato leaf disease recognition. *Electronics*, *10*(17), 2064.

13. Islam, T., Kundu, A., Lima, R. J., Hena, M. H., Sharif, O., Rahman, A., & Hasan, M. Z. (2023). 5 Review Analysis of Ride-Sharing Applications Using Machine Learning Approaches. *Computational Statistical Methodologies and Modeling for Artificial Intelligence*.

14. Islam, T., Kundu, A., Islam Khan, N., Chandra Bonik, C., Akter, F., & Jihadul Islam, M. (2022). Machine Learning Approaches to Predict Breast Cancer: Bangladesh Perspective. In *Ubiquitous Intelligent Systems: Proceedings of Second ICUIS 2022* (pp. 291-305). Singapore: Springer Nature Singapore.

15. Islam, T., Hosen, M. A., Mony, A., Hasan, M. T., Jahan, I., & Kundu, A. (2022, January). A proposed Bi-LSTM method to fake news detection. In *2022 International Conference for Advancement in Technology (ICONAT)* (pp. 1-5). IEEE.

final_and_Official_Thesis_2023.docx

ORIGINALITY REPORT

19%	14%	10%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	8%
2	Choyon Chandra Bonik, Flora Akter, Md. Harunur Rashid, Abdus Sattar. "A Convolutional Neural Network Based Potato Leaf Diseases Detection Using Sequential Model", 2023 International Conference for Advancement in Technology (ICONAT), 2023 Publication	3%
3	Submitted to Morgan Park High School Student Paper	2%
4	www.mdpi.com Internet Source	1%
5	Submitted to Daffodil International University Student Paper	<1%
6	"Proceedings of International Joint Conference on Advances in Computational Intelligence", Springer Science and Business Media LLC, 2023 Publication	<1%