FIELD-SPECIFIC DEEP LEARNING SOLUTIONS FOR ACCURATE POTATO LEAF DISEASE CLASSIFICATION

Nandhini S[1], S. Thilagavathi[2]

Abstract

This work employs a convolutional neural network (CNN) and deep learning to classify potato leaf diseases. The suggested methodology consists of preprocessing the leaf image data, utilizing that data to train a CNN model, and assessing the model's performance on a test set. The trial's findings show that the CNN model has an overall accuracy of 97.1% in correctly identifying two forms of potato leaf diseases: Healthy and Late Blight. The suggested approach might offer a dependable and effective way to diagnose potato infections, which is essential to ensuring food security and minimizing financial losses in agriculture. The program can correctly distinguish between various disease types even when there are serious infections present. The potential of deep learning techniques for classifying potato diseases is demonstrated by this work, which may aid in efficient and automated disease management in potato cultivation.

Keywords :

Potato Leaf Diseases, Convolutional Neural Networks, Image Classification

I. INTRODUCTION

As one of the world's staple crops, potatoes confront many difficulties, with infections being a major danger to both productivity and quality. Timely intervention and management of an illness depend heavily on effective disease identification. Convolutional Neural Networks (CNNs) have become a potent tool for picture classification tasks, such as identifying plant diseases, in recent years. In this work, we present a CNN-based approach for potato leaf disease classification. Our solution attempts to effectively classify various potato leaf illnesses by utilizing CNNs' capacity to automatically learn relevant features from photos. This would help farmers and agricultural specialists discover diseases early and make informed decisions. We hope to contribute to the creation of more reliable and effective agricultural disease management systems by utilizing deep learning techniques, which will ultimately increase crop output and food security.

Potato leaf diseases are a group of illnesses that greatly affect the well-being and yield of potato plants, which are among the most major food crops in the world [1]. These illnesses can take many different forms, such as bacterial, viral, or fungal infections, as well as physiological issues. Potato virus Y (PVY), X (PVX), late blight, early blight, and potato leaf roll virus (PLRV) are typical examples. If allowed to spread, these illnesses have the potential to seriously impair crop quality, lower yields, and jeopardize food security. Early detection and precise identification are necessary for the effective management of potato leaf diseases. This allows for the prompt implementation of intervention measures; such as crop rotation plans or targeted pesticide applications.

Convolutional neural networks (CNNs) have transformed a number of domains, including computer vision, thanks to their remarkable capacity to extract complex characteristics and patterns from images [2]. CNNs are a subset of deep learning algorithms created especially for effectively analyzing visual input. These systems consist of several layers, including convolutional, pooling, and fully linked layers. These layers cooperate to automatically extract hierarchical representations from input images. CNNs identify elements like edges, textures, and forms through convolution, gradually combining them into higher-level representations to enable complicated picture understanding [3]. CNNs are especially well-suited for applications like object identification, segmentation, and picture classification because of their architecture.

Classifying images into preset classes or categories according to their visual information is a key task in computer vision. It is essential for several applications, including autonomous driving, agricultural monitoring, and medical diagnosis [4]. Creating models or algorithms that can correctly categorize images is the aim of image classification

Department of Computer Science and Engineering[1]
Karpagam Academy of Higher Education, Coimbatore.[1]
nandhiniphd07@gmail.com[1]

Research Scholar[2]
NGM College[2]
thilagavathibluesky@gmail.com[2]

* Corresponding Author

in order to facilitate automated analysis and decision-making. Recent years have seen notable improvements in picture categorization efficiency and accuracy thanks to the development of deep learning techniques, especially Convolutional Neural Networks (CNNs) [5]. Intricate patterns and variances in visual data can be captured by CNNs because of their exceptional ability to learn hierarchical representations of picture information. Consequently, they have emerged as the mainstay of contemporary image categorization systems, surpassing conventional techniques across multiple fields.

II. LITERATURE SURVEY

This study was suggested by Gulshan Shrivastava [6] et al. The agricultural industry has recently witnessed a rise in interest in using technology-particularly machine learning-to address a number of issues, such as the early diagnosis of crop diseases. However, because of the intricate interactions between several elements such as crop species, disease symptoms, and environmental circumstances, early detection of potato leaf diseases remains a challenging task even with tremendous progress made in this area. This intricacy frequently makes it challenging to quickly and effectively detect certain illnesses. Researchers have been investigating novel approaches to address these issues, and machine learning techniques have emerged as potentially effective ones. The creation of a multi-level deep learning model especially intended for the identification of potato leaf disease is one such development. There are two main tiers in this concept that are intended to improve efficiency and accuracy. Initially, the model extracts potato leaves from photos of potato plants using the YOLOv5 image segmentation method.

According to David Camacho [7] et al.'s proposal in this research, managing potato diseases is important for agriculture since they can result in a large loss in crop yield. To reduce the loss, it is therefore essential to identify and classify potato leaf diseases as soon as possible; nonetheless, this is a labor-intensive process that takes time. Timely detection and classification of data using an accurate automated technique are required to address the issues mentioned before. Using an existing dataset called "The Plant Village Dataset," deep learning and machine learning techniques have been utilized to construct methods that categorize potato leaves into two groups. The five classes of potato leaves that this article proposes to categorize are Potato Leaf Roll (PLR), Potato Verticillium Wilt (PVw), Potato Late Blight (PLB), Potato Early Blight (PEB), and Potato Healthy (PH) class. The approach for classifying potato leaves is based on an improved deep learning algorithm. To train the suggested model, "The Plant Village" dataset is utilized, which contains images of two different states, such as Early Blight (EB) and Late Blight (LB), along with a healthy class for potato leaves. Additionally, a thorough collection of data has been made for the classes Potato Leaf Roll (PLR), Potato Verticillium Wilt (PVw), and Potato Healthy (PH).

Ankit Kumar Jain[8] et al. proposed this investigation. Conducting pertinent research is essential for the development of sustainable agriculture, given the advancements in agricultural technology and the use of artificial intelligence in plant disease diagnostics. Many diseases, such as late and early blight, have a significant impact on the quantity and quality of potatoes. It requires a lot of work and effort to manually interpret these leaf diseases. Potato crop yield can be increased by accurately and automatically diagnosing these illnesses during the budding phase, however this does require a high level of expertise. In the past, a number of models have been put up to identify different plant diseases. This work introduces a method for fine-tuning (transfer learning) pre-trained models, such as VGG19, in order to extract relevant features from the dataset. After that, the findings were examined using various classifiers. Of them, logistic regression outperformed the rest in terms of classification accuracy, reaching 97.8% across the test dataset.

According to Cecile Zachlod [9] et al., potato leaf blight is one of the most harmful plant diseases in the world because it reduces the quantity and quality of potato crops, which has an adverse effect on individual farmers as well as the agricultural industry. The application of artificial intelligence technologies to the early classification and detection of crop blight has broadened the possibility for improving and broadening plant protection. This work presents a classification architecture for potato leaf blight.

The technology is based on deep convolutional neural networks. The training dataset includes three different categories of potato leaves: healthy leaves, early blight leaves, and late blight leaves. The proposed 14-layer architecture comprises of two basic convolutional layers with different convolution window sizes for feature extraction and two fully linked layers for classification. The amount of dataset images increased from 1,722 to 9,822 in this research by the application of augmentation techniques, which resulted in a notable improvement in the overall testing accuracy. The recommended architecture's mean total testing accuracy was 98%. In order to guarantee the authenticity and correctness of the data given, this study used more than six performance metrics. The accuracy of the suggested architecture was determined to be better when the testing accuracy of the suggested approach was compared to relevant works.

According to Ziad Ismail et al.'s analysis, potatoes rank among the world's most significant food crops [10]. In Bangladesh, the cultivation of potatoes has grown significantly in popularity during the last few decades. Potato plants cannot grow adequately due to a number of ailments. There are obvious diseases on this plant's leaf area. Two common and widely distributed leaf diseases that attack potato plants (LB) are Early Blight (EB) and Late Blight. Still, early identification of these diseases would lead to increased agricultural yield. The best method to locate and examine these abnormalities in order to address this issue is through image processing. Using image processing and machine learning, this research proposes an autonomous technique to detect and classify illnesses in potato leaves.

III. RELATED WORK

The key to preventing crop loss and maintaining the value of agricultural output is early detection of plant diseases. Research on crop diseases entails research on the patterns that can be seen in the plant. Plant disease identification and health monitoring are essential for sustainable agriculture. Accurate observation is a very challenging way to monitor and identify plant diseases. Disease-related crop loss is estimated to account for 15–25% of agricultural production in India. Therefore, in order to address this issue properly, our nation must not only boost productivity but also guarantee food security and nutrition. We present our automated plant disease detection system, which was created by incorporating advanced "Deep Learning" models such as the Support Vector Machine (SVM).

IV. MATERIAL AND METHODS

All of us are familiar with potatoes as a vegetable. In India, growing potatoes has been increasingly popular in recent decades. However, diseases like early and late blight are making it harder to produce potatoes and driving up production costs. The popularity of potato farms has increased dramatically in Bangladesh during the past few decades. However, a variety of ailments are driving up farmer costs in the potato industry. On the other hand, the majority of crop-related diseases are to blame for the high cost of producing potatoes. It is throwing the farmer's calendar into disarray. In order to modernize the potato industry and speed up illness diagnostics, automation has been implemented. Potato leaf disease is a serious issue that has the potential to significantly reduce crop productivity, despite claims to the contrary. When a potato plant is sick, its leaves will show signs of early, late, and Septoria blight, among other diseases. If such outbreaks are detected early and enough action is taken, the farmer won't be at risk of suffering significant financial losses. A new model is proposed based on the findings of this work that uses image processing to precisely identify and detect illnesses in potato leaf stands. A Convolutional Neural Network (CNN) model is used to identify the disease in potato leaf photos, while other machine learning techniques can also be utilized. To predict potato leaf disease, this work employs a sequential CNN-based model. There was a 94.2% accuracy rate for this model in the research. Both normal and disordered potato leaves were used to test the suggested model's ability to distinguish between them. Following the application of the algorithm to the images, the potato tree's leaf is classified as either healthy or unhealthy.

A. Dataset

A useful and well-designed dataset is an essential and productive phase in the model's training process. The dataset is divided into three classes to help the algorithm recognize potato leaf diseases during training. Figure 2 shows our datasets for early blight and late brilliant potato leaves, which include 656 original photos, 152 healthy leaf samples, and 504 unhealthy image samples. Thanks to data augmentation techniques, the dataset's size can be enhanced to produce trustworthy results. 2176 augmented leaf samples of potato leaves were gathered after the use of augmentation techniques. To improve and modify the dataset, it was randomly mixed from the surviving datasets, 80 were used as training sets and the remaining 80 as testing sets.

Table I. Dataset Image With Augmentation

Dataset Class	Potato Dataset
Original Images	656
Augmented Images	2176
Training Dataset	1728
Testing Dataset	448
Classes	03

B. Image Pre-processing

Pre-processing images is a crucial step in creating a high-caliber model. Before training the model, the input images usually have complicated backgrounds and noise, making them unsuitable for training neural network models. Images can be enhanced for further processing by removing noise and distortion through the use of image pre-processing techniques. To save time and improve the model's performance, the input photos are downsized to 256x256.

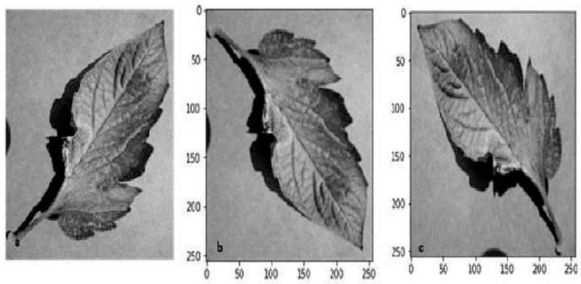


Fig 1. Generation of leaf images via data augmentation techniques: a) resized image, b) flip, c) rotation

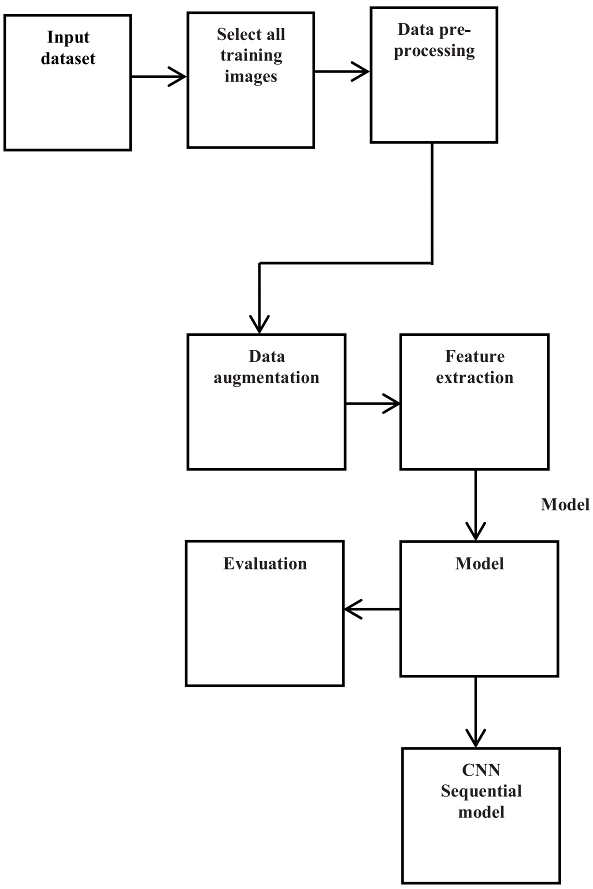


Fig 2. Proposed Workflow

C. Data Augmentation

A popular approach in computer vision and deep learning, data augmentation uses random changes such as scaling flipping, rotation, and cropping to increase the amount of data in datasets. In order to address the overfitting issue that arises when the training data of a proposed model yields irrelevant patterns, a larger dataset is necessary. Deep learning model augmentation techniques are a great way to enhance the robustness of the model and boost its performance. The amount of datasets was increased by the application of augmentation techniques, which employed random modifications such as scaling flipping, rotation, and cropping.

D. Disease Classification Using Cnn

Using convolutional neural networks (CNNs), high level features are extracted from an input dataset. In order to do classification tasks, it first gathers low-level information, which it then combines and maps to high-level characteristics. Well-known neural networks called convolutional neural networks are helpful for many different tasks, such as classifying and diagnosing leaf diseases [16]. To construct a CNN, there are often a few key processes. Convolution is applied to the input data using convolution layers, which produce a feature map. Through the use of convolution, the filters have been moved both vertically and horizontally. As demonstrated in the equation below, the model's non-linearity has been effectively demonstrated by using the activation function to correctly convert all negative values to zero.

$$() \ (0,)\ '() = \{ \quad 1, \ > 0 \quad 0, h$$

In order to reduce computational training, feature maps were down-sampled in the pooling layer, and flattening layers were used to transform multi-dimensional vectors to one-dimensional vectors. The findings of the previews for prediction and classification were then sent to the dense layer. Figure 4 illustrates our proposed model for potato leaf disease diagnostics, which makes use of convolutional neural networks with convolution layers, max-pooling layers, flattens, and dense layers.

The CNN model architecture's hyper-parameter setting is displayed in table 2.

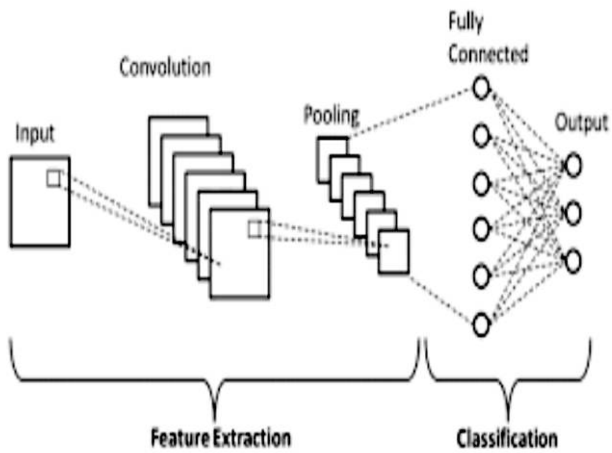


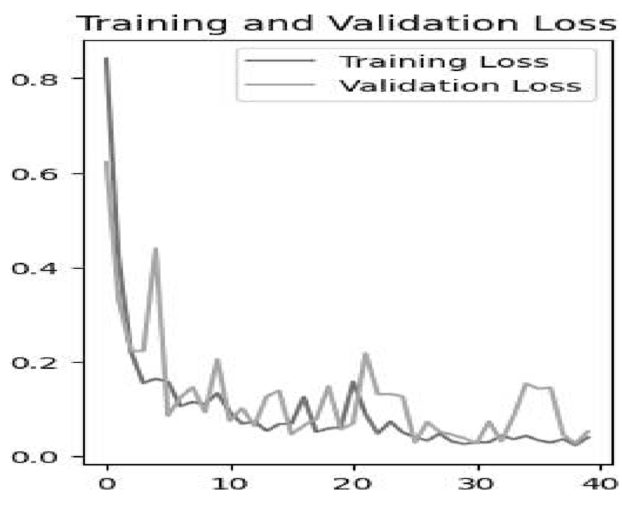
Fig 3. Architecture Convolutional

Table I. Cnn Model Architecture Hyper-parameters Configuration

Optimizer	Adam Optimization Algorithm
Batch size	32
Image size	256x256
Channels	3
Epochs	40
Loss	Sparse_categ-rical_crossent-opy

V. RESULT ANALYSIS

The CNN model that we developed can distinguish between healthy and diseased leaves. In addition, we implemented our tasks using a variety of techniques and system dependencies, such as installing Tensor Flow using Python libraries. Moreover, a variety of plant diseases exist. Thus, for our study, we have chosen three prevalent varieties of potato leaf diseases. We have classified and grouped our photos according to the different disorders. In order to demonstrate non-linearity in the network, we used rectified linear units (ReLU). Ultimately, the softmax function was employed in the final deep learning layer to classify and identify potato leaf illnesses. Adaptive moment estimation was then used to optimize the model. In contrast to conventional machine learning techniques, our research yields noteworthy outcomes for the categorization of potato leaf illnesses, making it noteworthy in the agricultural domain. Furthermore, in order to prevent the convolutional neural network model's training time, we decreased the amount of parameters. Our research employs more sophisticated techniques that save time and offer a deeper comprehension of plant leaf disease detection utilizing neural networks. In addition, it has been discovered that our suggested model is very successful scoring a 97%.



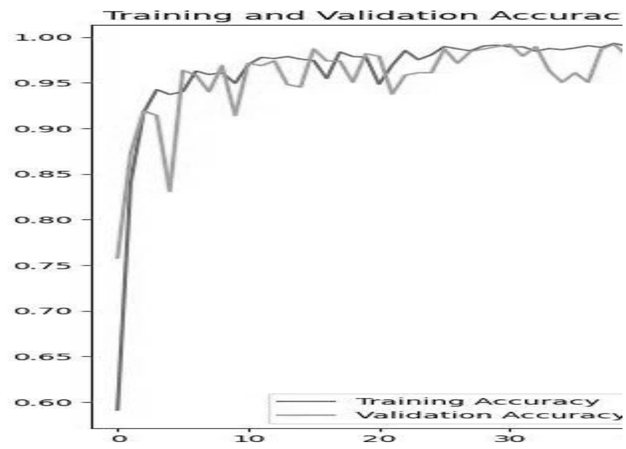


Fig 4. Training Process

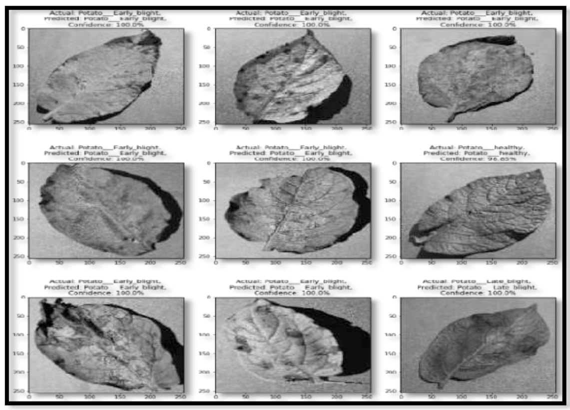


Fig 5. Testing Process

VI. CONCLUSION

The primary goal of this project is to create a machine learning method for detecting disease in potatoes on any surface (CNN). The results of the investigation clearly show that CNN is the most effective method for locating these kinds of items. Nonetheless, the validation accuracy of this model rises to 97.1%. For this study, both recent and historical datasets are incorporated. A project like this might be very beneficial to the world's agricultural industry. The farmers in Bangladeshi communities are mostly illiterate, thus they lack the knowledge needed to combat the disease. The study's findings could improve the lives of Bangladeshi potato farmers. The paper's main flaws are the data used and the accuracy of the model. Increased data can lead to higher accuracy.

VII. FUTURE WORK

We want to refine our model and build a website where anybody can view and submit pictures of the leaves on their plants. For common plant leaf diseases, our model will offer practical recommendations and remedies. It will also provide users with the contact details of the closest agricultural center office, enabling them to speak with experts in agriculture for more precise outcomes.

REFERENCES

[1] Multi-Level Deep Learning Model for Potato Leaf Disease Recognition, Gulshan Shrivastav. IEEE Trans. Anal Pattern. Mach. Intellectual. 2019, 43, 652–662

[2] David Camacho, applications, and software tools; a unique framework for detecting potato leaf disease utilizing an effective deep learning model. In the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15–20, 2019, Long Beach, CA, USA, proceedings, pp. 658–666.

[3] Ankit Jain. Deep Learning for the Identification of Potato Leaf Diseases. Pages 5561–5569 in the Proceedings of the IEEE International Conference on Computer Vision, held September 17–20, 2005 in Beijing, China.

[4] Cecil Zachlod. Artificial Intelligence in Potato Leaf Disease Classification: A Deep Learning Approach. IEEE Internet of Things Journal. 2019, 6, 5531–5539.

[5] Zihang Lin. Potato Disease Detection using Image Segmentation and Machine Learning. Journal of Comput. Eng. Soc. 2021, 2, 62–71.

[6] Massimo Stella. Convolutional neural networks are used to recognize and predict potato leaf diseases. IEEE Sens. J. 21, 17479–17491 in 2021.

[7] Logan, Austin P. Graph cut segmentation is used to identify early and late blight infections on potato leaves. In the 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS) Proceedings, held April 23–26, 2018, in Monterey, California, USA, pp. 1069–1073.

[8] Bioactive Compounds, Antioxidants, and Health Benefits of Sweet Potato Leaves, Kingstone Nyakurukwa, Hawaii, USA, July 21–26, 2017; pp. 7263–7271.

[9] Potato leaf disease prediction using an end-to-end hybrid deep learning framework, Ferdaous Benrouba and PLDPNet. In the IEEE Conference on Computer Vision and Pattern Recognition Proceedings, June 27–30, 2016, Las Vegas, NV, USA, pp. 779–788.

[10] Gebeyaw Astarkie Mulatu. Using a Sequential Model to Identify Potato Leaf Diseases. 2020 IEEE Access, 8, 1–11.

[11] Gebeyaw Astarkie Mulatu. Using a Sequential Model to Identify Potato Leaf Diseases. 2020 IEEE Access, 8, 1–11.

[12] Mr.Lawrence C. Ngugi "Leaf disease recognition using hand crafted features and shallow classifiers" 2020.

[13] Mr.V Suresh, "Plant Disease Detection using Image Processing": a survey," International Journa l of Computer Applications. 2020.

[14] Mitali V. Shewale , "Digital image sensor-based assessment of the status of oat (Avena sativa L.) crops after frost damage," Sensors 11(6), 2023.

[15] A. Ahila V. Prema, Sivasubramanian."Application of support vector machine for detecting rice diseases using shape and color texture features," 2009 International Conference on Engineering Computation, IEEE, Hong Kong, 2009, pp. 79–83.

[16] Abdallah Ali (2019, September) Plant Village Dataset, Version 1, Retrieved 22 February 2020, https://www.kaggle.com/xabdallahali/plantvillage-