Potato Leaf Disease and its Classification Using Deep Learning

Sahil Patil [1], *Aniket Korgaonkar* [2]
Shashank Nadankar [3], *and Archana Ekbote* [4]

Abstract

Potatoes are one of the most extensively consumed foods item, ranking as the 3rd largest staple food consumed throughout the world. Also, the demand for potato is expanding dramatically in the market, particularly due to the worldwide Coronavirus pandemic. However, potato diseases are the major cause of loss in the quality and quantity of the yield. Potato leaf blight is one of the most damaging global plant diseases because it impairs the productivity and quality of potato crop and badly impacts both individual farmers and the agricultural economy. Inappropriate classification and late diagnosis of the disease's type will severely impair the state of the potato plant. This study describes an architecture developed for potato leaf blight classification. This design depends on Deep Convolutional Neural Network (CNN). The methodology also takes use of Data Augmentation. The training dataset is visibly separated into three categories, namely, healthy leaves, early blight leaves and late blight leaves. The number of photos in the collection is 3000. The proposed design achieved an overall mean testing accuracy of 98%. The testing accuracy of the proposed approach was compared with that of comparable works, and the proposed architecture achieved improved accuracy compared to the related works.

Keywords : **Deep Convolution Neural Network, Leaf Disease Classification**

I. INTRODUCTION

Agriculture is an essential sector in countries like India as the economies of these countries are directly or indirectly dependent on agriculture. Using information technology, smart agricultural methods such as automation, precision farming, and eco-friendly farming are urgently needed. Management of crop disease is an important part of Smart Agriculture. Food safety and its nutrition improvement are some of the significant challenges faced by the agricultural sector every year. The presence of pests and diseases of potato plants during the growth period has reduced the quality and quantity of agricultural products. Early disease detection aids damage control and reduces the possibility of crop loss. Potato pests and diseases can lead to an early harvest where the harvest is done when the potatoes are still small, and crop failure is caused by spoilage of potato plants before harvest. These problems are mostly caused by the late identification of diseases in potato plants and errors in disease diagnosis.

Plants and agricultural lands are persecuted by diseases. Microorganisms, genetic disorders, and infectious agents such as bacteria, fungi, and viruses are the prime factors of these diseases. The scientific names of Late Blight and Early Blight are *Phytophthora Infestans* and *Alternaria Solani* respectively, which are significantly responsible for yield losses in potato plant. Preventing these diseases and reducing financial and

Paper Submission Date : June 20, 2023 ; Paper sent back for Revision : July 5, 2023 ; Paper Acceptance Date : July 10, 2023 ; Paper Published Online : August 5, 2023.

[1] S. Patil *(Corresponding Author), Student ;* Email : Sahil.192144105@vcet.edu.in ;
ORCID iD : https://orcid.org/0009-0007-2969-5060
[2] A. Korgaonkar, *Student* ; Email : Aniket.191964110@vcet.edu.in ; ORCID iD : https://orcid.org/0009-0008-5831-8018
[3] S. Nadankar, *Student* ; Email : Shashank.192064105@vcet.edu.in ; ORCID iD : https://orcid.org/0009-0007-8721-1511
[4] A. Ekbote, *Assistant Professor,* ; Email : Archana.ekbote@vcet.edu.in ; ORCID iD : https://orcid.org/0000-0001-8342-3612
[1, 2, 3, 4] Vidyavardhini's College of Engineering & Technology, K.T. Marg, Vartak College Campus Vasai Road, Vasai-Virar, Maharashtra - 401 202.

DOI : https://doi.org/10.17010/ijcs/2023/v8/i4/173264

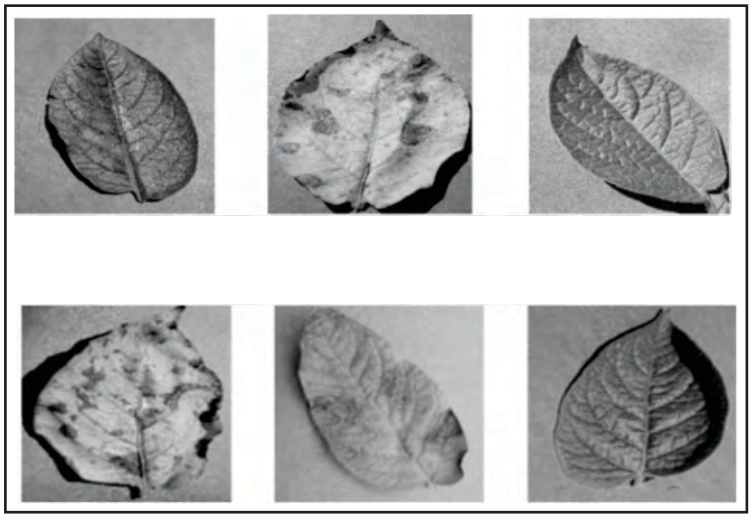

Fig. 1. Images of infected leaves from the dataset used for the project

productional losses is dependent on the early detection of the disease. So, detection and diagnosis of these fungal diseases on such vital crop can possibly increase its yield and farmers can make some profit.

Deep Learning approach focuses on successful classification of healthy leaves and infected leaves. This architecture is based on CNN model. The entire model is built using CNN having several layers. An app is built using React Native which has the model loaded in it. The farmers have to capture the picture of the potato leaf and the model will classify the image as healthy, early blight, and late blight along with percentage of confidence. An app-based approach is very convenient as farmers can easily diagnose the disease at an early stage and save the crop from disease.

II. LITERATURE SURVERY

In [1] the authors of the paper went for the Deep Learning approach. The architecture used in their study is VGG16 which is a pre-trained CNN model and various other architectures such as GoogleNet and ResNet50 which have various accuracies. Their research revolves around classification of two types of potato diseases using these pre-trained models. The models have an accuracy of 95%–97%.

The research article released by Monzurul Islam and Khan Wahid in [2] is a mix of machine learning and image processing for diagnosis of plant leaf disease. The segmentation approach and application of support vector machine exhibits disease categorization of images with an accuracy of 95%. With the support vector machine we extract the features, detect, and diagnose the disease.

Iqbal and Hassan in [3] used an image segmentation approach. In the paper image segmentation is done over 450 photos of healthy and sick potato leaf in the dataset. They made use of seven classifiers method for classification of the leaves. Among them only Random Forest classifier gave the greatest accuracy of 97%.

In [4] the researchers proposed approach decorrelation. Stretching is employed to increase the color differences in the input images. Fuzzy c-means clustering segments disease-affected area, then adds backdrop with the same color characteristics to the photos. Their proposed approach obtains an accuracy of 93% for 27 photos.

The approach presented by Khalifa et al. in [5] depends on 14 layers, comprising two major convolutional layers for feature extraction with variable

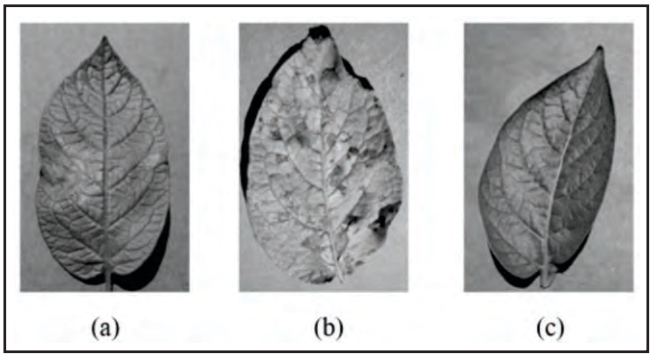

Fig. 2. Sample images of potato leaf (a) Late Blight affected (b) Early Blight affected (c) Healthy

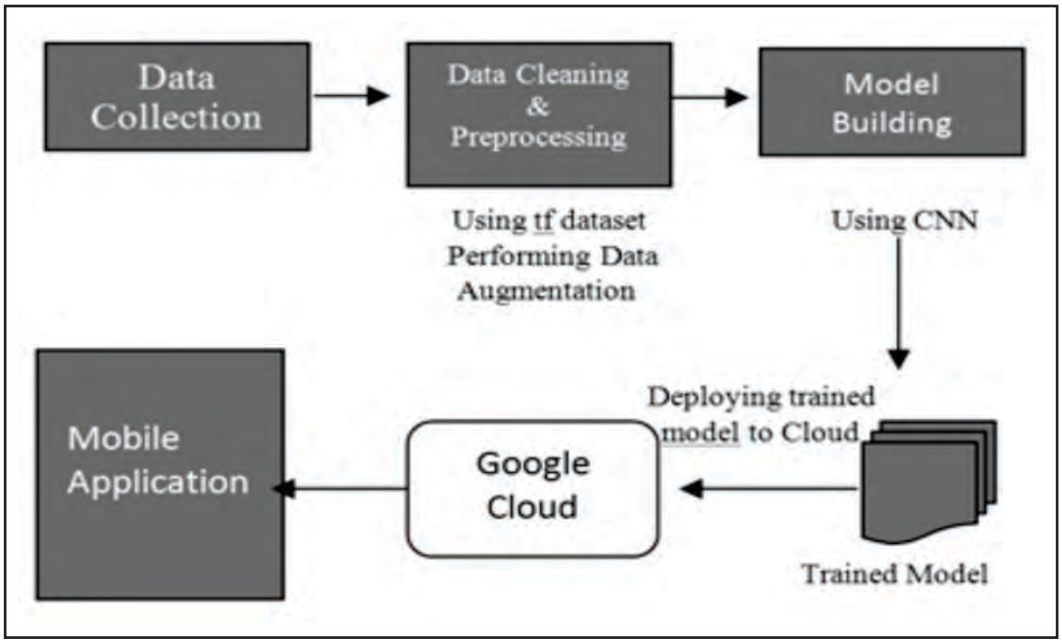

Fig. 3. Flow diagram of the system

convolution window sizes followed by two fully connected layers for classification. In this method, data augmentation process were utilised to expand the amount of data set from 1,722 images to 9,822 images. The proposed approach yields an accuracy of 98%. 6 performance metrics were utilised in this research to assure the correctness of the results acquired.

The work published by Rashid et al. in [6] focuses on creating a Multi-level Deep Learning model. Some of the present models cannot detect crop species and agricultural disease. At the very first stage of their procedure, potato leaves from the potato plant are extracted using YOLOv5, which is an image segmentation technique. Then at the following layer deep learning techniques using convolution neural network were created for detection of early blight and late blight in potato leaves. The collection contained a total of 4,068 photos taken in various areas. Their proposed approach attained an accuracy of 98.75%.

III. DATA SET

The dataset is taken from freely available "Plant Village" (www.plantvillage.org), an image database. It contains 54,306 photos of sick and healthy plant leaves of 14 distinct crop types.

We analyzed only 3,000 images which are divided into 3 class labels:

1) Late Blight potato leaf

2) Early Blight potato leaf

3) Healthy potato leaf

IV. METHODOLOGY

A. Architecture

The flow diagram of the system is given in Fig. 3.

B. Data Pre-Processing

First, we augment the given data to increase the quantity of data. Data augmentation is a process of slightly modifying the original copies of the images of the dataset. It includes geometrics and colour specific transformation like flipping, resizing, and cropping. Data augmentation is used to prevent models from overfitting. The images are flipped horizontally and vertically. The colour, contrast, and lighting of the images can be changed. We just expand the size of our dataset by doing these modifications. After that we create a batch size of 32 images with 3 channels and 50 epochs.

C. Model Building

After data pre-processing, the next step is building the

model for making classifications. The model is built using Convolution Neural Network (CNN) which is a popular Neural Network for solving image classification problems. We create the CNN model using Max Pooling and Conv2D layer. "Relu" is used as an activation function inside the hidden layers.

D. Validation

Our dataset is divided into training, testing, and validation. We fit our model over the training data and find the results, and based on it, the accuracy scores are computed.

E. Mobile Application

After model creation the next step is to deploy our trained model to Google Cloud Functions. We have also created a React Native mobile application. The mobile application will call the function using http and will serve the request. The mobile application is very handy as the farmers must just click the picture of the potato leaf and diagnose whether the plant is infected or it is healthy. The UI of the app is very simple and it is very easy to use. The user has to just upload the picture of the leaf or he can click a picture of the plant by using his mobile camera. The app calls the function and gives the output whether the leaf has late blight or early blight along with the confidence.

V. RESULTS AND DISCUSSION

In our experimental study, a dataset of 3000 images of potato plant leaves were collected and trained against our model. Our dataset consists of 1000 images of early blight, 1000 images of late blight and remaining 1000 images were healthy plant leaves images. In our

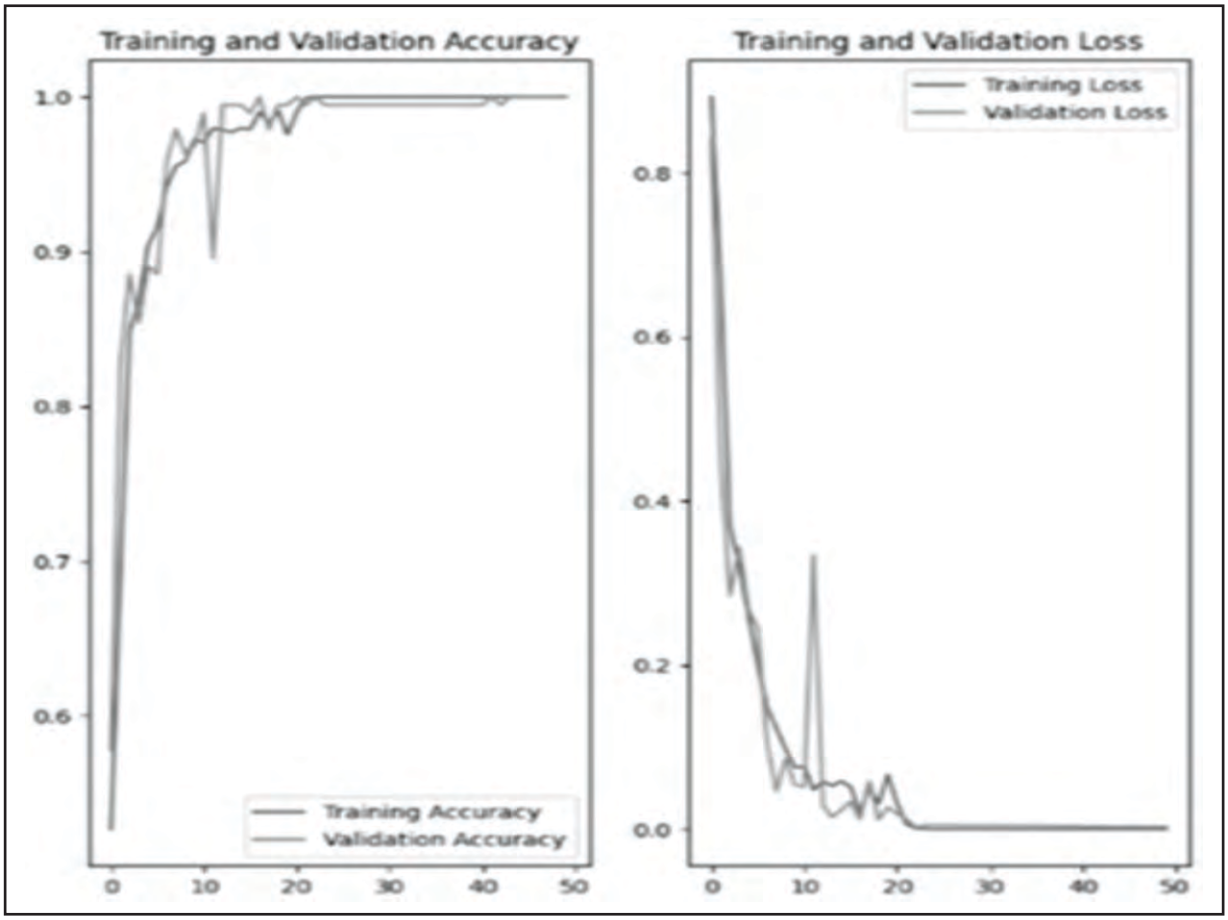

Fig. 4. Train-test accuracy and loss visualization

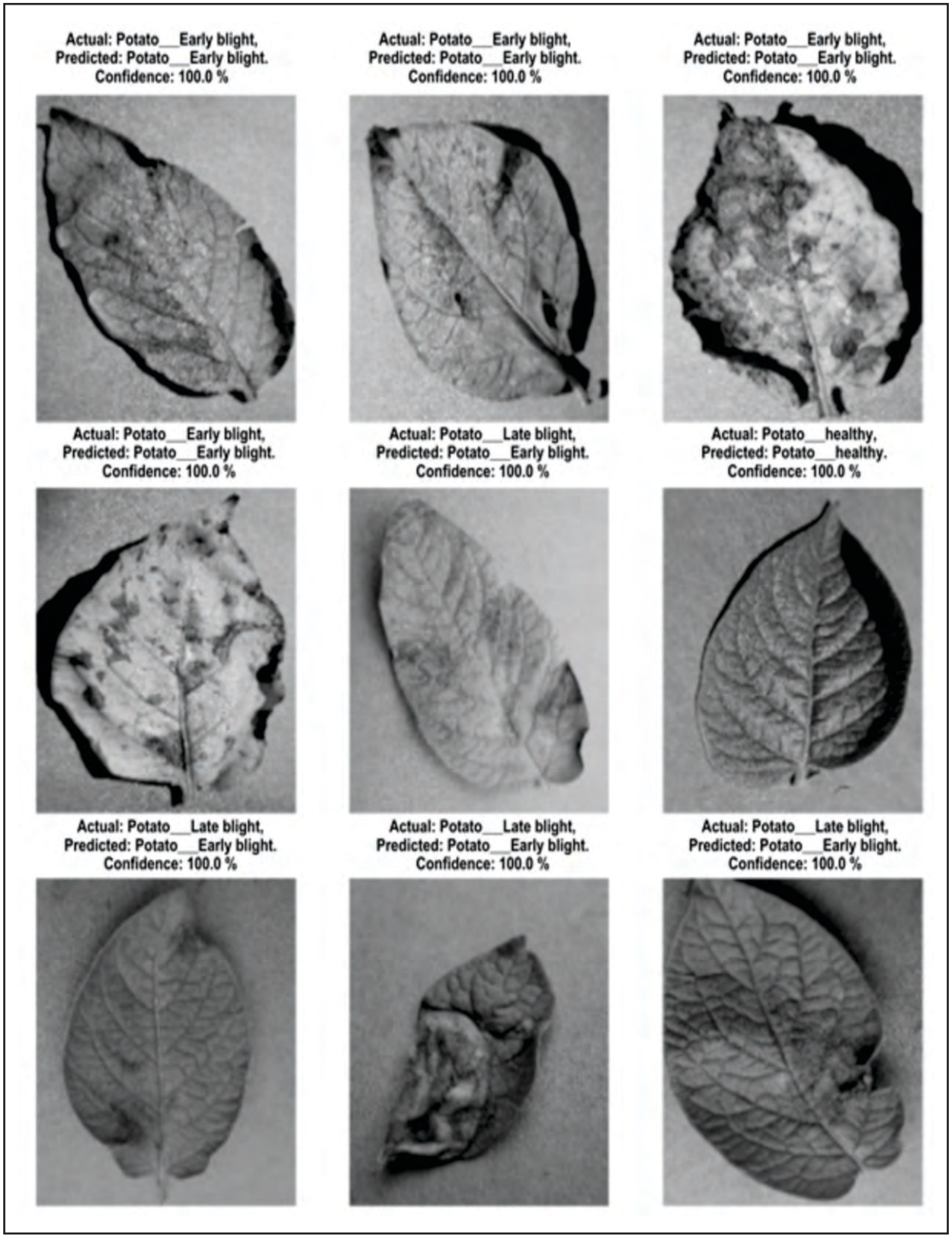

Fig. 5. Results of the model

Fig. 6. The UI of the application

Fig. 7. Model predicting outcome

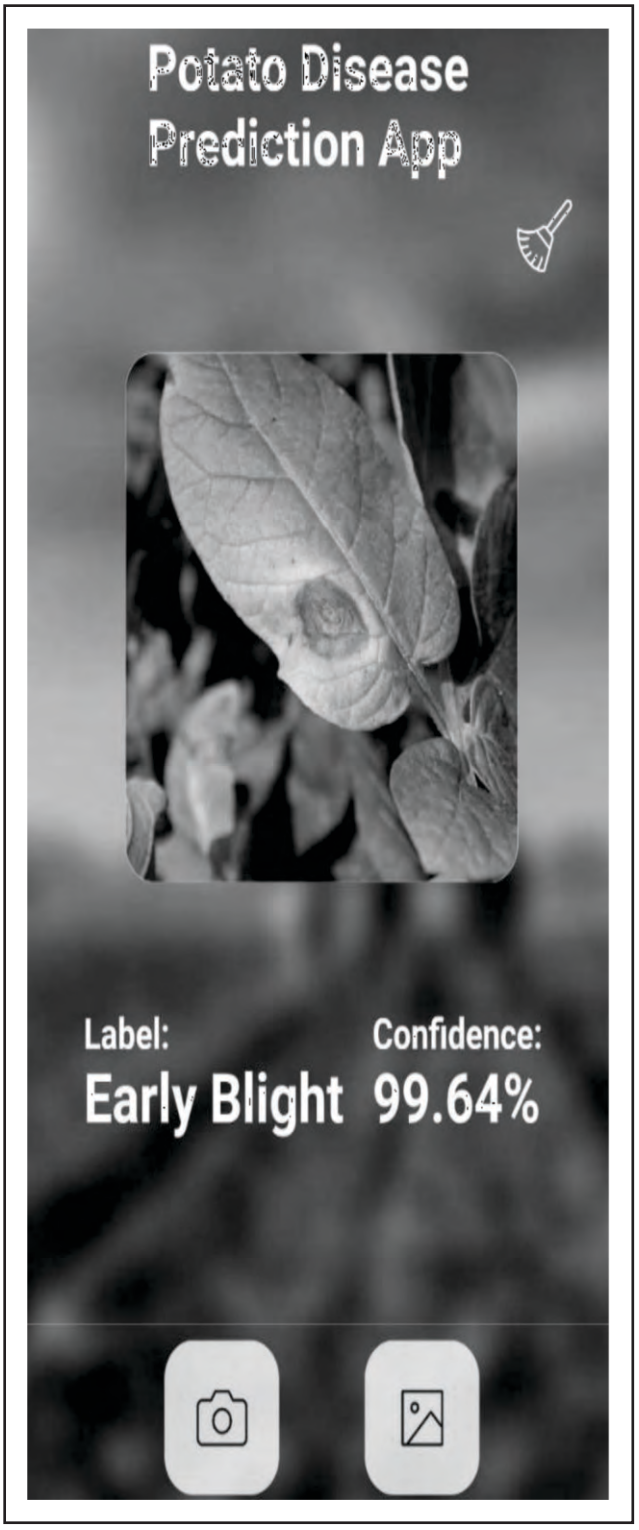

Fig. 8. Early Blight Detection

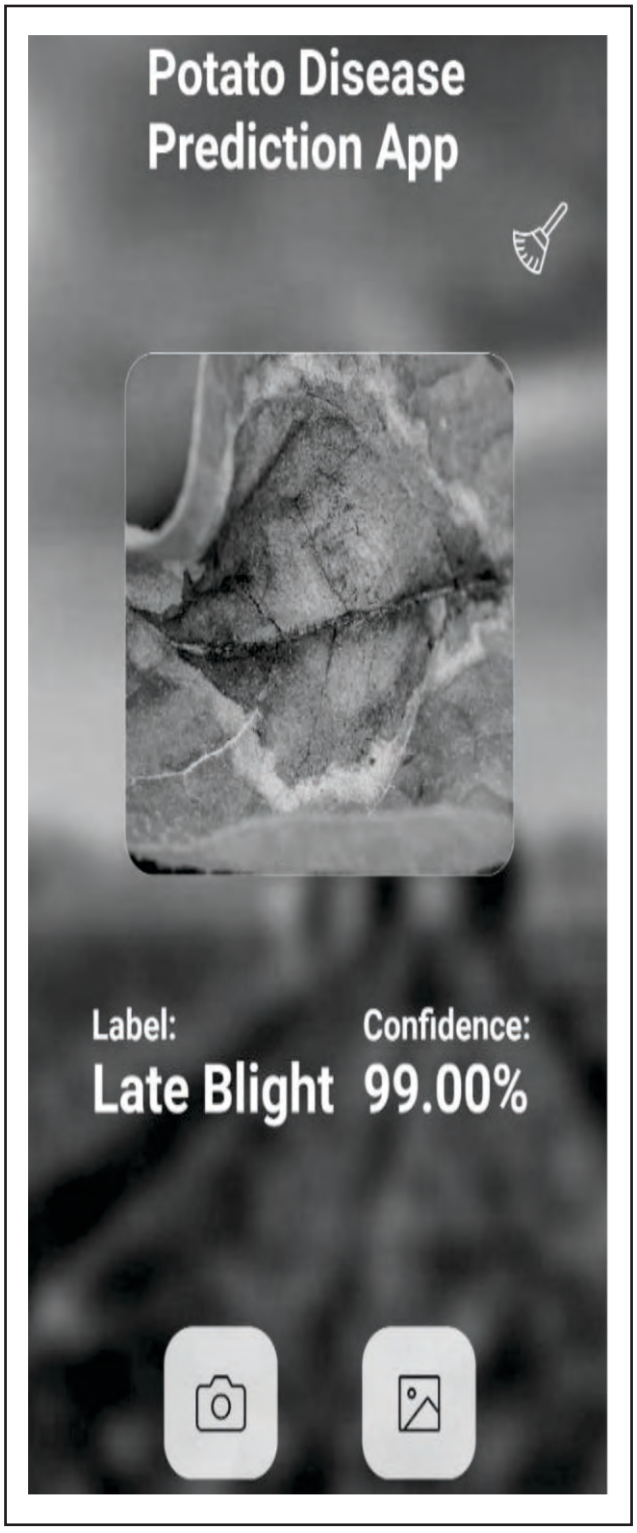

Fig. 9. Late Blight Detection

		Predicted Label		
		Potato _ Early_ Blight	Potato _ Late_ Blight	Potato _ Healthy
True Label	Potato _ Early_ Blight	73	59	68
	Potato _ Late_ Blight	74	65	61
	Potato_ Healthy	68	59	73

Fig. 10. Confusion matrix of the dataset

Classification Report

	precision	recall	f1-score	support
Potato___Early_blight	0.34	0.36	0.35	200
Potato___Late_blight	0.36	0.33	0.34	200
Potato___healthy	0.36	0.36	0.36	200
accuracy			0.35	600
macro avg	0.35	0.35	0.35	600
weighted avg	0.35	0.35	0.35	600

Fig. 11. Classification report of the dataset

experiment, the dataset has been split into training, testing, and validation dataset. The training data comprises of 80% of images of the total dataset, 10% is used for testing, and the remaining is used for validation. At 80-20% train test split, the model shows an accuracy of 98%, which is the best accuracy as compared to accuracies in other papers. VGG16, a pre-trained model performs worst in terms of accuracies. Here, in this paper means the correct classification of the infected area of the leaf.

Also, a confusion matrix is used to for evaluating the performance of the model which we have created. It summarizes and visualizes the performance of the model. In Fig. 10, we can observe the matrix having been predicted and the true labels.

Classification Report measures the quality of the predictions from the given CNN model. The report shows the main classification metrics precision, recall and *F*1 score on a per-class basis.

VI. CONCLUSION

Crop diseases are a widespread hazard to food security across the globe. However, with Artificial Intelligence and powerful deep learning techniques detection and categorization of illness lowers these threats and

eliminates them at a very early stage. In this research, we have employed the concept of Neural Networks and have constructed a model to classify situations in potato leaves. The proposed architecture consists of 14 layers: numerous convolution layers for feature extraction using "ReLu" as an activation function coupled with Max Pooling layer followed by a Flatten layer and a Dense layer containing "softmax" activation function. Augmentation procedures were also employed to boost the value of the dataset. The model attained an overall mean testing accuracy of 99%. Finally, the results were compared with other similar efforts in this sector. Our methodology also involves designing an application which can help farmers in diagnosing plant diseases at an early stage and boosting productivity. Our methodology will give an attainable, timesaving and methodical way of illness classification.

VII. FUTURE SCOPE

In the future we want to extend our research to detection of multiple potato leaf diseases, enhance the "Plant Village" dataset, develop an IoT-based real-time monitoring system.

AUTHORS' CONTRIBUTIONS

Sahil Patil contributed to data preprocessing, including tasks like data augmentation and dataset partitioning. He also played a significant role in constructing the Convolution Neural Network (CNN) and implementing activation functions. Furthermore, he actively participated in the validation process and the development of the mobile application for disease diagnosis.

Aniket Korgaonkar performed an extensive literature review, summarizing existing research on the classification of potato leaf diseases. He conducted comparisons of various models and their respective accuracies, providing valuable insights into the current state of the art research in this domain. Aniket also made contributions to the model development phase by incorporating Max Pooling and Conv2D layers, along with the "Relu" activation function.

Shashank Nadankar played a pivotal role in the collection and management of the dataset sourced from "Plant Village," which comprised 3,000 images categorized into late blight, early blight, and healthy potato leaves. He actively engaged in data preprocessing, including data augmentation and dataset creation. Additionally, Shashank contributed to the training and evaluation of the models.

Dr. Archana Ekbote (Teaching faculty guide) provided guidance and supervision throughout the research project as a mentor. She offered expertise and advice on various aspects of the research, encompassing methodology design, model architecture, and the overall structure of the paper. Dr. Archana's input ensured research's quality and rigour.

CONFLICT OF INTEREST

The authors confirm that they have no affiliations or involvements with any organization or entity having either a financial or non-financial interest in the subject matter or materials discussed in the manuscript. No conflicts of interest are present.

FUNDING ACKNOWLEDGMENT

The authors did not receive any financial support for the research, authorship, or publication of this article. The research was conducted independently without external funding or sponsorship..

REFERENCES

[1] D. Kothari, H. Mishra, M. Gharat, V. Pandey, M. Gharat, and R. Thakur, "Potato leaf disease detection using deep learning," *Int. J. Eng. Res. Technol.*, vol. 11, no. 11, doi: 10.17577/IJERTV11IS110029.

[2] M. Islam, A. Dinh, K. Wahid, and P. Bhowmik, "Detection of potato disease using image segmentaion and multiclass support vector machine," in *2017 IEEE 30th Can. Conf. Elect. Comput. Eng.*, Windsor, ON, Canada, 2017, pp. 1–4, doi: 10.1109/CCECE.2017.7946594.

[3] M. A. Iqbal and K. H. Talukder, "Detection of potato disease using image segmentation and Machine Learning," in *2020 Int. Conf. Wireless Commun. Signal Process. Netw.*, Chennai, India, 2020, pp. 43–47, doi: 10.1109/WiSPNET48689.2020.9198563.

[4] S. Biswas, B. Jagyasi, B. P. Singh, and M. Lal, "Severity identification of Potato Late Blight Disease

from crop images captured under uncontrolled environment," in *2014 IEEE Can. Int. Humanitarian Technol. Conf.*, Montreal, QC, Canada, 2014, pp. 1–5, doi: 10.1109/IHTC.2014.7147519.

[5] N. E. M. Khalifa, M. H. N. Taha, L. M. A. E-Maged, and A. E. Hassanien, "Artificial Intelligence in potato leaf disease classification: A Deep Learning approach," in *Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges. Studies in Big Data*, vol. 77. Dec. 2020, A. E. Hassanien, A. Darwish, Eds., Springer, Cham, doi: 10.1007/978-3-030-59338-4_4.

[6] J. Rashid, I. Khan, G. Ali, S. H. Almotiri, M. A. AlGhamdi, and K. Masood, "Multi-level Deep Learning model for potato leaf disease recognition," *Electron.,* vol. 10, no. 17, 2064, 2021, doi: 10.3390/electronics10172064.

About the Authors

Sahil Patil is a student of Bachelor of Engineering (Information Technology) at Vidyavardhini's College of Engineering & Technology, Vasai. He has a keen interest in the intersection of Computer Science and Agriculture, particularly in the application of Deep Learning techniques for crop disease classifications.

Aniket Korgaonkar is an Information Technology Engineering student at Vidyavardhini's College of Engineering & Technology, Vasai. His research interests encompass Machine Learning, image processing, and their applications in agriculture. He is deeply passionate about pushing the boundaries of AI to solve real-world problems and drive innovation.

Shashank Nadankar is a dedicated Bachelor of Engineering (Information Technology) student. His passion lies in the exciting realms of Artificial Intelligence (AI) and Machine Learning (ML). He has been actively involved in research in these fields, striving to push the boundaries of technology and innovation.

Dr. Archana P. Ekbote received her Ph.D. from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, Maharshtra, India in 2021. She has reviewed many papers for different journals. She has over 20 years of teaching experience. Her research interests include Signal and Image Processing and Pattern Recognition.