International journal of application or innovation in engineering and management

ijaiem.com/Dec 2024/ Volume 13/Issue 2/Article No-1/17-23
ISSN: 2319-4847

Web-Based Convolutional Neural Network-Based Automated Disease Classification for Potato Leaves

K.Rasagna K.Rasagna27@gmail.com,Algeria

Article Info

Received: 17-09-2024 ***Revised: 21 -10-2024*** ***Accepted: 06-11-2024*** ***Published:15/12/2024***

ABSTRACT:
Agriculturally important potatoes are vulnerable to blight and other illnesses. To solve this problem, we suggest classifying photos of potato plant leaves as healthy or sick using a Convolutional Neural Network (CNN). In order to train the CNN to be more generalizable, a larger and more varied dataset is used. An easy-to-use online software that incorporates the best model allows users to submit photos and get instantaneous predictions. With this method, potato crops may be inspected for diseases at an earlier stage, which will allow for faster response times and less crop loss.

KEYWORDS: Classification, data augmentation, late blight, early blight, CNN, web application, potatoes

INTRODUCTION:

Potatoes (Solanum tuberosum) are an essential global food crop, serving as a dietary staple for millions due to their rich content of vitamins, carbohydrates, and minerals. In 2020, global potato production exceeded 370 million metric tons, with India, China, and Russia being the leading producers. Potatoes' versatility makes them a favoured ingredient in various cuisines. However, successful cultivation hinges on plant health, as diseases like Late Blight and Early Blight can severely impact both yield and quality, posing significant challenges for farmers.Early Blight, triggered by Alternaria solani, a fungus, primarily affects leaves which are older, leading to dark, concentric lesions that reduce the plant's photosynthetic efficiency. Late Blight, triggered by the pathogen Phytophthora infectants, is more devastating, often destroying entire fields rapidly under humid conditions. This pathogen was responsible for the 19th-century Irish Potato Famine, underscoring the critical need for effective disease management. Traditional methods for detecting these diseases are labour-intensive, time-consuming, and dependent on subjective judgments, which can lead to delayed responses and increased reliance on harmful chemical fungicides.Current progressions in artificial intelligence (AI) and machine learning provide new solutions for disease recognition. Convolutional Neural Networks (CNNs) have demonstrated their effectiveness in task of image processing and can accurately detect subtle disease symptoms in potato leaves. However, the effectiveness of CNN models relies heavily on the quality and balance of the training data. Many datasets used for plant disease detection are imbalanced, with more images of diseased plants than healthy ones, leading to biased models. To address this, the dataset was expanded and balanced by integrating additional images from various sources, capturing diverse conditions such as lighting and leaf positioning.Although the expanded dataset initially led to a decline in model performance due to increased variability, data augmentation techniques such as flips, rotations, and brightness adjustments were introduced to improve the model's robustness. The project aims to create a reliable CNN-based system for detecting potato leaf diseases, while also exploring optimizers like ADAM and NADAM to enhance:e model performance. The development of a web-basedapplication integrating this model will allow for real-time disease diagnosis, avoiding yield loss and diminishing the need for chemical pesticides, eventually leading to sustainable agriculture.

LITERATURE REVIEW

Pratham Gupta's research[1] utilized Convolutional Neural Networks (CNNs) to recognize diseases in potato plants using images from the PlantVillage dataset. The study aimed to improve classification accuracy by leveraging deep learning techniques. The results demonstrated that CNNs are effective tools for diagnosing potato plant diseases, contributing to improved agricultural disease management. However, the research emphasized the importance of expanding the dataset

International journal of application or innovation in engineering and management

ijaiem.com/Dec 2024/ Volume 13/Issue 2/Article No-1/17-23
ISSN: 2319-4847

in both size and diversity to enhance the model's accuracy and reliability. The authors also suggested exploring alternative deep learning architectures and advanced image processing methods to further improve performance.

Dr. Suman Kumar, Swarnkar [2] developed a system using Convolutional Neural Networks (CNN) to recognize and categorize potato leaf ailments. Approximately 2,000 samples of ailing and healthy leaves of potato were sourced from Kaggle for this purpose. The study aimed to establish a fast, automated method for disease detection through CNN. The process included image preprocessing, training the CNN model, and evaluating its accuracy in distinguishing among ailing and healthy leaves. This CNN model achieved a testing accuracy of 91.41%, outperforming other techniques in detecting potato plant diseases. The research also highlighted the need for more advanced image processing methods and a larger dataset to further enhance the system's accuracy and dependability.Abhinav Baranwal [3] explored the application of deep learning techniques to classify diseases in potato plants. Images of potato leaves affected by various diseases were gathered from publicly accessible sources to train the model. The study aimed to accurately categorize different potato plant diseases using deep learning. The process included collecting, preprocessing the images, and training the model to detect diseases such as leaf spot, late blight, and early blight. This model attained an 85% of accuracy for classification. The research emphasized the need for a larger datasetand suggested experimenting with different deep learning architectures to improve the model's performance and generalizability.Priya Khobragade [4] focused on detecting potato leaf ailments utilizing CNN. The study utilized around 2000 instances of healthy and ailing potato leaves obtained through Kaggle. The objective was to diagnose potato diseases using CNN on leaf pictures. The CNN model achieved an 91.41% accuracy in disease recognition. The study highlights the need for improved accuracy and suggests exploring additional machine learning models to enhance performance.Husnul Ajra [5] introduced a method for detecting plant leaf diseases and offering preventive measures by utilizing AlexNet, ResNet-50 , and two Convolutional Neural Network (CNN) models, along with image processing techniques. The study focused on addressing challenges faced by farmers, particularly those in remote areas, by incorporating digital farming solutions. The methodology was applied to datasets of tomato and potato leaves from Kaggle to detect unhealthy signs. ResNet-50 and AlexNet were used for feature extraction and classification, achieving overall accuracies of 95.3% and 96.5% with AlexNet, and 96.1% and 97% with ResNet-50. These models effectively distinguished between healthy and unhealthy leaves and identified specific diseases. Additionally, a graphical interface provided preventive measures for detected diseases, helping to raise farmers' awareness of plant health.

PROPOSED SYSTEM:

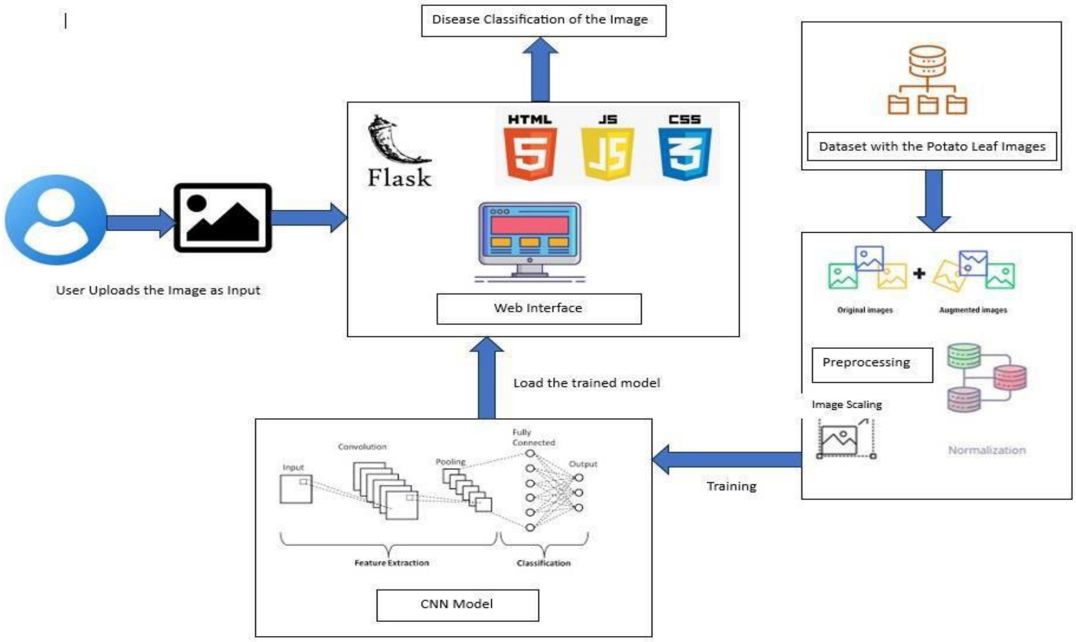
A. Proposed methodology

The proposed methodology involves designing a Convolutional Neural Network (CNN) to sense and classify plant leaf diseases. Initially, a dataset, containing samples for both healthy and ailing leaves, will be collected and pre-processed to ensure consistency in size, quality, and format. Following this, a CNN model will be developed, leveraging its abilityto extract essential features from the images. The model will undergo training using labelled data, where various augmentation techniques, such as rotations and flips, will be applied to improve generalization. After training, the performance of the model will be evaluated using accuracy, precision, recall, and F1 score metrics. To further enhance this system's reliability, optimizers like ADAM or NADAM will be employed to fine-tune the model. Additionally, the system will be integrated into a web application, enabling real-time disease detection for farmers by allowing them to upload leaf images for immediate analysis and classification.

Proposed architecture:

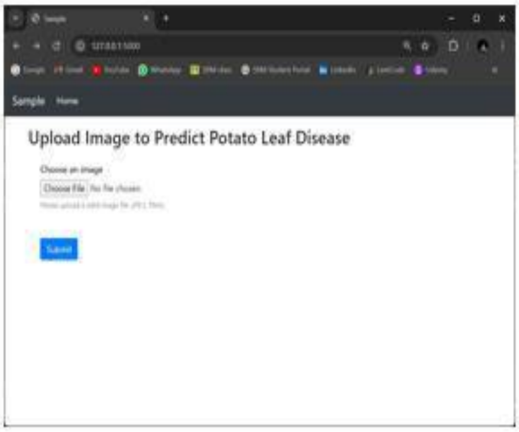
The proposed architecture for the potato leaf disease classification system integrates several components to ensure efficient data processing and accurate predictions, refer Fig. 3.1. At its core, the system uses a Convolutional Neural Network (CNN) constructed to categorize images of leaves of potato into three: "Early Blight," "Late Blight," and "Healthy" categories. This architecture leverages the strengths of deep learning to effectively analyse and learn from the images provided, facilitating precise disease identification.

ijaiem.com/Dec 2024/ Volume 13/Issue 2/Article No-1/17-23
ISSN: 2319-4847

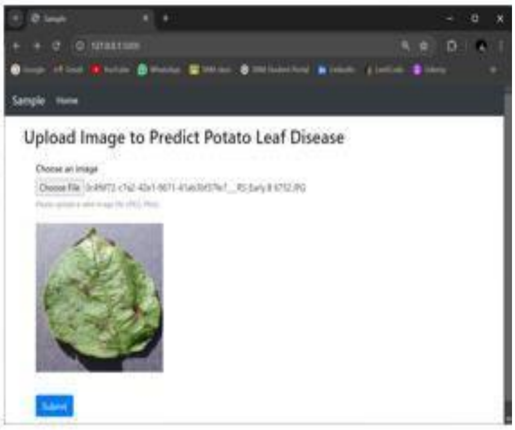


To enhance the classification performance of the CNN model, the initial dataset created by Hafiz Nouman, which had a limited number of healthy potato leaves, only 152 images, was significantly expanded. By incorporating additional datasets from Kaggle and images obtained through web scraping, the dataset was increased to include 2,000 instances for both "Early Blight" and "Late Blight," and 1,964 instances for "Healthy." This expansion aimed to improve the diversity of the dataset, which is crucial for enhancing the model's generalization capabilities. The broader range of images now better represents the variability of conditions under which potato leaves may exhibit disease symptoms.

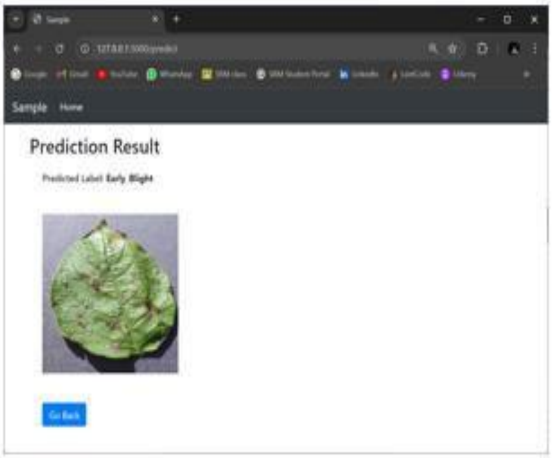
Despite this expansion, the initial evaluation of the extended dataset revealed a decline in performance metrics, compared to the original dataset. This drop was primarily attributed to the increased complexity and variability introduced by the additional images, which posed challenges to the model's capability to generalize effectively across dissimilar leaf conditions and appearances.To mitigate this decline, data augmentation techniques were applied to the expanded dataset. These techniques includedhorizontal and vertical flips, rotations, zooms, and brightness adjustments, simulating various real-world conditions anddiversifying the training data. By artificially increasing the dataset size and variety, data augmentation helped improve the robustness of the model, allowing it to better handle the variability introduced by the expanded dataset.The architecture also features a user-friendly web-based application that allows farmers and agricultural professionalsto easily upload leaf images for classification, refer Fig 3.2. Built on the Flask framework, the backend of thepplication processes the uploaded images and invokes the trained CNN model for prediction. The frontend, developed using HTML, CSS, and JavaScript, ensures an intuitive and accessible experience for users, allowing them to receive immediate feedback and disease identification. The application aims to empower users by providing a practical tool for early disease detection, thereby enhancing their ability to manage crop health effectively.



(a) HOME PAGE



(b) UPLOAD THE IMAGE AS INPUT



(c) THE PREDICTION SHOWN ON THE RESULT PAGE

FIG. 3.2 WEB APPLICATION

Overall, the proposed architecture encompasses a comprehensive approach that integrates advanced machine learning techniques, effective data handling strategies, and a user-friendly interface. By uniting these elements, the system equips farmers and agricultural professionals with timely and actionable insights for disease recognition and supervision, eventually contributing to upgraded agricultural productivity and sustainability.

B. Algorithm Used:

Step 1: User Image Input: User uploads the image of a potato leaf on the website. In the backend the image is pre-processed so it is in the proper form to be given as input to the CNN model.

Step 2: Load the pretrained CNN Model: The best performing CNN model on the expanded dataset is saved. Load the model.

Step 3: Make Prediction: Give the pre-processed image as input to CNN to make prediction.

Step 4: Display the Result: Once the prediction is done, display the results to the User on the Result Page.

International journal of application or innovation in engineering and management

ijaiem.com/Dec 2024/ Volume 13/Issue 2/Article No-1/17-23
ISSN: 2319-4847

RESULTS

The performance of the potato leaf disease classification model was evaluated across three stages: the original dataset, the expanded dataset, and the improved results following data augmentation.

Initially, the model was trained using the original dataset, which contained a limited number of images, particularly in the "Healthy" category. Despite this limitation, the model achieved reasonably good performance metrics, including accuracy, precision, recall, and F1 score, as seen in Table 4.1. The results demonstrated the model's capability to effectively classify potato leaf images despite the relatively small dataset.

TABLE 4.1: MODEL PERFORMACE FOR THE ORGINAL DATASET

Optimizer	Accuracy	Precision	Recall	F1 Score
ADAM	0.9655	0.9653	0.9655	0.9653
SGD	0.9526	0.9521	0.9526	0.9522
RMSPROP	0.9526	0.9529	0.9526	0.9515
ADAGRAD	0.9397	0.9377	0.9397	0.9372
NADAM	0.9698	0.9697	0.9698	0.9696

To enhance the dataset, additional images were sourced from Kaggle and web scraping, significantly increasing the total size. However, this expansion led to a decline in the model's performance metrics across various optimizers, as shown in Table 4.2. The increased variability and complexity introduced by the new images posed challenges to the model's generalization ability, resulting in lower accuracy and other performance measures compared to the original dataset.

TABLE 4.2: MODEL PERFORMACE FOR THE EXPANDED DATASET

Optimizer	Accuracy	Precision	Recall	F1 Score
ADAM	0.9028	0.9041	0.9028	0.9025
SGD	0.9125	0.9128	0.9125	0.9125
RMSPROP	0.8882	0.8898	0.8882	0.8878
ADAGRAD	0.8882	0.8884	0.8882	0.8883
NADAM	0.9141	0.9155	0.9141	0.9141

To counteract the performance decline witnessed with the expanded dataset, data augmentation techniques were applied. These techniques, which included horizontal and vertical flips, rotations, zooms, and brightness adjustments, diversified the training data and improved the model's robustness. As depicted in Table 4.3, the application of data augmentation led to a notable recovery and improvement in the model's performance metrics, bringing them closer to or even surpassing those achieved with the original dataset.

TABLE 4.3: MODEL PERFORMACE FOR THE EXPANDED DATASET WITH DATA AUGMENTATION

Optimizer	Accuracy	Precision	Recall	F1 Score
ADAM	0.9698	0.9699	0.9698	0.9697
SGD	0.9337	0.9337	0.9337	0.9336
RMSPROP	0.9547	0.9550	0.9547	0.9547
ADAGRAD	0.7114	0.7367	0.7114	0.7059
NADAM	0.9681	0.9681	0.9681	0.9681

In conclusion, while the transition from the original to the expanded dataset initially resulted in reduced performance, the strategic application of data augmentation successfully enhanced the model's capability to categorize diseases of potato leaf. This demonstrates the importance of these techniques in improving machine learning model outcomes.

CONCLUSION AND FUTURE WORK:

The CNN model for classifying potato leaf ailments represents significant progress in agricultural technology, effectively identifying leaf samples as "Late Blight," "Healthy," or "Early Blight." By integrating this model into an intuitive web application, farmers and agricultural professionals can take timely actions to reduce crop losses and enhance overall yield.Looking ahead, there are opportunities for further enhancement. Implementing advanced preprocessing techniques and exploring sophisticated CNN architectures like ResNet could improve model performance on the expanded dataset. Additionally, broadening the model to detect more potato diseases and incorporating IoT sensors for real-time data collection would create a comprehensive crop management tool, providing valuable insights to farmers and boosting agricultural productivity.

REFERENCES:

[1] P. Gupta, H. Waghela, S. Patel, N. Rajgor, S. Sange, and V. Korade (1923). Disease Classification in Potato Plants with the Use of a Convolution Neural Network. Page 1–6, 2023 ICACTA, 2023 International Conference on Advanced Computing Technologies and Applications, Mumbai, India. The DOI for this article is 10.1109/ICACTA58201.2023.10792414. 2. Swarunkar, the doctor, Arijeet Dutta, Abhisek Chakraborty, and Nachiketa Mohanty. Using Convolutional Neural Networks for Disease Detection and Classification in Potato Leaves." Thirdly, Yogesh, A. Goyal, M. Mishra, and A. Baranwal (2022). Deep Learning for Disease Classification in Potato Plants. Pages 673–671 from the 2022 COM-IT–CON International Conference on Machine Learning, Big Data, Cloud, and Parallel Computing in Faridabad, India. This article has the DOI: 10.1109/COM-IT-CON54601.2022.9850701. 4. In a study conducted by P. Khobragade, A. Shriwas, S. Shinde, A. Mane, and A. Padole in 2022. Potato Leaf Disease Detection Using CNN. Page 1–5 of the 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON) held in Bangalore, India, 2022. The conference proceedings may be found at DOI: 10.1109/SMARTGENCON56628.2022.10083986. 5. In a 2020 publication by H. Ajra, M. K. Nahar, L. Sarkar, and M. S. Islam. The Use of Convolutional Neural Networks (CNNs) and Image Processing for the Prevention of Damage to Plant Leaves. 2020 ETCCE, Bangladesh, pages 1-6, Emerging Technology in Computing, Communication and Electronics. Achieving a DOI of 10.1109/ETCCE51779.2020.9350890