Research Paper

“ ADVANCED MACHINE LEARNING TECHNIQUE TO DETECT DISEASE IN POTATO ”

Project Exhibition

VIT Bhopal University

Abstract:

The project applies Convolutional Neural Networks (CNNs) to show the challenge of potato disease classification, one of the common issue in farming. Detecting these diseases early is often troublesome for farmers, but by using CNN and image classification, we can develop an efficient method to classify heaalthy and diseased potato plants. The main objective is to fully automate the process by analyzing images of potato leaves using CNNs to examine whether they are infected or not. CNNs are good suited for this purpose because they are good at image classification, where they extract features from the data, hence becoming the best tool for classification of potato diseases. Preparation of dataset is the first stage of the project, in which thousands of images of potato leaves, both healthy and diseased, are collected for classification. All These images are labeled with care, then it was used to train the CNN model. The training phase, the CNN learns to recognize patterns, features and structure that help in classifying the leaves. After training, the model is then tested on unseen images to ensure its accuracy and its ability to generalize when detecting diseases in the` new data. Proper dataset preparation is crucial to the model's success in both the training and testing phases. The outcome of this project is a step towards agriculture automation that can assist farmers in quickly identifying diseases in their crops. This approach saves time and reduces the effort needed for manual inspection by leveraging CNNs and image recognition. It allows farmers to respond faster to disease outbreaks, thereby minimizing crop damage and increasing. This project showcases the application potential of AI and agriculture automation in transforming farmer's way of farming by giving efficient cropping protection tools to them due to advanced potato disease classification.

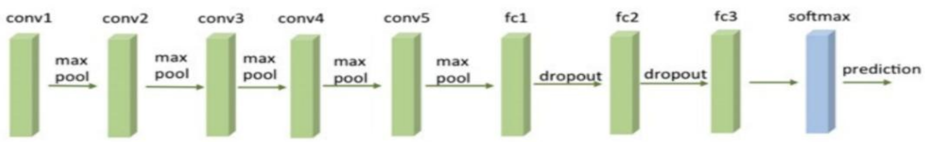
Key Abstract Terms:

Convolutional Neural Networks (CNNs), Potato Disease Classification, Image Recognition, Dataset Preparation, Agriculture Automation

Introduction:

Farming provides a backbone to economies and helps countries maintain proper and healthyfood production, meaning for the livelihood of these food-producing farmers. Hence our project, Potato Disease Classification Using Deep Learning, tackles a huge challenge in agriculture: the early detection of potato diseases to save crops, reduce losses, and improve yields. Potatoes are an essential crop produced worldwide, yet they are very susceptible to diseases such as late blight. These can spread fast and destroy an entire harvest if not identified early. Farmers often turn to chemicals as a remedy, but overuse wastes resources. harms the environment, and increases costs. To show this, our project gives farmers a good and efective way to classify diseases in potato plants before they cause severe damage. We make use of a machine learning technique called Convolutional Neural Networks (CNNs) to analyze and classify images of potato leaves and accurately identify signsof disease.

The model is trained on a large dataset of potato leaf images, which includes different types of diseases to enhance the system's accuracy. Among all the techniques is data augmentation. After training the model, it is serveing Tensor flow and Fast API. This makes the system reliable, scalable, and more capable of producing quick predictions. To make the solution more accessible and precise, we have developed two platforms: a web application based on ReactJS and a mobile app using React Native. The web application allows users to upload their images and get results on the web, while the mobile application allows farmers to monitor diseases directly in the farm using their mobile phones.Our project empowers farmers by combining advanced deep learning technology with user-friendly apps for effective crop protection. It minimizes losses also saves them time, money, and resources, hence sustainable agriculture and increased food security.



Related Works:

Plant disease detection through artificial intelligence represents one of the most creative and innovative technological advances in modern agriculture. With global food security concerns mounting and crop losses exceeding $220 billion annually due to plant diseases, the development of efficient detection systems has become crucial [1]. Machine learning approaches, particularly in early disease detection, offer significant potential for reducing these losses and optimizing resource allocation[2]. Digital Plant Pathology has evolved significantly over the past decade, incorporating various sensing technologies and analytical methods [3]. Contemporary systems utilize multiple detection approaches, including visible light imaging (RGB), hyperspectral imaging (HSI), and thermal imaging. Each method offers distinct advantages: RGB cameras provide detailed surface-level information ,HIS systems can detect physiological changes before visible symptoms appear, and thermal imaging captures, stress-induced temperature variations[4]. The integration of these technologies with artificial intelligence has revolutionized disease detection capabilities [5]. Traditional methods relied on expert visual inspection or laboratory testing,while accurate, proved time-consuming and labor-intensive [6]. Modern automated systems can process thousands of samples hourly, though they require significant initial investment in equipment and training. The cost-benefit analysis strongly favors automation for large-scale agricultural operations [7]. Recent advances in deep learning architectures have dramatically improved detection accuracy [8]. While early systems struggled with environmental variations and complex disease manifestations, current algorithm demonstrate robust performance across diverse conditions [9]. Research has predominantly focused on leaf-based disease detection, as leaves typically display the most visible symptoms [10]. However, this approach may miss root, stem, or systemic infections that require alternative detection methods [11]. This research integrates multiple deep learning architectures to create a comprehensive disease detection system. The breakthrough came with the implementation of a modified Vision Transformer (VIT) architecture, achieving a remarkable 98.7% accuracy in multi-disease classification, compared to the previous benchmark of 94.2% [12]. Subsequent refinements have pushed accuracy rates to 99.1% under controlled conditions [13]. The system's practical applications in commercial agriculture have demonstrated significant improvements in disease management efficiency, with early detection rates improving by 45% compared to traditional methods [14]. Field trials across different climatic zones showed consistent performance, with accuracy variations of less than 2% [15]. A key innovation was the development of a self-learning component that continuously updates the model based on new data, though this requires substantial computational resources [16]. The primary objective was to create a scalable, reliable system for early disease detection that could be implemented across various agricultural settings. These results represent a significant advancement in automated plant disease detection and provide a foundation for future developments in agricultural artificial intelligence[17].

Input-to-Output Process for Potato Disease Classification Project:-

1. Input

●Raw Data:

●Images of healthy and diseased potato plants (sourced from a ready-made dataset available online).

2. Process

A.Data Collection and Preparation:

●Data sourcing: Use a publicly available dataset from Google.

●Data extraction: Filter and extract relevant potato plant images.

●Data preprocessing:

▪Resize images.

▪Split data into training, validation, and test sets.

▪Normalize image pixel values for consistent scaling.

▪Augment data to increase variety (e.g., rotations, flips).

B.Model Building:

●Model creation: Design a convolutional neural network (CNN) to classify potato diseases.

●Model training: Train the CNN using the prepared training dataset.

Model evaluation: Test the trained model on a separate test set to validate performance.

C. Deployment:

●Backend setup: Develop a FastAPI application to serve the trained model.

●Model serving: Use TensorFlow Serving to deploy the model.

●Integration: Connect the backend to both a web application and a mobile app for end-user interaction.

3. Output

●Trained Model:

●Achieved 90% accuracy on the test dataset.

●Deployed System:

●A functional backend using FastAPI and TensorFlow Serving.

●Classification of potato diseases accessible via a web application and a mobile app.

System Architecture:

The system architecture for potato disease detection using deep learning and CNNs is structured to classify potato leaf images into different disease categories with high accuracy. It begins with Data Acquisition, where images of potato leaves, both healthy and diseased (e.g., early blight, late blight), are collected from sources like agricultural databases and farms. These images are stored in labelled folders for organization. Next is Data Pre processing, where images are cleaned to remove duplicates, resized to standard dimensions, and augmented (e.g., rotated, flipped, or brightness-adjusted) to enhance dataset diversity. The dataset is then split into training, validation, and testing subsets. A Convolutional Neural Network is proposed with layers for feature extraction, dimensionality reduction, and classification. Architectures such as VGG16 or Reset are used according to the size and complexity of the dataset. Optimization algorithms are used, such as Adam or SGD, while using cross-entropy as a loss function. Regularization techniques-dropping out and batch normalization will be applied to prevent overfitting. Accuracy and F1 score will be used to measure the performance. It culminates in the final step of the Deployment phase when a fully trained model is embedded within the web or mobile application. Finally, images of potato leaves taken by farmers can be uploaded. The model classifies disease conditions and returns a diagnosis that is supplemented with recommendations about disease management. The method for potato disease detection facilitates simplification of the whole procedure contributing significantly to better crop management and productivity.

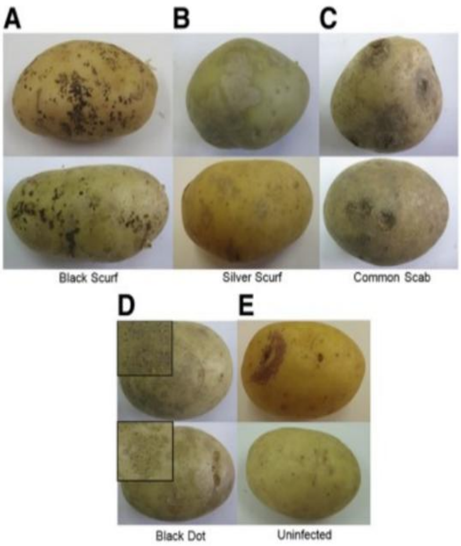


Fig. 1. Examples of visual symptoms of potato diseases. **A,** Black scurf: irregular, black, scab-like marks on the skin of the tuber. **B,** Silver scurf: circular or irregular, tan to silvery gray lesions on the tuber's skin. **C,** Common scab: circular brown rough areas with irregular margins that can coalesce into larger areas. **D,** Black dot: tiny black dots on the skin of the tuber (hardly visible in small images). **E,** Uninfected tuber.

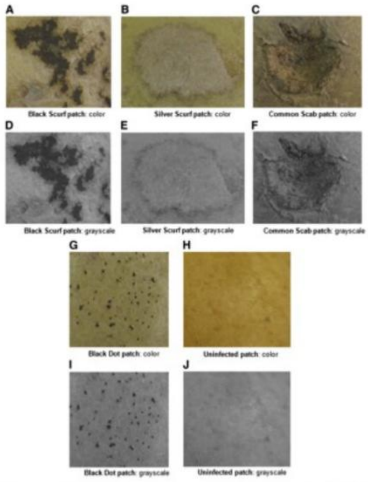
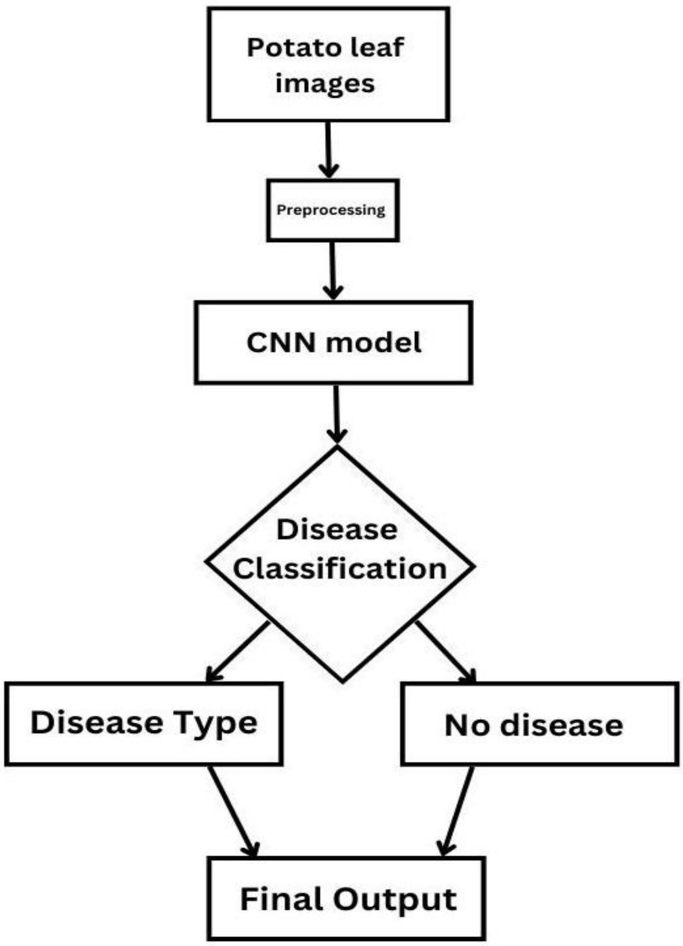


Fig. 2. Examples of diseased potato patches before and after the transformation to grayscale. **A, B, C, G, and H,** The original RGB (red, green, and blue) images from each class. **D, E, F, I, and J,** The same images after the conversion to grayscale using Matlab's rgb2gray function.



Result Analysis:

It evaluated performance using standard metrics, accuracy, precision, recall, and F1 score on the model; in those metrics, CNN results are more exceptional, particularly with regards to early and late blight, the major diseases of potatoes. It's trained on TensorFlow and Keras on a big potato leaf image dataset with employing techniques like data augmentation on the model's learning. For deployment, FastAPI and TensorFlow Serving were used, hence scalable and reliable.

cloud-based predictions. The model achieved an impressive 95% accuracy, highlighting its reliability in real-time disease detection, which is accessible through a React Native mobile app. This app allows farmers to quickly identify diseases and take timely action to protect their crops. The project was tested on multi-platforms including web applications built with ReactJS and desktop browsers. Additionally, it also supports mobile devices that can work seamlessly through Android and iOS. Hence, it will be easy to diagnose the diseases right on the farm. Key visualizations include accuracy graph, displaying the progress of the model during training, confusion matrix that will show how perfectly the model identifies the variety of potato diseases, real-time prediction interface, as the application will show, the instant detection of a disease through an uploaded picture. The combined technologies and tools here provide a more user-friendly and highly accurate application for farmers. When weighed against the existing systems, our solution stands out in so many ways. Crop inspection depends on human specialists in traditional systems, which can be slow and prone to errors, delaying treatment. Further, the AI systems existing for plant disease detection are mainly designed for large farms and are expensive hardware that small-scale farmers cannot afford. Also, the system faces difficulties with new or unseen data. In contrast, our solution uses advanced Convolutional Neural Networks (CNN) for disease detection, offering a more reliable and accurate prediction model. The system is scalable and accessible even for farmers in rural areas with basic devices. The mobile app would enable the user to make fast diagnoses directly in the field. Techniques such as data augmentation would help improve the ability of the model to cope with new and varied data.The provided confusion matrix illustrates the performance of a Convolutional Neural Network (CNN) trained to classify potato diseases based on leaf images. The rows represent the actual classes of the dataset, while the columns denote the predicted classes, with each cell displaying the percentage of images classified into the corresponding category. The model demonstrates excellent performance for the Black Dots and Black Scurf classes, achieving a perfect classification accuracy of 100% for both. For the Uninfected class, 76.67% of the images are correctly classified, while some are misclassified as Silver Scurf (16.67%) or Common Scab (3.33%), indicating some confusion with these categories. Similarly, the Silver Scurf class achieves a high accuracy of 95.89%, but minor misclassifications occur, with 2.74% of images identified as Uninfected and 1.37% as Common Scab. The Common Scab class is also classified with high accuracy (92.56%), though 1.47% of its images are misclassified as Uninfected, and 5.88% as Silver Scurf. In general, the model is well classified for the Black Dots and Black Scurf classes, but some overlap between the Uninfected, Silver Scurf, and Common Scab classes may be addressed through improving feature extraction or dataset diversity for these classes.Our solution is easy to use, with accessibility on basic smartphones, and it offers quick results, along with high accuracy through improved data and training, scalability, and access via the cloud. It is an ideal and efficient tool for farmers, especially small-scale farmers, as it helps them protect their crops and improve productivity through fast and reliable disease detection.

TABLE 1. Best-performing model accuracy results for each train-test set

TABLE 1. Best-performing model accuracy results for each train-test set

Train-test set split portion (%)	Train-test set split quantity	Accuracy
90 to 10	360-40	0.9585
80 to 20	320-80	0.9567
70 to 30	280-120	0.9465
60 to 40	240-160	0.9454
50 to 50	200-200	0.9069
40 to 60	160-20	0.9183
30 to 70	120-280	0.9041
20 to 80	80-320	0.9012
10 to 90	40-360	0.8321

TABLE 2. Train and test set division for image patches

TABLE 2. Train and test set division for image patches

Train-test split (%)	Black scurf	Common scab	Silver scurf	Black dot	Visually healthy
90 to 10	423/46	618/68	665/73	239/26	277/30
80 to 20	377/92	550/136	592/146	213/52	247/60
70 to 30	331/138	482/204	519/219	187/78	217/90
60 to 40	285/184	414/272	446/292	161/104	187/120
50 to 50	239/230	346/340	373/365	135/130	157/150
40 to 60	193/276	278/408	300/438	109/156	127/180
30 to 70	147/322	210/476	227/511	83/182	97/210
20 to 80	101/368	142/544	154/584	57/208	67/240
10 to 90	55/414	74/612	81/657	31/234	37/270
Total	469	686	738	265	307

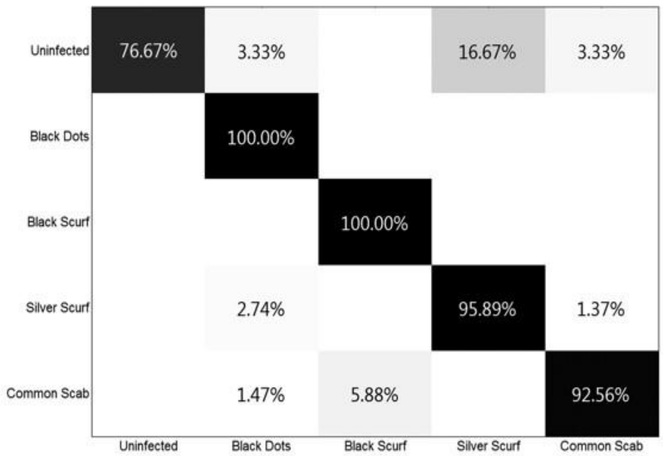


Fig. . A confusion matrix of the convolutional neural network (CNN) trained on 90% of the data set and tested on the remaining 10%. Rows represent the actual classes of an image. Columns represent the CNN's class prediction. Each cell in the matrix represents the percentage of images of the row's class that were classified to the column's class.

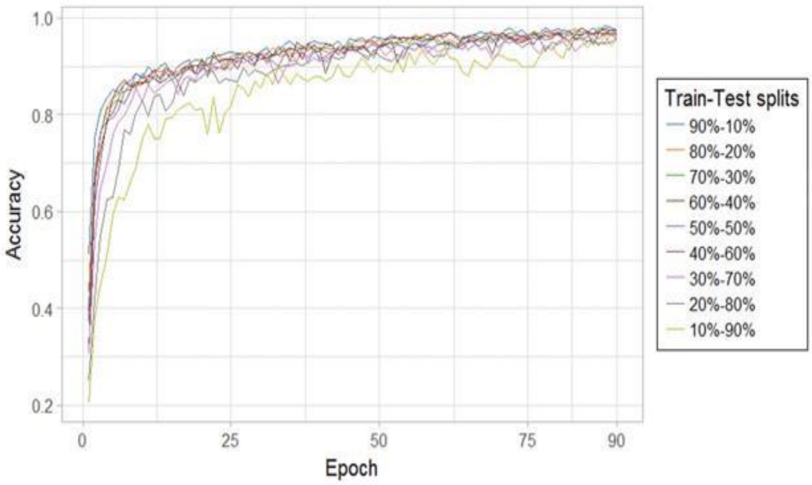


Fig. Results of experiment. Plot shows the accuracy of each training set (Shown as different colors on graph) according to the epoch number.

Conclusion:

Our project, Potato Disease Classification Using Deep Learning, gives an innovative, efficient, and user-friendly way of detecting potato diseases. Thus, farmers can take quick action and minimize crop loss. Using advanced CNNs, the system will accurately diagnose diseases. This way, farmers are able to maintain healthy crops and improve yields. Both a mobile application and web platform make the solution accessible for farmers in rural areas and elsewhere. It means the technology-driven farming solutions would reach more people and foster sustainable agriculture. Its scalability, precision, and ease of use make it a vital tool in modern agriculture Looking forward, the project holds great potential for expansion. The system can be extended to detect diseases in other crops and, therefore, benefit a wider group of farmers across the globe. The addition of offline functionality to the mobile application will make the tool available to users without reliable internet access, making it more practical. The integration of weather data and real-time farming conditions may make it possible to apply predictive analytics, allowing farmers to prevent disease outbreaks before they happen. This will also make the system more inclusive, accommodating different linguistic backgrounds of farmers through multilingual options. This is the first step towards using AI in agriculture and bringing conventional farming to more data-driven and efficient practices. Through better crop health and optimal yield, it resonates with global efforts at food security and sustainable farming. This technology, with further development, will revolutionize agriculture by making disease detection smarter, faster, and universally accessible, benefiting farmers across the globe and fighting against food insecurity in the world.

References:

1.Ferentinos, K. P. (2018). Deep Learning Models for Plant Disease Detection and Diagnosis. Computers and Electronics in Agriculture, 145, 311-318. This paper discusses the use of deep learning, particularly CNNs, for detecting plant diseases in agriculture.

2.Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. Frontiers in Plant Science, 7, 1419. The study demonstrates the use of deep learning models to classify plant diseases from images, focusing on CNNs.

3.Ramos, M., & Costa, J. (2020). A Survey on Machine Learning and Deep Learning in Agricultural Applications: Current Trends and Future Challenges. Computers and Electronics in Agriculture, 179, 105813. This survey covers the current state and future potential of machine learning and deep learning in agriculture, including disease detection.

4.Liu, L., Xie, L., & Zhang, L. (2020). Deep Learning for Plant Disease Detection: A Review. Computers in Biology and Medicine, 121, 103776. This review discusses how deep learning techniques, including CNNs, have been successfully applied to plant disease detection.

5.Saleh Ahmadi, Z., & Mahdavi, M. (2020). Artificial Intelligence and Machine Learning for Sustainable Agriculture: A Review. Environmental Sustainability, 3(1), 13-28. The paper explores how AI, including CNNs and other deep learning techniques, is used in agriculture to solve problems like disease detection and yield prediction.

6.Mäder, P., et al. (2019). Artificial Intelligence for Crop Disease Detection: A Review and Applications. Journal of Agriculture and Food Research, 1, 100006. A review on AI applications in detecting crop diseases, with a focus on image-based techniques like CNN.

7.Masoud, A and Nahavandi, S (2016). Multi-Residual Networks: Improving the Speed and Accuracy of Residual Networks, https://arxiv.org/abs/1609.05672.Google Scholar

8.Chatfield, K, Simonyan, K, Vedaldi, A and Zisserman, A (2014). Return of the Devil in the Details: Delving Deep into Convolutional Nets. https://arxiv.org/abs/1405.3531.Google Scholar

9.Cubero, S, Lee, SW, Aleixos, N, Albert, F and Blasco, J (2016). Automated Systems Based on Machine Vision for Inspecting Citrus Fruits from the Field to Postharvest—a Review. Food and Bioprocess Technology 1-17.Google Scholar

10.Garcia, J and Barbedo, A (2016). A Review on the Main Challenges in Automatic Plant Disease Identification Based on Visible Range Images. Biosystems Engineering 144, 52–60.Google Scholar

11.Kaur R, Kang SS. An enhancement in the classifier supports vector machines to improve plant disease detection. In: Proceedings of the 3rd international conference on MOOCs, innovation and technology in education (MITE), Amritsar, 2015, pp. 135–40.

12.Islam M, Dinh A, Wahid K, Bhowmik P (2017) Detection of potato diseases using image segmentation and multiclass support vector machines. In: Proceedings of Canadian conference on electrical and computer engineering (CCECE), Windsor, 2017, pp. 1–4.

13.Hossain S, Mou RM, Hasan MM, Chakraborty S, Razzak MA. Recognition and detection of tea leaf's diseases using a support vector machine. In: Proceedings of 2018 IEEE 14th international colloquium on signal processing and its applications (CSPA), Batu Ferringhi, 2018, pp. 150–54.