Grenze International Journal of Engineering and Technology, January Issue

Recognition of Plant Diseases by Leaf Image Classification

Dr. A.R. Surve[1], Mrs. M. S. Dabade[2], Rutuja Patil[3] and Shivani Jadhav[4]
[1-4]Name of Institution/Department, City, Country Department of Computer Science and Engineering, Walchand College of Engineering, Sangli, Maharastra, India.
Email: anil.surve@walchandsangli.ac.in, manisha.dabade@walchandsangli.ac.in, rutuja.patil@walchandsangli.ac.in, shivani.jadhav@walchandsangli.ac.in

***Abstract*—Global food security depends on minimizing crop damage through timely detection of plant diseases. In addition to increasing crop yields, early disease detection reduces pesticide use, which is often applied without a precise understanding of the disease. Achieving global food security requires not only better crop varieties, but also effective disease detection. For millions of small and medium-sized farms around the world, disease detection has traditionally relied on manual inspection by farmers or experts, a time-consuming, costly, and impractical process. Artificial intelligence's deep learning branch offers a promising solution to this problem. Recently, deep learning has gained significant attention both in academic and industrial circles due to its advantages in automatic learning and feature extraction. A hot topic in agricultural plant protection, particularly in disease detection and pest assessment, it finds widespread applications in fields like image and video processing, voice recognition, and natural language processing. In plant disease detection, deep learning can overcome the limitations of manual feature selection, allowing for a more objective approach. As a result, disease detection is more accurate and research and technological advancements are accelerated. Using deep learning and advanced imaging techniques, this paper discusses current challenges and trends in detecting plant leaf diseases. We hope this work serves as a valuable resource for researchers focused on plant disease and pest detection, while also addressing some of the challenges and issues that still need to be addressed in this evolving field.**

***Index Terms*— Deep learning, Plant Leaf Disease Detection, visualization, small sample.**

I. INTRODUCTION

Agriculture is fundamental in shaping human lives and their economic status. It serves as the main origin of income of that mainly covers 58% of Indian population. On a global scale, India holds the second position in terms of agricultural output. As of 2018, it was noted that over half of India's workforce was engaged in agriculture, contributing around 18-20% to the nation's GDP. This makes India a prominent player in agricultural production and efficiency. Given that a significant portion of the population relies on this sector, tackling its challenges is crucial.

The agricultural sector confronts numerous problems, including outdated farming practices, improper fertilizer and manure usage, limited water resources, and plant diseases. Plant diseases, in particular, have a detrimental effect on crop health, influencing both growth and yield. The harm caused by these diseases can result in a 20-30% decrease in crop production, affecting both quality and quantity.

Grenze ID: 01.GIJET.11.1.317_1

Thus, early and precise disease detection is vital to mitigate substantial losses in agricultural output. Traditional methods of disease detection are not only prone to errors but labor-intensive also, leading to inappropriate treatments. However, with technological advancements, automated techniques for identifying and diagnosing plant diseases have become increasingly viable, offering more accurate solutions. The proposed system for detecting plant diseases is designed to focus on 14 different plant varieties, including apple, blueberry, cherry, corn, grape, orange, peach, pepper, potato, raspberry, soybean, squash, strawberry, and tomato. This system utilizes Deep Learning, particularly Convolutional Neural Networks (CNN), to analyze input images and classify the plants based on identified diseases, facilitating effective treatment strategies.

II. Related Work

Eftekhar Hossain et al. [1] proposed a system for detecting diseases in plant leaves utilizing the K-nearest neighbor (KNN) classifier. Their methodology involved extracting features from images of infected leaves to facilitate classification. The KNN classifier effectively identified prevalent plant diseases, including bacterial blight, early blight, bacterial spot, and leaf spot across various species, achieving an accuracy rate of 96.76%. Similarly, Sammy et al. [2] developed a system based on Convolutional Neural Networks (CNN) for classifying various plant diseases. Their research centered on nine distinct leaf diseases impacting tomato, grape, corn, apple, and sugarcane plants. The system was trained for approximately 50 epochs with a batch size of 22, utilizing the Adam optimizer and categorical cross-entropy. This model reached a remarkable accuracy of 96.5%. Ch Usha Kumari et al. [1] introduced a system employing K-means clustering combined with an Artificial Neural Network (ANN) for detecting plant diseases. The system computed features such as contrast, correlation, energy, mean, standard deviation, and variance. Their study underscored the significant effects of plant diseases on agricultural productivity, noting that delays in detection could exacerbate food insecurity. Traditionally, the identification of plant diseases relied on the expertise of agricultural professionals or seasoned farmers, who based their assessments on observation and experience. However, this method is subjective, time-consuming, labour-intensive, and often leads to inaccuracies. Novice farmers may misinterpret symptoms, resulting in inappropriate pesticide applications. Early detection is crucial for effective disease management and contributes to informed decision-making in agriculture. Despite advancements in detection systems, a notable limitation in this study was that only four diseases were examined, with the average accuracy being lower compared to other models.

Merecelin et al. [2] conducted a comprehensive study on disease detection in apple and tomato plants using CNN models, utilizing a dataset comprising 3,663 images of infected leaves, with their model achieving an accuracy of 87%. In a separate study, Jiayue et al. [3] established a disease recognition system for tomato fruits utilizing the YOLOv2 CNN model. This regression-based model incorporates a target detection algorithm recognized for its speed and accuracy, achieving a mean Average Precision (MAP) of 97%. However, one of the main challenges faced was the necessity for image tuning to enhance performance. Li et al. [5] focused on five types of apple leaf diseases, such as speckled deciduous disease, yellow leaf disease, round spot disease, mosaic disease, and rust disease. They extracted eight distinct features from images of infected apple leaves, including color, texture, and shape, employing a Backpropagation Neural Network (BPNN) for classification, which yielded an average recognition accuracy of 92.6%. Similarly, Guan et al. [6] investigated rice leaf diseases—specifically blast, stripe blight, and bacterial leaf blight—using step-based discriminant analysis alongside a Bayesian discriminant method. They extracted 63 features, including morphological, color, and texture attributes, achieving a recognition accuracy of 97.2%.

Deficiencies and limitations as follow:

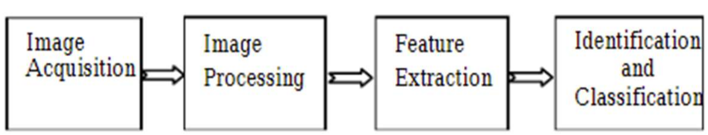

Figure 1

- The procedures are intricate, highly subjective, time-consuming, and labour-intensive.
- There is a heavy reliance on spot segmentation.
- The process depends significantly on manual extraction of features.
- The models performance or algorithm in recognizing diseases is difficult to assess in more complex environments.

Given these challenges, the development of an intelligent, rapid, and precise plant disease detection system is of great significance. Recently, deep learning technology has advanced considerably in the realm of plant disease identification. Deep learning (DL) is beneficial as it can autonomously extract features from images and classify disease spots on plants, removing the necessity for manual feature extraction and classifier design that traditional image recognition techniques require. This method, known as end-to-end processing, allows for the original characteristics of images to be effectively represented, making DL exceptionally powerful for recognizing plant diseases. Several factors have fuelled the swift adoption of deep learning in plant disease detection: the emergence of larger datasets, the prevalent use of multi-core graphics processing units (GPUs), and improvements in training deep neural networks facilitated by tools like NVIDIA's CUDA (Compute Unified Device Architecture). Among the various deep learning techniques, Convolutional Neural Networks (CNNs) are now regarded as the most efficient for image recognition, showcasing outstanding performance in image processing and classification tasks. CNNs were first implemented for plant image recognition focusing on leaf vein patterns to classify three leguminous plant species: white bean, red bean, and soybean. In another study, Mohanty et al. trained a deep learning model to identify 14 crop species and 26 related diseases.

Despite these encouraging outcomes, the limited diversity in datasets remains a challenge. Extensive datasets, often comprising thousands of images, are vital for effectively training CNNs. Unfortunately, in the domain of plant disease recognition, such datasets are still insufficient. A viable solution is transfer learning, which permits the adaptation of pre-trained CNNs for new tasks by retraining them with smaller, domain-specific datasets that differ from the larger datasets used initially. For detecting plant leaf diseases, CNN models that are pre-trained on datasets like ImageNet can be fine-tuned for this specific application. This integration of deep learning and transfer learning offers a promising approach to address the challenges posed by limited datasets in plant disease detection research. Robert G et al. [6] developed a system that utilized CNNs to identify diseases in tomato leaves. They reported 80% confidence score of the F-RCNN trained model, while the Transfer Learning model reached an impressive accuracy of 95.75%. The automated image capturing method demonstrated an accuracy of 91.67%. In a separate study, Halil et al. [7] deployed two deep learning network architectures—AlexNet and SqueezeNet—on the Nvidia Jetson TX1 for training and validation. The AlexNet model achieved an accuracy of 95.6%, while the SqueezeNet model attained 94.3%. Meanwhile, Sabrol et al. [8] applied a straightforward classification method for various plant diseases, which yielded promising results during the classification process.

III. Key Technologies

A. Deep Learning

Deep learning is sub branch of machine learning focuses on extracting high-level features from raw data through multiple layers. It simulates the way the human brain processes information by passing inputs through different levels of abstraction. This approach allows a computer to identify complex patterns and relationships in the data. Deep learning models often rely on neural network architectures, with the term "deep" referring to the presence of numerous hidden layers within the network. In contrast to traditional neural networks, which usually contain 2-3 hidden layers, deep neural networks may consist of as many as 150 layers or even more.

B. Convolutional Neural Network

Convolutional Neural Network (CNN) is a type of Deep Neural network is the, which is highly effective for processing 2D data like images. CNNs remove the need for manual feature extraction in image classification tasks by combining learned features with input data and utilizing 2D convolutional layers. Instead, CNNs automatically extract significant features directly from images during training. These features are not pre-defined but evolve as the network learns from different image datasets. A typical CNN is composed of several layers that handle image processing, including the input layer, output layer, convolutional layers, fully connected layers, pooling layers, and a softmax layer.

C. Disease Datasets

Common datasets for plant diseases include: 1) PlantVillage, an open-access resource, which contains 54,309 images of diseased leaves from 14 different fruit and vegetable crops, such as apple, blueberry, cherry, grape, orange, peach, potato, tomato, soybean, corn., strawberry, bell pepper, raspberry, and pumpkin.This dataset includes 26 different disease types—comprising 17 fungal diseases, 4 bacterial diseases, 2 mycotic diseases, 2 viral diseases, and 1 caused by mites. Additionally, it features 12 images of healthy leaves. 2) The Plant Pathology Challenge brings attention to issues like small sample sizes and imbalanced datasets, both of which negatively impact disease recognition performance in deep learning models. To improve these models for detecting leaf diseases, expanding the dataset is crucial. Data augmentation is often necessary in real-world applications but must be carefully applied—altering colors is not advisable since color differences are key to identifying diseases. Two common techniques are used to augment datasets.

Traditional Augmentation

Typical methods for data augmentation include physical expansion techniques (such as rotation, resizing, image translation, and distortion), web crawling, variational autoencoders (VAEs), and autoregressive models, among others. However, traditional expansion methods often produce samples with several drawbacks, including low quality, insufficient diversity, and inconsistency.

Generate Adversarial Networks (GANS)

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. in 2014, are a class of generative models. Since their introduction, various versions have emerged, including DCGAN, CGAN, PGGAN, LAPGAN, InfoGAN, WGAN, F-GAN, SeqGAN, and LeakGAN, among others. The main goal of these models is to generate synthetic data that closely mimics the characteristics of the training distribution. GANs consist of two key components: the generator and the discriminator. The architecture of these components is depicted in Fig. 3.

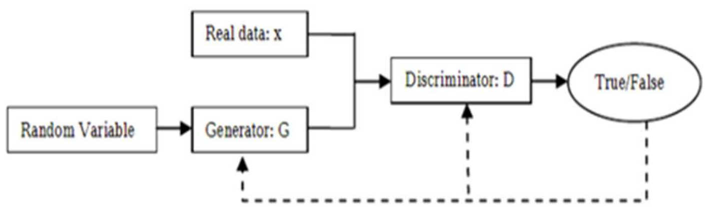

Figure 2

IV. PROPOSED MODEL

Each year, the agricultural sector faces significant declines in productivity and substantial crop losses due to the delayed identification of plant diseases. Farmers often rely on visual inspections of leaves, which are frequently insufficient for accurately diagnosing specific diseases; thus, expert consultations become necessary for reliable assessments. These delays in disease recognition adversely affect crop yields and, consequently, the livelihoods of farmers. To address this challenge, the proposed model introduces an automated system designed to detect plant diseases, enabling timely interventions to safeguard crop health.

- Collection of Plant Image Dataset: The first step is gathering a dataset consisting images of 14 different plant varieties, which includes 38 distinct classes of plant diseases.
- Image Preprocessing: The images undergo preprocessing through various convolutional layers to upgrade their quality and prepare them for further analysis.
- Leaf Classification: Finally, the leaves are classified based on their condition, determining whether the plant is healthy or diseased.

V. DESIGN A SYSTEM ARCHITECTURE

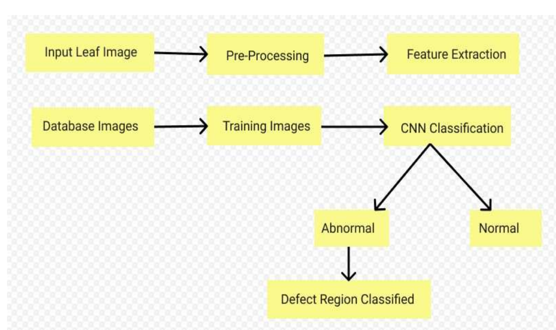

Figure 3 Type Sizes for Camera-Ready Papers

The architecture of the proposed system includes several key components: obtaining data from a comprehensive dataset, processing it through multiple convolutional layers, and classifying plant diseases. During the classification stage, the system determines whether the plant image is healthy or diseased.

VI. Experimental Analysis and Results

A. Data Analysis

The dataset used in this project is the New Plant Diseases dataset, sourced from the Kaggle website. It contains a collection of images representing both diseased and healthy plant leaves.

B. Result Analysis

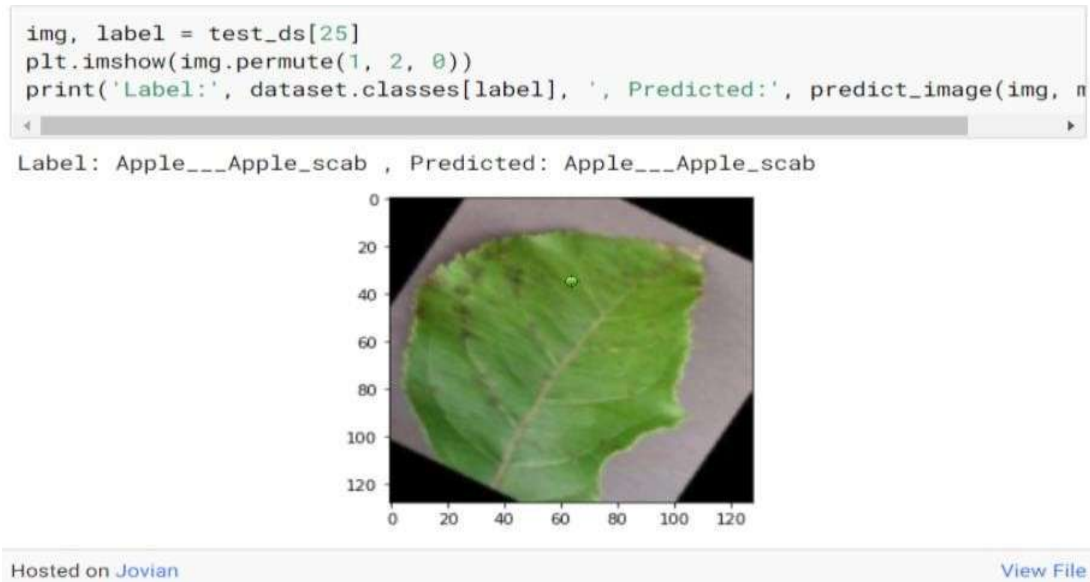

Figure 4

Figure 5

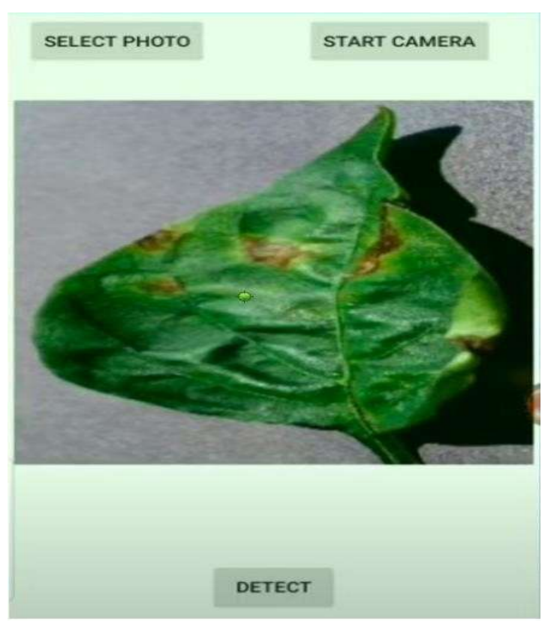

Figure 6

VII. IMPLEMENTATION DETAILS

The development platform for this project is Python, which provides essential libraries required to execute the code. The key libraries utilized include PyTorch, Pandas, NumPy, and Keras.

VIII. CONCLUSIONS

A large segment of India's population depends on agriculture, making it important to detect and diagnose leaf diseases which leads to crop losses, as agriculture is a key driver of the country's economic growth. This project leverages deep learning, utilizing Convolutional Neural Networks (CNNs), to create a system capable of identifying and recognizing 13 different plant leaf diseases. The CNN architecture is streamlined with a minimal number of layers, ensuring effective classification across seven distinct disease categories. For model training, the New Plant Diseases dataset is employed, offering a wide array of images that enhance the system's accuracy and reliability. Additionally, a user-friendly Graphical User Interface (GUI) is developed, allowing users to effortlessly navigate and upload images from the dataset. Users can select an image, which is then processed for analysis, with the system displaying the predicted disease directly on the interface, offering instant feedback. The CNN, specifically trained for detecting and recognizing plant leaf diseases, achieves a remarkable accuracy rate of 94.8%, successfully classifying nearly all images. Although a few outliers exist, the system overall proves highly effective in helping farmers and agricultural stakeholders manage crop health more efficiently.

REFERENCES

[1] Ch. Usha Kumari, S. Jeevan Prasad, G. Mounika, “Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN “Proceedings of the 3rd International Conference on Computing Methodologies and Communication (ICCMC)”, Erode, India, 2019, pp 10951098

[2] Mercelin Francis, C. Deisy, “Disease Detection and Classification in Agricultural Plants Using Convolutional Neural Networks — A Visual Understanding”, Proceeding of the 6 th International Conference on Signal processing and Integrated network, Noida, India, 2019, pp 1063-1068.

[3] Jiayue Zhao, Jianhua Qu, “A Detection Method for Tomato Fruit Common Physiological Diseases Based on YOLOv2 “Proceeding of the 10th International Conference on Information Technology in Medicine and Education”, Qingdao, China, China, 2019

[4] Robert G. de Luna, Elmer P. Dadios, Argel A. Bandala, “Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition” Proceeding of the IEEE Region 10 Conference, Jeju, South Korea, 2018, pp 1414-1419.

[5] Halil Durmus, Ece Olcay Gunes, “Disease detection on the leaves of the tomatoplants by using deep learning” Proceeding of the 6 th International Conference oftheAgro-Geoinformatics, 2016, Fairfax, VA, USA

[6] H. Sabrol , K. Satish, “Tomato plant disease classification in digital images using classification tree” Proceeding of the International conference on communication and signal processing, Melmaruvathur, India, 2016 pp 12421246

[7] Usama Mokhtar, Mona A.S. Ali, Hesham Henfy, “Tomato leaves diseases detection approach based on support vector machines” Proceeding of the 11th International Computer Engineering Conference, Cairo, Egypt, 2015 pp 246-250

[8] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). "Using deep learning for image-based plant disease detection." Frontiers in Plant Science, 7, 1419.

[9] Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). "Deep learning for tomato diseases: Classification and symptoms visualization." Applied Artificial Intelligence, 31(4), 299-315.

[10] Ferentinos, K. P. (2018). "Deep learning models for plant disease detection and diagnosis." Computers and Electronics in Agriculture, 145, 311-318.

[11] Patel, N., Patel, P., & Prajapati, Y. (2020). "Plant leaf disease detection using machine learning." In International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 272-275.

[12] Amara, J., Bouaziz, B., & Algergawy, A. (2017). "A deep learning-based approach for banana leaf diseases classification." In Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband, 79-88.

[13] Rangarajan, A. K., Purushothaman, R., & Ramesh, A. (2018). "Tomato crop disease classification using pre-trained deep learning algorithm." Procedia Computer Science, 133, 1040-1047.

[14] Barbedo, J. G. A. (2018). "Factors influencing the use of deep learning for plant disease recognition." Biosystems Engineering, 172, 84-91.

[15] Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). "A comparative study of fine-tuning deep learning models for plant disease identification." Computers and Electronics in Agriculture, 161, 272-279.

[16] Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). "Deep neural networks based recognition of plant diseases by leaf image classification." Computational Intelligence and Neuroscience, 2016, Article ID 3289801.

[17] Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition." Sensors, 17(9), 2022

[18] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). "Deep learning in agriculture: A survey." Computers and Electronics in Agriculture, 147, 70-90.
[19] Singh, V., Srivastava, P., & Singh, B. K. (2022). "Tomato leaf disease detection using deep learning and machine learning techniques." Computers, Materials & Continua, 72(2), 3877-3894.
[20] Chen, J., Zhang, D., Li, H., & Zhang, Y. (2021). "Tomato disease recognition based on deep learning model." Journal of Ambient Intelligence and Humanized Computing, 12, 931-938.
[21] Saleem, M. H., Khanchi, S., Potgieter, J., & Arif, K. M. (2021). "Image-based plant disease identification by deep learning meta-architectures." Plants, 10(1), 28.
[22] Prajapati, H. B., Shah, M., & Dabhi, V. K. (2021). "Detection and classification of rice plant diseases using image processing and SVM classifier." Procedia Computer Science, 167, 516-528.
[23] Prajapati, H., & Patel, V. (2021). "Detection and classification of leaf disease using novel deep learning model." IEEE Access, 9, 77963-77973.