IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

" Smart Agriculture: ML And CNN For Early Detection Of Leaf Diseases"

Prof. Dhiraj S. Kalyankar[1], Mr. Ayush D. Kadu2, Mr. Chaitanya G. Bhagat3, Ms. Anjali P. Tambade [4], Mr. Kaustubh S. Dhoran [5]

Department of COMPUTER SCIENCE AND ENGINEERING, DRGITR, Amravati, Maharashtra, India

Abstract:

This paper provides an overview of the use of deep learning in plant protection, specifically in the identification of crop leaf diseases. Deep learning has gained significant attention due to its advantages in feature extraction and machine learning, making it a popular tool in various fields such as image and video processing, speech processing, and natural language processing. By applying deep learning to plant disease identification, it can provide more objective and efficient extraction of disease characteristics and improve research efficiency and technology transformation. our research aims to summarize recent advances in deep learning research related to crop leaf disease identification, highlighting current trends and challenges in this area, and serving as a useful resource for researchers studying plant pest detection. We are using CNN algorithm which provided 97% accuracy to detect disease.

Keywords: CNN, feature Extraction, plant leaf disease detection. Accuracy.

1. INTRODEUCTION

Plants play a crucial role in providing energy and are a key component in addressing the global warming crisis. However, several plant diseases have the potential to cause disastrous economic, social, and ecological consequences. To meet the growing demand for food, global crop production must increase by at least 50% by 2050. Currently, the majority of crop production occurs in Africa and Asia, where most growers are family-run businesses with limited horticultural experience. Consequently, pests and diseases often cause yield losses of over 50%. The traditional method of human analysis through visual inspection is no longer sufficient for identifying agricultural diseases. The primary objective of this project is to detect the most common diseases in plant leaves, specifically those of tomato, potato, and pepper. This system can identify 15 different forms of diseases in these three plants. The user can upload an image of a leaf, and if the leaf is infected, the system will display the name of the disease and suggest the appropriate pesticide after clicking the predict button. If there is no disease present on the plant leaf, the system will show a message that reads "There is no disease on the plant." Additionally, the system displays the percentage of affected areas and recommends pesticides based on that proportion.

1.1 Need of Plant Detection

Plant diseases can have a devastating impact on both farmers and consumers. When plants become infected, farmers may experience a significant loss in income, while consumers may face shortages and higher prices. The economic cost of plant leaf diseases alone is estimated at a staggering 60 billion dollars. One of the most critical requirements for this project is that the detection of disease must be easy for farmers to perform in a simple manner.

Farmers cannot rely on visual inspection alone to predict the onset of disease in plants. Therefore, the proposed system must enable farmers to identify the disease and recommend effective pesticides to prevent further spread of the disease. If a single plant leaf becomes infected, it can quickly spread to other leaves and potentially destroy the entire yield.

1.2 Problem Statement

To optimize the use of Machine Learning algorithms and streamline the process of detecting plan diseases, as well as assessing their impact on crop yield and recommending suitable pesticides, with the aim of reducing both time and cost for farmers.

1.3 Scope

This project has the potential to significantly benefit rural farmers by assisting them in protecting their crops from diseases. Farmers often suffer significant losses due to crop diseases, and this system aims to prevent such situations from occurring. Moreover, the project has been implemented in a regional language to enhance its accessibility and comprehension for farmers.

Additionally, the system can recommend effective pesticide treatments and appropriate dosages based on the current disease situation. It can also provide farmers with insights into the impact of crop diseases on crop yield, allowing them to make informed decisions. Overall, this project has the potential to be a valuable tool for farmers in combating crop diseases and ensuring the health and productivity of their crops.

1.4 Convolutional Neural Network

Convolutional neural networks (CNNs) may seem like an unconventional blend of biology and math with a touch of computer science, but they have revolutionized computer vision and become one of the most influential advancements in this field. In 2012, Alex Krizhevsky employed neural nets to win the annual ImageNet competition, reducing the classification error record from 26% to 15%, a remarkable feat at the time.

Since then, numerous companies have integrated deep learning at the heart of their services. Facebook uses neural nets for automatic tagging algorithms, Google for photo search, Amazon for product recommendations, Pinterest for home feed personalization, and Instagram for search infrastructure. Despite their versatility, CNNs are widely recognized for their effectiveness in image processing. In the realm of image processing, let's examine how CNNs can be utilized for image classification

2. LITERATURE SURVEY

The use of smart farming systems that incorporate advanced technologies such as deep learning and computer vision can greatly improve the efficiency and productivity of agriculture, including the cultivation of tomatoes. Tomato farming involves various factors such as soil quality, environmental conditions, and sunlight exposure, which can make it challenging to avoid diseases. However, the development of an innovative solution that utilizes an automated image capturing system can help detect and recognize leaf diseases in tomato plants. The system uses a motor-controlled image capturing box to capture images of the four sides of every tomato plant, which are then analysed using a deep convolution neural network (CNN) to identify the presence of three diseases, namely Phroma Rot, Leaf Miner, and Target Spot. To train the CNN, a dataset containing images of both diseased and healthy tomato leaves was collected. The system achieves a high accuracy of 95.75% using Transfer Learning disease recognition model.

The F-RCNN trained anomaly detection model produced a confidence score of 80%, which indicates the level of certainty in identifying the presence of a disease. The automated image capturing system was implemented in real world settings and was able to achieve an accuracy rate of 91.67% in identifying tomato plant leaf diseases. Overall, the use of smart farming systems and advanced technologies such as deep learning and computer vision can greatly improve the efficiency and productivity of tomato farming by enabling the early detection and prevention of diseases, which can help increase the quality and quantity of tomato production. In This paper The field of agriculture has great impact on our lives.

Agriculture is the most important sector of our economy. It is difficult for farmers to identify leaf diseases, so their yields are lower. However, videos and leaf images give agronomists a better view and can provide better solutions. Therefore, problems related to crop diseases can be solved. It is important to note that if the productivity of the crop is poor, there is a high chance that it will not be able to provide Plant Leaf Diseases Detection Using ML And CNN Dept of CSE, DRGIT&R, Amravati Page 4 good nutrition.

Thanks to improvements and developments in technology, the devices are smart enough to identify and detect plant diseases. Recognize diseases and treat them more quickly to reduce the negative impact on the harvest. This article mainly investigates the use of image processing technology for the detection of plant diseases. This paper obtains an open dataset image composed of 5000 images of healthy and diseased plant leaves, and uses semi-supervised techniques to detect crop types and four disease types. Scientists used the convolutional neural network (CNN) to divide plant leaf diseases into 15 categories, including three classes for healthy leaves and 12 classes for diseases found in various plants, such as bacteria, fungi, and others. With a training accuracy of 98.29% and a testing accuracy of 98.029% for all used data sets, they were able to achieve excellent accuracy in both.

In addition to feature extraction and dimensionality reduction techniques, researchers have also explored the integration of advanced machine learning algorithms and deep learning architectures for more accurate and robust disease detection and classification. Regarding image analysis, the current survey focuses on a particular subset of DL models and techniques since there are very few of this type of survey in The agricultural field, especially about CNN utilization.

Machine learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool in this domain due to its ability to automatically learn discriminative features from raw input data. For instance, S. Mohanty et al. proposed a deep learning-based approach for the identification and classification of plant diseases using a CNN architecture trained on a large dataset of plant images. Their model achieved impressive results, outperforming traditional methods and demonstrating the potential of deep learning in plant disease diagnosis.

3. ARCHITECTURE AND WORKING

3.1 Existing System

Now a day, we regularly use digital camera and other electronic devices day today life. Then the automatic plant disease identification has been widely applied as a satisfactory alternative.

Most of the cases following traditional machine learning approaches such as support vector machine (SVM), CNN and K-means clustering have complex image pre-processing and feature extraction steps, which reduce the efficiency of disease diagnosis.

3.2 Proposed System

Agriculture is one of the most important occupations in the world. It plays an important role because food is a basic need of all living beings on this planet. In this proposed system, deep learning region-based convolutional neural network (R CNN) method is used for recognition. They have two phases, the training phase and the testing phase. At the initial stage, they perform image acquisition, preprocess the images, and train the images using R-CNN.

The second step is the classification and identification of leaf diseases. For training purposes, images are extracted from the dataset, while real time images are available for testing. Foliar disease diagnosis is performed from images uploaded to the system or database. Real-time environment input requires image pre-processing and then performs feature classification to detect disease diagnosis and get disease name.

.2.1 Architecture Diagram

In the architecture shown below, once the user uploads or captures an image of a plant leaf, the system processes the image through a pre-trained deep learning model stored in the database. Based on the classification, the system provides the user with the disease name and a diagnosis report. If the disease is recognized, relevant treatment suggestions and preventive measures are also displayed to assist the user in managing the plant's health.

3.3 Proposed model

The model is developed based on the IP and ML approaches for detection of leaf disease in presented in this section. The proposed model (DWT + PCA + GLCM +CNN) using computer vision and machine learning approaches for leaf disease detection is shown in Fig 3.1. The tomato samples having six disorders are considered to evaluate its accuracy and to recognize the leaf disease as Healthy or Unhealthy.

As a part of image processing, the samples of tomato are resized to 256 × 256 pixels to maintain equal in their size throughout the experiment. The HE and K-means clustering are employed to maximize the quality and segment the leaf samples. Based on the K-means clustering response, the leaf is diseased or not can be predicted at the early stage of operation. The boundaries of the leaf samples can be extracted using contour tracing.

The DWT, PCA and GLCM are used to extract the informative regions/features of the samples. In the next stage as a part of machine learning approaches the SVM, KNN and CNN are used to classify the features and the performance of the model is recorded.

When we provide a new input image first the module extracts the leaf features. Then it goes through the CNN model. It then compares of the features with a already trained dataset. Then it can goes through dense CNN and therefore the leaf features are extracted separately. Then the module will predict whether the plant leaf is affected by any disease or not. It shows the output from one among the38 classes which are predetermined and trained. Then the output is going to be during a textual format.

In the next stage as a part of machine learning approaches the SVM, KNN and CNN are used to classify the features and the performance of the model is recorded.

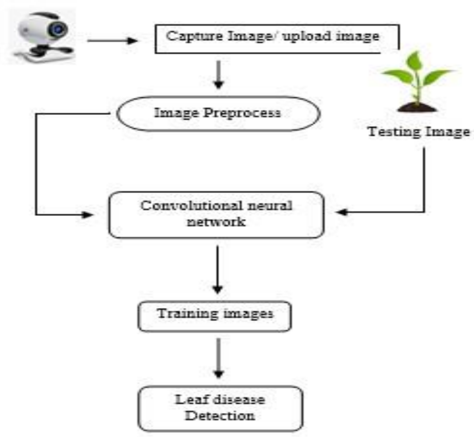

Fig. 1: Architecture of Plant Leaf Disease Detection

When we provide a new input image first the module extracts the leaf features. Then it goes through the CNN model. It then compares of the features with a already trained dataset It shows the output from one among the38 classes which are predetermined and trained. Then the output is going to be during a textual format.

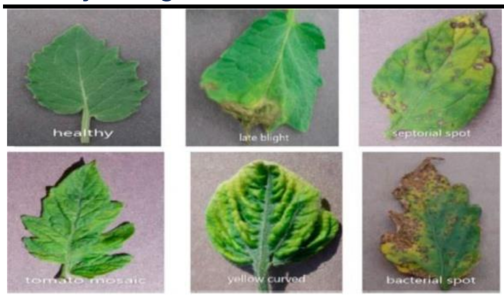

Fig 3.2: Sample of tomato having disorders.

Fig 3.3: K-mean clustering.

.3.1 Dataset

The village database of tomato leaf is considered, the plants which are affected from variety of diseases. The images of tomato leaf having six disorders are taken to carry out the experiments for detection of leaf disease.

3.3.2 Preprocessing

K-means clustering technique is applied on leaf images to find out the infected region. The K-mean clustering is used to get the data center of the image and make the clusters of that image and calculates the center distance from the other cluster. Samples of leaf after applying k-mean clustering algorithm is shown in Fig. Contour tracing is performed on digital leaf samples to ex- tract their general shape information. After extracting the contour, its characteristics is analysed and used for pattern classification.

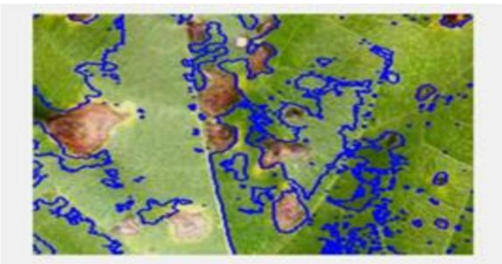

Fig 3.4: Contour tracing.

It of- ten helps for determining the efficiency of feature extraction process. The images appeared after performing contour tracing is shown in Fig. 3.4 there are also many deep learning methods available to detect different types of leaf disease detection and classify into various categories of leaf disease, such as Bacterial spot, Early blight, Late blight, Mold, Septoria leaf spot, Spider mites, Two spotted spider mite, Target Spot, Mosaic virus, Yellow Curl Virus, and Healthy.

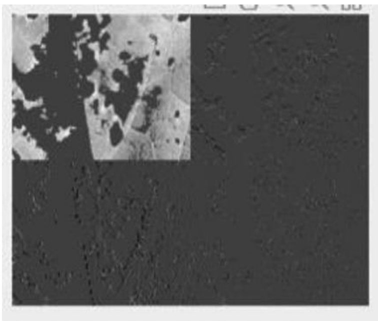

Fig 3.5: DWT decomposition.

3.3.3 Feature Extraction

Discrete Wavelet Transform: The DWT is applied on enhanced tomato samples to extract the useful features. The DWT decomposes into sub-bands of lower (LL, LH, HL) and higher frequency (HH) components. The component of DWT carries maximum availability of information when compared with higher frequency components of DWT as shown in Fig 3.5.

Gray Level Co-occurrence Matrix: The optimal features are selected obtained from wavelet decomposition is carried out by Principal Com- ponent Analysis. The GLCM uses in the distribution of higher order of gray values that are defined with neighbourhood criterion. The several properties are derived from the GLCM technique for extraction of leaf features.

The most used texture-based features are as follows. Homogeneity, Autocorrelation, Dissimilarity, Entropy, Sum of squares, Average, Variance, Entropy. The features obtained using DWT, GLCM and PCA are combined to form feature vector which are provided as an in- put sample to the classifiers to recognize classify the images.

4. APPLICATION

1. Data augmentation techniques:

Data augmentation techniques increase the diversity of the data during training by artificially generating additional samples from the real dataset. Furthermore, image augmentation is a technique that creates new data from existing data to help train a deep neural network model. The accuracy improved as well. In another study, used resizing and image segmentation methods to increase the size of the dataset from 1567 images to 46,409 images. The accuracy improved by 10.83% over the no expanded dataset.

2. Transfer learning:

Transfer learning is a machine learning technique in which we reuse a previously trained model as the base for a new model on a new task. As a result of the new datasets, it will just retrain a few layers of pertained networks which helps to reduce the amount of data required.

Introduced the INC-VGGN DL architectural features to Plant Leaf Diseases Detection Using ML And CNN Dept of CSE, DRGIT&R, Amravati Page 18 detect plant diseases, which used transfer learning by changing the pre-trained VGG Net. This method was developed using the VGG16 CNN model, which was pre-trained on the public dataset. The study provided a satisfactory result, with a 94.5% recall rate and a 95.00% accuracy rate.

3. Citizen science:

In 1995, the concept of citizen science was proposed. Nonprofessional participants collect data as part of a scientific study in this technique. Farmers submit the collected images to a server for plant disease and pest classification, after which the images are correctly labelled and analysed by an expert.

4. Data sharing:

Data sharing is another way of increasing datasets. Several studies are now being conducted worldwide on accurate disease detection. The dataset will become more accurate if the different datasets are shared. This situation will encourage more significant and satisfying study results.

Plant Village and Kaggle are the most commonly employed public datasets in the literature on DL methods for plant disease classification and detection. describes these datasets as the most widely applied public datasets. Plant leaf disease detection using machine learning (ML) and Convolutional Neural Networks (CNNs) is a rapidly evolving field with practical applications in agriculture.

The goal is to identify and classify plant diseases at an early stage to prevent significant crop losses. Below is an overview of how these technologies are applied in this domain.

5. ADVANTAGES & DISADVANTAGES

Advantages:

a. High Accuracy

CNNs are particularly good at image recognition tasks due to their ability to automatically learn relevant features from images. This leads to high accuracy in identifying plant diseases compared to traditional methods.

CNNs can detect subtle variations in the leaf's colour, texture, and shape, which helps in accurately classifying various diseases.

b. Automation

Once trained, CNN-based systems can automatically classify plant diseases without human intervention. This automation helps reduce labour and human errors that can occur in manual inspection of plants.

Farmers can use mobile applications with built-in disease detection models to monitor crops in real time.

c. Scalability

CNNs can be trained on large datasets and deployed at scale for different types of plants and diseases, making the system highly adaptable for large-scale farming.

d. Non-invasive and Real-time Detection

Using image-based disease detection does not harm plants, unlike some traditional methods where samples need to be taken from leaves for lab testing. It enables early detection of diseases, helping farmers to act before the disease spreads, thus saving crops and reducing losses.

e. Improved Crop Management

Integration of ML with geographic information systems (GIS) and weather data can provide holistic insights into disease patterns, improving crop health monitoring and disease prediction.

Disadvantages:

a. Need for Large and Diverse Datasets

CNNs require a large number of labeled images of both diseased and healthy leaves for training. Collecting and annotating such datasets can be time-consuming and expensive. If the dataset is biased or lacks diversity, the model may not generalize well, leading to poor performance on new, unseen data.

b. Complexity in Model Training

Training CNNs requires substantial computational resources, including high-performance GPUs, and can be time-intensive, especially with large datasets.

c. Overfitting

Techniques like data augmentation, regularization, or using simpler models are required to prevent this issue.

d. Limited by Image Quality

The accuracy of detection heavily depends on the quality of the input images. Poor lighting, resolution, or camera angles can affect model performance. Images captured in the field under varying conditions (e.g., light, shadows) may lead to inaccuracies compared to controlled lab environments.

e. Disease Similarities and Complexities

Some plant diseases have very similar visual symptoms, making it challenging for even CNNs to distinguish between them. This can result in misclassification of diseases. In some cases, diseases may not manifest visually in the early stages, so image-based detection may miss these early signs.

6. FUTURE SCOPE

The future scope in plant leaf disease detection offers opportunities for further innovation and practical application in agricultural settings. Moving forward, advancements in image processing techniques, coupled with deep learning algorithms, can lead to more accurate and efficient disease detection systems.Researchers can explore the development of novel feature extraction methods and model architectures tailored specifically for detecting subtle signs of disease in plant leaves Additionally, the application of artificial intelligence (AI) techniques beyond disease detection opens up new avenues for enhancing agricultural sustainability and productivity.

Researchers can explore the use of AI-driven decision support systems to optimize crop management practices, resource allocation, and pest control strategies. Overall, the future of plant disease detection lies in the convergence of cutting-edge technologies and interdisciplinary research efforts aimed at addressing the complex challenges facing modern agriculture. The analysis of the proposed model is well suited for CNN machine learning classification technique with a desired accuracy compared to other state of the art method.

In future, the model can be improved using fusion techniques for extraction of significant features and examined for other leaf samples of data-sets. The proposed model uses computer vision techniques including RGB conversion to gray, HE, K-means clustering, contour tracing is employed in preprocessing stage. The analysis of the proposed model is well suited for CNN machine learning classification technique with a desired accuracy compared to other state of the art method. In future, the model can be improved using fusion techniques for extraction of significant features and examined for other leaf samples of datasets.

7. CONCLUSION

To prevent losses, small holder farmers are dependent on a timely and accurate crop disease diagnosis. In this study, a pre-trained Machine Learning and Convolutional Neural Network was fine-tuned and the model was deployed online. The final result was a plant disease detection app. This service is free, easy to use and requires just a smart phone and internet connection. Thus, the user's needs as defined in this paper have been fulfilled. A thorough investigation exposes the capabilities and limitations of the model. Overall, when validated in a controlled environment, an accuracy of 97.2% is presented. This achieved accuracy depends on a number of factors including the stage of disease, disease type, background data and object composition. Due to this, a set of user guidelines would be required for commercial use, to ensure the stated accuracy is delivered. As the model was trained using a plain background and singular leaf, imitation of these features is best. Augmentation and transfer learning in this case, proved beneficial to the model, helping the CNN to generalize more reliability. While this improved the model's ability to extract features r, it was not enough when the model was presented with 'in field' imagery. In this case, the classifier ranked an accuracy of just 44% Above all, this highlights the importance of diversifying the training dataset to include alternative background data, additional plant anatomy and varying stages of disease. In the future, work should be focused on diversifying training datasets and also in testing similar web applications in real life situations. Without such developments, the struggle against plant disease will continue.

8. REFERENCES

1. Md. Tariqul Islam, "Plant disease detection using CNN model and image processing," International Journal of Engineering Research and Technology, 09 (10), 291–297, 2020.

2. S.B. Gharpankar, "Plant Leaf Disease Detection using Convolutional Neural Networks," International Journal of Research in Applied Science and Engineering Technology, 8 (11), 494–499, 2020.

3. S. Kumar, K. Prasad, A. Srilekha, T. Suman, B.P. Rao, J.N. Vamshi Krishna, "Leaf Disease Detection and Classification based on Machine Learning," IEEE International Conference on Smart Technologies in Computing, Electrical and Electronics, 361–365, 2020.

4. S. Nema, A. Dixit, "Wheat Leaf Detection and Prevention Using Support Vector Machine," International Conference on Circuits and Systems in Digital Enterprise Technology, 1–5, 2018.

5. Vidyashree Kanabur, Sunil S. Harakannanavar, Veena I.

Puranikmath, Dattaprasad Torse, "Detection of Leaf Disease Using Hybrid Feature Extraction Techniques and CNN Classifier," Springer Computing, 1213–1220, 2019.

6. K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, "Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system," IEEE Transactions on Intelligent Transportation Systems, 22 (7), 4337–4347, 2020.

7. S.H. Pushpa, A. Ashok, "Diseased leaf segmentation from complex background using indices based histogram," IEEE International Conference on Communication and Electronics Systems, 1502–1507, 2021.

8. S. Hossain et al., "Recognition and detection of tea leaf's diseases using support vector machine," IEEE International Colloquium on Signal Processing and Applications, 150–154, 2018.

9. M. Arya, K. Anjali, D. Unni, "Detection of unhealthy plant leaves Using image processing and genetic algorithm with Arduino," IEEE International Conference on Power, Signals, Control and Computation, 1–5, 2018.

10. K. Yu, L. Tan, L. Lin, X. Cheng, Z. Yi, T. Sato, "Deep-learning-empowered breast cancer auxiliary diagnosis for 5GB remote E-health," IEEE Wireless Communications, 54–61, 2021.

11. B. Mishra, S. Nema, M. Lambert, S. Nema, "Recent technologies of leaf disease detection using image processing approach-review," IEEE International Conference on Innovations in Information, Embedded and Communication Systems, 1–5, 2017.

12. M.H. Saleem, J. Potgieter, K.M. Arif, "Image-based plant disease identification by deep learning meta-architectures," Plants (Basel), 9 (11), 1–23, 2020.

13. J. Zhang et al., "3D reconstruction for motion blurred images using deep learning based intelligent systems," CMC-Comput. Mater. Continua, 66 (2), 2087–2104, 2021.

14. R. Prakash, G.P. Saraswathy, G. Ramalakshmi, "Detection of leaf diseases and classification using digital image processing," IEEE International Conference on Innovations in Information, Embedded and Communication Systems, 1–4, 2017.

15. W. Wang, X. Li, L. Wang, "Identification of plant leaf diseases based on improved convolutional neural network," Sensors, 20 (16), 4236, 2020.

[1] V. Gupta et al., "Automated segmentation of powdery mildew disease from cherry leaves using image processing," IEEE International Conference and Workshop on Bioinspired Intelligence, 1–4, 2017.