A CNN Based Approach For Detection Of Grape Leaf Diseases

*Smita Rani Sahu[1], Bhogi Saikrishna[2], Kothakota Priyanka[3], Koppala Pavan Sai[4], Gujjuru Rohith[5]

[1,2,3,4,5]Dept of Information Technology, Aditya Institute of Technology and Management,Tekkali,Andhra Pradesh, India

[1*]smitasahu57@gmail.com
[2]bhogisaikrishna@gmail.com
[3]priyankakothakota786@gmail.com
[4]pavansai6868@gmail.com
[5]rohithrohi14216@gmail.com

Abstract: Plant leaf disease detection has become increasingly important in ensuring sustainable agriculture and maintaining crop health. Since plant illnesses are quite widespread, finding infections in plants is an important job in the agricultural industry. Manual inspection, which is labour-intensive and subjective, is the basis for traditional plant disease detection. It can be inaccurate and has a limited scope. Faster and more accurate detection is provided by more recent techniques like deep learning and machine learning. They can handle a broader range of diseases, making them an appealing option for large-scale, efficient plant disease management. Every nation must automate its agricultural sector. Plant diseases are typically characterised by visual symptoms, and in recent years, a number of deep learning models have produced exceptional results in the classification of plant diseases. Diseases that affect grape plants, such as leaf blight, black measles, and black rot, lower crop yields Early intervention is essential to address this crop disease. A proper diagnosis is required. This paper uses the Grape Leaf image dataset, which comprises 8845 images with four different classes, and applies a deep learning-based convolutional neural network to perform disease prediction. Additionally, various optimisation strategies and activation functions were employed to bring out the differences in convolutional neural network (CNN) model performance. CNN-Nadam with a sigmoid activation function outperforms other CNN optimizers with 99.45% accuracy, according to an analysis of the experiment results. Therefore, quick action would help minimise losses in plant productivity. Revenue, economic expansion, and agricultural productivity will all increase as a result.

Keywords: Agriculture, Grape, Diseased plant, Convolutional Neural Network, Deep Learning, Image Classification, Nadam optimizer.

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_8

1 Introduction

A significant drop in the quantity and quality of agricultural products has occurred recently as a result of diseases affecting different plant species, which have presented challenges for the agricultural sector. In those effected plants, leaves play a crucial role as they are important for photosynthesis, the primary function supporting a plant's overall health. Grape diseases called Black Rot, Leaf Blight, Black Measles are common grape leaf diseases affects the grape crops seriously. Plant leaf disease detection has historically relied on skilled professionals conducting visual inspections, which frequently results in higher prediction errors [1]. To reduce the human efforts researchers applies the machine learning algorithms such as k-means clustering to detect the plant diseases [2]. When it comes to feature extraction, machine learning is not very good, image preprocessing is complex, and plant detection accuracy is quite low. Due to CNN's ability to extract features from photos, plant diseases have been detected on leaves in recent years [3]. Earlier predefined model is used to detect grape leaf disease due to incompatibility of the predefined model gets less accuracy and predictions not accurate [4]. This paper develops CNN for the classification of diseases affecting grape leaves with the goal of producing an accurate, mean, and affordable method for detecting the diseases that affect plant leaves. CNN, a deep learning technique, is used to identify diseases from images of leaves. The superiority of Convolutional Neural Networks (CNNs) over predefined models for plant leaf disease detection is highlighted in this research paper. Examining CNNs' benefits—such as their adaptability, flexibility, and ability to learn complex features—the study offers a thorough comparison. The study supports CNNs as a cutting-edge approach by analysing performance metrics, training requirements, computational complexity, and interpretability.

2 Literature Survey

There are primarily two types of studies in the literature on plant disease detection. Using traditional machine learning techniques to analyse image data, such as K-Nearest Neighbours (KNN), Support Vector Machine (SVM), and Decision Tree (DT), is the first group. Using a deep learning technique with automatic feature extraction, the second group will directly use the plant image data set for decision classification.

Garima Shrestha et al. developed procedure for plant disease detection using CNN by using plants leaves image dataset got overall accuracy of 88.88% and it only deals only for 15 cases of diseased plant leaves and it takes more time complexity [3].Rahul S.et al. created method for grape leaf disease detection using a deep learning-based VGG16 model and grape leaf image dataset obtained an overall accuracy of 78.28%; however, it could only distinguish between 10 different grape leaf diseases with lower accuracy. [4]. Deep learning techniques have recently been applied to the categorization of plant diseases. Using a dataset of maize leaves, Malusai et al. developed a procedure for the reorganisation and classification of maize leaf diseases using convolutional neural networks. The procedure achieved an overall accuracy of 92.85%, but it was only applied to three diseases of maize leaves, resulting in lower accuracy. [5]. Using a paddy leaf disease dataset, Md. Ashiqul et al. developed a procedure for an automated convolutional neural network-based approach for paddy leaf disease detection. Using Inception-ResNet-V2, they obtained an overall accuracy of 92.68%. The chosen model provided more negative predictions than was appropriate for the selected dataset [6]. Anwar et al. developed procedure for plant disease detection using AI based vgg-16 model by using plant leaf dataset got overall accuracy of 91% and the model presents challenges for plant disease diagnosis because it depends on particular lighting conditions and intricate backgrounds in input photos [7]. Using a dataset of plant leaf disease images, R. Sujatha et al. developed a procedure for comparing the performance of deep learning compared to machine learning in plant leaf disease detection. This procedure achieved an overall accuracy of 89.5% and was limited to the detection of citrus plant diseases, where it performed less accurately [8]. Using a small dataset of images, Mehmet et al. developed a procedure for the automatic detection and classification of leaf spot disease in sugar beetroot using deep learning algorithms. The procedure achieved an overall accuracy of 93.48% and is limited to the detection of sugar beetroot leaf spot diseases [9]. In order to detect soybean leaf disease from a synthetic image more quickly, Keke et al. developed a procedure called multifeatured fusion. Even though detecting soybean leaf disease is difficult, R-CNN was able to detect it with 83.34% overall accuracy using the soybean leaf image dataset. However, the accuracy of their detection was lower when they used only synthetic soybeans [10]. Pallapothala et al. developed procedure for the Rice leaf disease classification using CNN by using Rice leaf disease image dataset got overall accuracy of 78.2% and it is used only for four rice leaf diseases and got less accuracy [11]. The process for a novel multi-head CNN design that uses the plant village dataset to identify plant diseases was developed by Yasin Kaya et al. Its overall accuracy was 92.17%, and it is only used for DL-based plant disease detection by fusing RGB and segmented images [12].

3 Material and Methods

Several stages are needed to implement CNN architectures: gathering datasets is the first, followed by performance analysis and mappings for visualisations. began by gathering datasets and dividing them into distinct ratios for training, testing, and validation, in that order.

3.1 Dataset

In the grape fields, phone cameras captured the 8845 images that comprised the four different classes present in the grape leaf disease dataset. The information in the table below shows the classes that make up the dataset, and the figure below shows them.

Healthy Black Rot Black Measles Leaf blight

Fig.1. The example photos from the data collection on grape leaf disease

Table 1. An explanation of diseases that affect grape leaves

Class	No. of images in dataset	Description
Healthy leaf	1692	Healthy grape leaves are generally vibrant green, with a smooth surface. They have a consistent colour throughout, and there are no visible signs of lesions, spots, or discoloration.
Leaf blight	1722	Leaf blight is a fungal disease that affects grapevines. It typically manifests as irregular, brown lesions on the leaves. These lesions may have a reddish or purplish border, and in severe cases, they can cause defoliation.
Black measles	1920	Black measles, also known as black spot or anthracnose, is a fungal disease characterized by small, dark spots on leaves. These spots may coalesce, forming larger lesions. The disease can also affect other parts of the grapevine, such as stems and fruit clusters.
Black rot	1888	Black rot is a fungal disease that affects various parts of the grapevine, including leaves, fruit, and stems. It often starts as small, reddish-brown lesions on leaves, which can enlarge and develop a characteristic black centre. The disease can lead to significant crop loss if not managed.

3.2 Proposed Method

The model must first be loaded with the diseased grape leaf dataset. The collection consists of a variety of images related to diseases, each with a label. To increase efficiency, resizing and rescaling of photos is done in the data preprocessing step. The dataset is then split up into sets for testing, validation, and training. The second stage involves training a convolutional neural network architecture on photos of plant leaves across a number of epochs. Subsequently, the trained model's performance is assessed based on a number of factors, including f-score, accuracy, precision, and recall. In the stage of classification, the deep learning model recognizes disease on the picture.

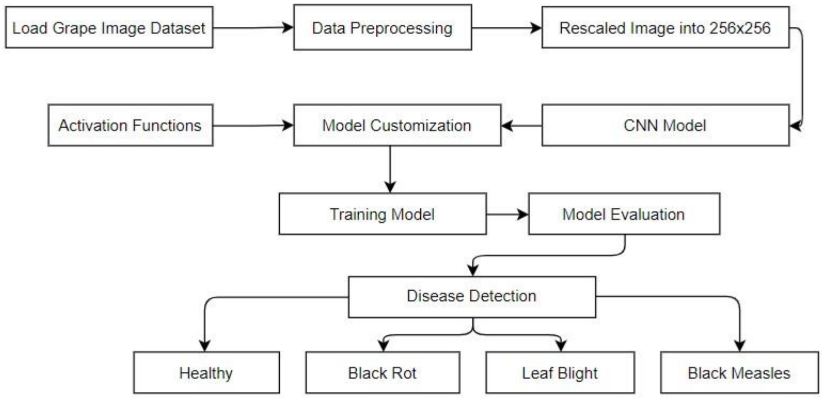

Fig.2. Overview of The Proposed Approach

The CNN architecture proposed by the proposed model consists of several layers: max-pooling, dropout, flatten, convolution with filters and rectified linear units (ReLU) activations, dense fully-connected, and output with sigmoid activation function. Kernel sliding is used by the convolution layer to extract features from the input image.

The pooling layer reduces the size of the data. A dropout layer is used to deactivate a portion of randomly selected nodes in order to prevent over-fitting. The flatten layer converts the multi-dimensional data into a one-dimensional array. In the end, the fully connected layer integrates the data, and the output layer classifies the data using a sigmoid function that enhances integration and yields high training and validation accuracy.

Convolutional Neural Network

Convolutional neural networks (CNNs), in particular, are a kind of deep learning system that can be very helpful for processing image and video data; in fact, the idea behind CNNs came from the human visual cortex. Features can be extracted from photographic data by humans with skill. These layers move across the input image using tiny filters, or kernels. In order to identify particular elements like edges, textures, or shapes, the filter compares itself to various regions of the image. In doing so, feature maps capturing these identified features are produced. By calculating the maximum or average value of a small neighbourhood of pixels, these layers down sample the feature maps. This keeps important information while reducing the size of the data.

These resemble the layers of a conventional neural network. They predict things about the image, such as its content, the presence of objects, or even possible actions, using the pooling layers' output as input.

The Sigmoid Function curve appears to be in the shape of a S.

The main reason we use the sigmoid function is that it exists between 0 and 1. It is specifically used for models whose output is a probability prediction. Since probability only occurs in the interval between 0 and 1, the sigmoid is the optimal choice. A differentiable function exists. This suggests that we can calculate the slope of the sigmoid curve between any two points.

Nadam, a different take on Adam that uses the Adam optimizer in conjunction with the Nesterov approach, has a somewhat quicker training time than Adam. NAdam optimizer, an enhancement of Adam, integrates Nesterov momentum and adaptive learning rates. It utilizes first and second moment estimates, adjusting the learning rates individually for each parameter during neural network training. The algorithm includes bias correction to counter initialization bias. The update rule involves a combination of momentum and adaptive scaling, facilitating faster convergence and improved performance.

4 Experimental Results

A dataset of grape images was used to train the suggested CNN model. The dataset is separated into three groups: testing, validation, and training. With 2.01% loss and 99.98% accuracy, the suggested model performed well. The optimal classification accuracy is obtained by utilising the activation function as a sigmoid and the Nadam optimizer.

Accuracy:

The simplest metric to determine the overall robustness of the model is accuracy. It calculates the proportion of accurately predicted cases to all instances.

$$Accuracy = \frac{TrueNegatives + Truepositives}{TruePositive + FalseNegative + FalseNegative}$$

Precision

Precision focuses on the right cases among the anticipated positive cases. It is defined as the ratio of all positive predictions to correctly predicted positive observations. When assessing the model's potential for false positives, accuracy plays a role.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Recall

The percentage of true positives (actual positives) that a model correctly identified is calculated using recall. Recall is best used when the cost of a false negative is significant.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

F1-score:

The precision and recall harmonic mean are represented by this. In terms of math, it looks like this:

$$F1 - Score = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

Table 2. Comparison of Different Optimizers

Optimizer	Accuracy	Precision Score	Recall Score	F1-Score
Nadam	99.99%	99.57 %	99.57 %	99.57%
RMSprop	99.91 %	99.14%	99.14%	99.14%
Adam	99.98 %	99.45%	99..45%	99.45%
Adagrad	96.47%	85.15%	85.15%	85.15%
Adadelta	70.44 %	66.24%	66.24%	66.24%

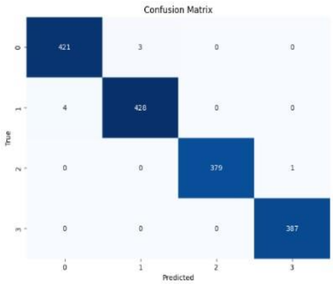

Fig.3. Adam

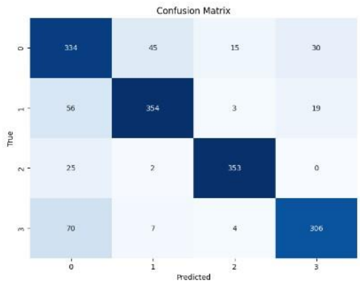

Fig.4. Adagrad

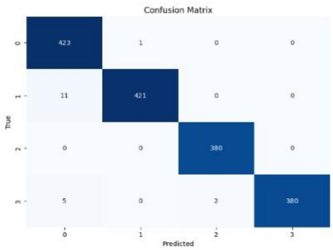

Fig.5. RMSprop

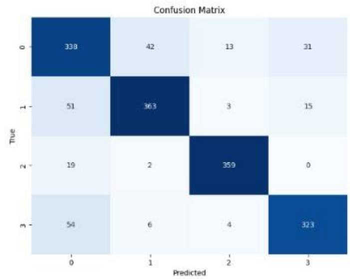

Fig.6. Adadelta

The performance of optimizers across 50 epochs with a Sigmoid activation function is shown in the table. With remarkable training (99.98%) and great test accuracy (99.45%), Nadam's selection is supported by these high numbers. Adam has an outstanding 98.05% training accuracy, but only slightly outperforms 86.26% in the test set. This choice points out how crucial it is to have reliable performance measurements that lead us to models with the best possible balance and dependability. Adagrad and Adadelta perform poorly, but RMSprop consistently shows correctness, which is consistent with the selection criteria.

Accuracy and Loss Curve

The previously mentioned curves display the model's real-time training history over the course of 50 training epochs. Training accuracy and validation accuracy increased while train loss and validation loss gradually decreased.

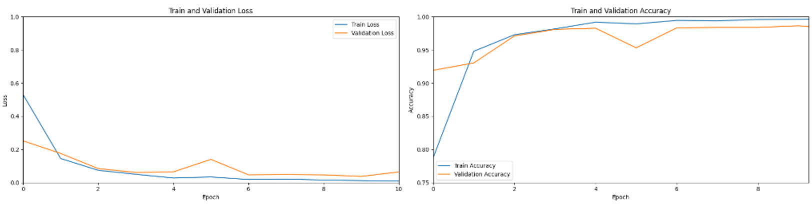

Fig.7. Accuracy and validation curve

Confusion Matrix

The confusion matrix, which shows the classifier's right and wrong predictions with count values, appears in the above figure. The matrix's diagonal cells show how many forecasts were right, while the non-diagonal elements show the number of predictions that went wrong.

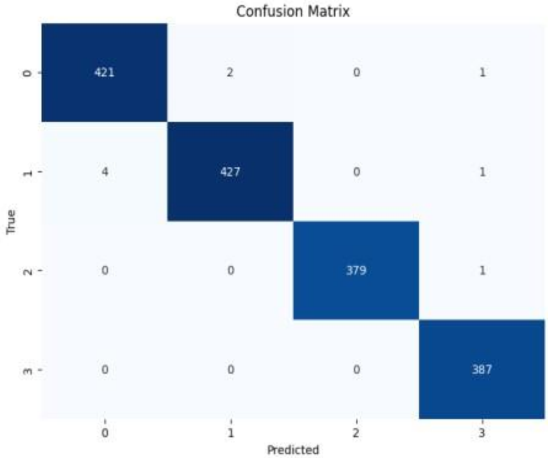

Fig.8. Confusion matrix of CNN model with Nadam optimizer

Remember that [421, 427, 379, 387] are the confusion matrix's diagonal elements. The good performance of the suggested model in terms of classifying grape plant diseases is thus validated by these high values. It is evident that the trained model had little trouble classifying all types of illnesses. For example, in black-rot disorders, 421 out of 424 cases were correctly predicted by the model; only 3 cases were misclassified. The results showed that the suggested model for predicting grape plant diseases was highly stable.

5 Conclusion and Future Scope

In conclusion, the traditional but time-consuming and subjective method of plant leaf disease detection via manual inspection is an essential part of sustainable agriculture. The shortcomings of manual inspection, such as its limited scope and inaccuracy, highlight the necessity for sophisticated methods. Convolutional neural networks (CNNs), one of the most recent advances in deep learning and machine learning, present a viable substitute for more effective and precise plant disease management. This study uses a CNN-based method to predict grape plant diseases by utilising the Grape Leaf image dataset. After extensive testing, the implementation discovered that CNN-Nadam performs better than other optimizers and achieves an astounding accuracy of 99.45% when using a sigmoid activation function. This study's effective use of deep learning models to classify plant diseases highlights the potential for automation in agriculture, which could result in early disease intervention, reduced crop losses, and eventually higher financial gains and agricultural productivity. The larger objective of automating the agriculture sector and guaranteeing sustainable food production in the face of pervasive plant diseases is supported by putting these findings into practice.

The future scope includes broadening the CNN-based disease classification system from grape leaves to various crops, implementing real-time monitoring, and developing user-friendly mobile applications for farmers and also developing real-time IOT based system for continuous monitoring. Integrating precision agriculture, continuous model training, and climate change adaptation enhances the system's versatility. Emphasis on global collaboration, model explain ability, and automated treatment recommendations aims to provide accessible tools for timely decision-making, contributing to sustainable agriculture and minimizing crop yield losses due to diseases.

References

1. M. Duot, L.M Nelson, R.C. Tyson (2013) Predicting the spread of postharvest disease in stored fruit,with application to apples based.
2. Kai Tian, Jiuhao Li, Jiefeng Zeng, Asenso Evans, Lina Zhang, Segmentation of tomato leaf images based on adaptive clustering number of k-means clustering.
3. Garima Shrestha, Deepsikha, Majolica Das, Naiwrita day (2020), Plant Disease Detection Using CNN.
4. Rahul S. Pachade, Dr. Avinash Sharma, Dr. Manoj Patil. Grape leaf disease detection using deep learning based vgg16 model
5. Malusi Sibiya and Mbuyu Sumbwanyambe (2019), A Computational Procedure for the Recognition and Classification of Maize Leaf Diseases Out of Healthy leaves using convolutional neural network.
6. Md. Ashiqul Islam1 , Md. Nymur Rahman Shuvo2 , Muhammad Shamsojjaman3 Shazid Hasan4 , Md. Shahadat Hossain5 , Tania Khatun6 (2021), An Automated Convolutional Neural Network Based Approach for Paddy Leaf Disease Detection.
7. Kumar, Voruganti Naresh, U. Sivaji, Gunipati Kanishka, B. Rupa Devi, A. Suresh, K. Reddy Madhavi, and Syed Thouheed Ahmed. "A Framework For Tweet Classification And Analysis On Social Media Platform Using Federated Learning." Malaysian Journal of Computer Science (2023): 90-98.
8. R. SujathaH, Jyotir Moy Chatterjee, NZ Jhanjhi, Sarfraz Nawaz Brohi (2020), Performance of deep learning vs machine learning in plant leaf disease detection.
9. Mehmet Metin Ozguven , Kemal Adem (2019), Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms.
10. Keke Zhan , Qiufeng wu, Yiping chen (2021), Detecting soybean leaf disease from synthetic image using multi-feauture fusion faster R-CNN.
11. Madhavi, K. Reddy, A. Vinaya Babu, A. Anand Rao, and S. V. N. Raju. "Identification of optimal cluster centroid of multi-variable functions for clustering concept-drift categorical data." In Proceedings of the International Conference on Advances in Computing, Communications and Informatics, pp. 124-128. 2012.
12. Yasin Kaya, Ercan Gursoy (2023), A novel multi-head CNN design to identify plant diseases using the fusion of RGB images