The Bioscan 18(3): 163-166, 2023
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
www.thebioscan.com

ANALYSIS OF DISEASE DETECTION IN COTTON PLANT LEAVES USING CONVOLUTIONAL NEURAL NETWORKS.

S ADITI APURVA*
IIIT Ranchi, Ranchi - 834004
e-mail: sapurva.2023dr103@iiitranchi.ac.in

KEYWORDS
CNN
Cotton
Diseases
Crop health

Received on :
20.03.2023

Accepted on :
26.07.2023

***Corresponding author**

ABSTRACT
Cotton (*Gossypium* spp.) is a vital crop that contributes significantly to global textile and oil seed industries. However, cotton cultivation faces considerable challenges due to various diseases affecting its leaves. The accurate and timely detection of diseases in cotton plant leaves is of paramount importance for ensuring optimal crop health and maximizing agricultural productivity. The CNN model was trained using a dataset that contained 10,000 leaf images, representing three diseases and healthy samples. The model was successful in detecting these diseases with a 98% accuracy rate.

INTRODUCTION

Cotton (*Gossypium spp.*) is a vital global crop, serving as a primary source for textiles and oilseed production. Cotton cultivation is facing many challenges, including the threat of diseases that can have a significant impact on crop yield and quality (plantdiseasehandbook,2021). Some commonly found major diseases in the cotton plant leaves in India are bacterial blight , curl virus, Fusarium wilt, alternaria leaf blight etc and some commonly found minor cotton plant diseases in India are Cercospora Leaf spot, Tobacco Streak Virus etc (https://agritech.tnau.ac.in/ ; Mehboob-ur- Rahman, 2017). To ensure sustainable cotton production and effective disease management, it is crucial to detect these diseases in a timely and accurate manner. In recent years, advancements in artificial intelligence and machine learning, particularly Convolutional Neural Networks (CNNs), have opened new avenues for revolutionizing disease detection in agriculture. In this paper a CNN model is proposed to carry out the experiment for the diseased leaves image detection process and to find better accuracy (Rai,2023; Udawant,2021; Shruthi, 2019; Vasavi1,2022). The CNN model which is discussed in the paper has come into the picture various times with wide range of datasets.

Convolutional Neural Network (CNN), is a class of deep neural networks, which has shown remarkable success in various image analysis tasks, ranging from object recognition to medical diagnosis. Its ability to automatically learn and extract intricate features from images has led to significant improvements in accuracy and efficiency. In the context of agriculture, CNNs have been increasingly employed for the detection and classification of plant diseases, including those affecting cotton plants (Prashant Udawant,2017; Adi,2021).

Detecting diseases in cotton plant leaves using CNN involves training models to recognize distinctive patterns and visual cues associated with various diseases. This approach holds great promise for transforming traditional disease diagnosis methods, which often rely on manual inspection and expert judgment, into automated, data-driven processes. By harnessing the power of Neural Networks for researchers and farmers can potentially detect diseases at early stages, enabling timely intervention and mitigation measures.

This paper aims to explore the application of Convolutional Neural Networks in the detection of diseases in cotton plant leaves. We will delve into the underlying principles of CNN, discuss the importance of robust datasets for training, examine pre-processing techniques to enhance model performance, and highlight recent advancements in disease detection using CNNs.

MATERIALS AND METHODS

Dataset Collection and Pre-processing

Acquire a diverse and representative dataset of high-resolution images of cotton plant leaves with varying stages of health and diseases.

Label the images with corresponding disease categories (e.g., Alternaria leaf spot, Fusarium wilt, Bacterial blight) for supervised learning.

Split the dataset into training, validation, and testing sets to assess the model's performance.

Apply data augmentation techniques (e.g., rotation, scaling, flipping) to increase the dataset's size and improve model generalization.

Normalize the pixel values of images to bring them within a common range (e.g., [0, 1]).

Model Architecture

Three convolutional layers are present in the architecture, followed by a max-pooling layer. This combination enables hierarchical feature extraction from input images, progressing from low-level features to higher level representations (Lee *et al.*, 2015). Filters are used by convolutional layers to capture features of varying complexity. The feature maps are down sampled by the subsequent max-pooling layers, which preserve essential information while reducing spatial dimensions. After creating the feature maps, they are flattened into a 1D vector to be inputted into two fully connected, dense layers. Using sequential connections, these dense layers, which have 128 neurons and 4 neurons, are connected. They learn high-level features and produce the final classification output, representing predicted probabilities for each class (Tugrul *et al.*, 2022).

Convolutional layers: These layers are designed to learn spatial hierarchies of features automatically and adaptively from the input images.

Max Pooling layers: These layers down sample the spatial dimensions of the feature maps, reducing computation and increasing the network's receptive field.

Flattening: Converts the 2D feature maps to a 1D vector before feeding them into the fully connected layers.

Fully connected layers: Neurons in these layers are fully connected to all the neurons in the previous layer. They process the features extracted by the convolutional layers.

Dropout: Helps prevent overfitting by randomly setting a fraction of input units to 0 during each update.

Output layer: Uses SoftMax activation for binary classification with two classes (diseased and not diseased).

Selection of an appropriate CNN architecture is based on the complexity of the problem.

Sequential model is selected for creating and predicting the output.

The selected model consists of three convolutional layers with max-pooling, followed by fully connected (dense) layers.

Used VGGNet for transfer learning by initializing the CNN with already trained weights from a large chunk of data.

Fine-tuning the model's weights to adapt it to the specific cotton leaf disease detection task.

Model Training of CNN

Implemented the chosen CNN architecture using a deep learning framework (e.g., TensorFlow).

Compiled the model with an appropriate loss function (e.g., categorical cross-entropy) and an optimizer (e.g., Adam).

Trained the model on the training dataset using mini-batch gradient descent. Monitor training progress with metrics such as accuracy and loss.

Utilized the validation dataset for hyper parameter tuning and prevented overfitting.

Applied techniques like learning rate scheduling and early stopping to optimize model convergence.

Model Evaluation of CNN

Evaluated the trained model's performance on the testing dataset using appropriate evaluation metrics (e.g., accuracy, precision, recall, F1-score).

RESULTS AND DISCUSSION

There are a total of three convolutional layers in the model. Convolutional Layer 1 applies 32 filters to the input images, resulting in feature maps of sizes None, 222, 222, and 32. Feature maps of sizes (None, 109, 109, and 64) are produced by Convolutional Layer 2 by applying 64 filters to the previous layer's output. The output of the previous layer is processed by Convolutional Layer 3 with 128 filters, leading to feature maps of size (None, 52, 52, 128). A convolutional layer is followed by three max-pooling layers in the model. Down sampling is performed by Max Pooling Layer 1 on the output of Convolutional Layer 1, which results in feature maps of sizes (None, 111, 111, 32). Down-sampling the output of Convolutional Layer 2 results in feature maps that are larger than None, 54, 54, 64. Down-sampling the output of

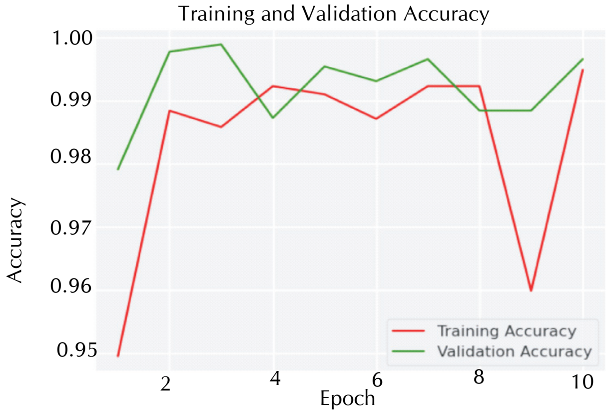

Figure 1 Training Accuracy Vs Validation Accuracy Graph

Figure 2 Training Loss Vs Validation Loss Graph

Convolutional Layer 3 results in feature maps that are size (None, 26, 26, 128). By flattening the 3D feature maps into a 1D vector, the data for the fully connected layers is prepared. Flatten Layer is used to transform feature maps into a 1D vector of size (None, 86528). Two fully connected dense layers follow the flattened layer in the model. 128 neurons are present in Dense Layer 1 and it is fully linked to the flattened input. It is trained to create high-level representations of the features that are extracted by convolutional layers. The number of classes or categories in the classification task corresponds to the number of neurons in the final dense layer. The final output of the model is represented by the predicted probabilities for each class. The model was assembled using the Adam optimizer, which is a stochastic gradient descent variation that has adaptive learning rates. The loss function for multi-class classification was determined to be categorical cross-entropy.

Two categories of leaves were collected for diseased cotton plants: diseased leaves and healthy leaves. Four different classes of leaves were taken for the experiment, 3 classes were having dataset of diseased leaves namely bacterial blight, curl virus, and curl virus and 1 class was having data for the healthy leaves respectively.

The CNN-based disease detection model demonstrates promising results in accurately identifying diseases in cotton plant leaves. The model achieves a high accuracy rate,

```
Epoch 1/10
107/107 [==============================] - 579s 5s/step - loss: 0.9992 - accuracy: 0.5998 - val_loss: 0.6166 - val_accuracy: 0.7835
Epoch 2/10
107/107 [==============================] - 297s 3s/step - loss: 0.6622 - accuracy: 0.7572 - val_loss: 0.5630 - val_accuracy: 0.8005
Epoch 3/10
107/107 [==============================] - 297s 3s/step - loss: 0.5718 - accuracy: 0.7817 - val_loss: 0.3419 - val_accuracy: 0.8906
Epoch 4/10
107/107 [==============================] - 288s 3s/step - loss: 0.4681 - accuracy: 0.8291 - val_loss: 0.4933 - val_accuracy: 0.8151
Epoch 5/10
107/107 [==============================] - 269s 3s/step - loss: 0.3385 - accuracy: 0.8777 - val_loss: 0.2348 - val_accuracy: 0.9257
Epoch 6/10
107/107 [==============================] - 272s 3s/step - loss: 0.3086 - accuracy: 0.8935 - val_loss: 0.1718 - val_accuracy: 0.9391
Epoch 7/10
107/107 [==============================] - 288s 3s/step - loss: 0.2461 - accuracy: 0.9163 - val_loss: 0.2508 - val_accuracy: 0.9087
Epoch 8/10
107/107 [==============================] - 285s 3s/step - loss: 0.2118 - accuracy: 0.9204 - val_loss: 0.1468 - val_accuracy: 0.9585
Epoch 9/10
107/107 [==============================] - 255s 2s/step - loss: 0.1998 - accuracy: 0.9274 - val_loss: 0.0581 - val_accuracy: 0.9795
Epoch 10/10
107/107 [==============================] - 253s 2s/step - loss: 0.2147 - accuracy: 0.9216 - val_loss: 0.1380 - val_accuracy: 0.9549
```

Figure 3 Epoch Result

Classification Report:

	precision	recall	f1-score	support
bacterial_blight	1.00	0.97	0.98	448
curl_virus	0.98	0.99	0.98	417
fussarium_wilt	0.98	0.99	0.98	419
healthy	0.97	0.99	0.98	425
accuracy			0.98	1709
macro avg	0.98	0.98	0.98	1709
weighted avg	0.98	0.98	0.98	1709

Figure 4 Classification Report

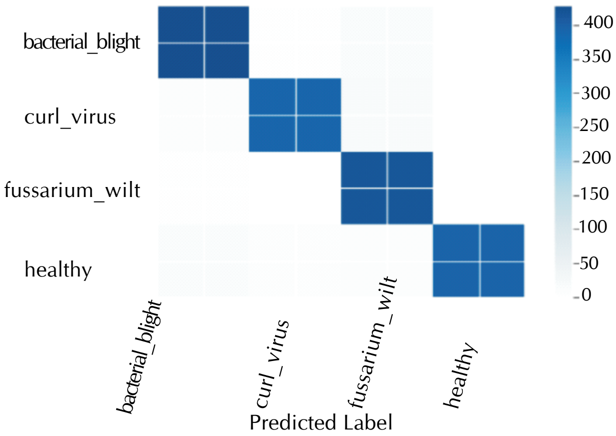

Figure 5 : Confusion Matrix of Proposed CNN Model

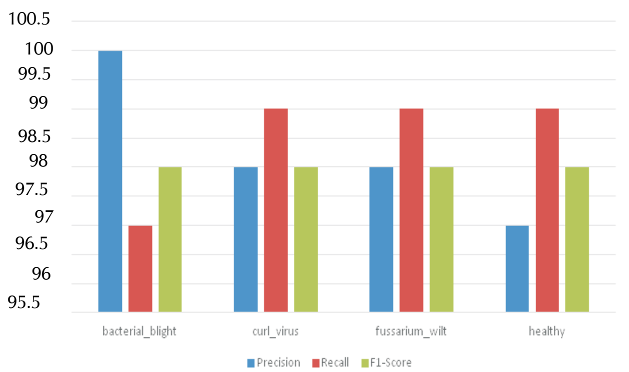

Figure 6 Accuracy Metrics Result for CNN Model

outperforming ANN methods and showcasing its potential as an innovative solution for disease management in cotton cultivation.

Results were compared based on the accuracy metrics such as accuracy, precision, recall, f1-score.

Accuracy (ACC): It is calculated as the ratio of correctly predicted instances to the total instances in the dataset. It is expressed as a percentage.

$$\text{Accuracy} = \frac{(\text{Number of Correct Predictions})}{(\text{Total Number of Predictions})}$$

Precision: It is also known as Positive Predictive Value which measures the accuracy of the positive predictions made by the model. It is the ratio of correctly predicted positive instances to the total instances predicted as positive.

Precision = (True Positives) / (True Positives + False Positives)

Recall (Sensitivity or True Positive Rate): It measures the ability of the model to correctly identify positive instances. It is the ratio of correctly predicted positive instances to the total actual positive instances.

Recall = (True Positives) / (True Positives + False Negatives)

F1-Score: It is the harmonic mean of precision and recall. It provides a balance between precision and recall, particularly useful when class distribution is imbalanced.

F1-Score = 2 * (Precision * Recall) / (Precision + Recall)

S. Bavaskar, V. Ghodake, G. Deshmukh, P. Chillawar and A. Kathole.2022. "Image Classification Using Deep Learning Algorithms for Cotton Crop Disease Detection," 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballari, India. pp. 1-8.doi: 10.1109/ICDCEC E53908.2022.9792911.

Convolutional Neural Networks offer a powerful and transformative approach to detecting diseases in cotton plant leaves (Adi,2021; Tugrul *et al*., 2022). This study showcases the potential of CNNs in automating disease diagnosis, potentially leading to improved disease control, enhanced crop yield, and sustainable cotton production (Suriya, 2023; Dahiya,2022; Yadhav,2020; Arivazhagan ,2018; Lee,2015; Zeiler, 2014). Further research and integration efforts can unlock the full potential of CNNs in revolutionizing agriculture practices.

Hence based on the above obtained result, it was concluded that Convolutional Neural Network is accurate in predicting the correct output for image detection task with an accuracy of 98%(Rai, 2023;Udawant,2021;Shruthi,2019; Vasavi,2022).

CONCLUSION

Therefore, it can be stated that the integration of Convolutional Neural Networks in disease detection holds promise for revolutionizing cotton agriculture by providing accurate, efficient, and automated solutions for identifying diseases in plant leaves. As we delve into the intricacies of this technology and its application in the context of cotton cultivation, we can anticipate transformative advancements in disease management strategies and the sustainability of cotton production.

REFERENCES

Arivazhagan, S. and VinethLigi, S. 2018. Mango leaf diseases identification using convolutional neural network. *Int J Pure Appl Math.* **120:**11067–11079.

Bacterial Blight in Cotton : https://www.cropscience.bayer.us/articles/dad/bacterial-blight-in-cotton

http://www.sciencedirect.com/topics/agricultural-andbiological-sciences/plant-diseases.

https://doi.org/10.3389/fpls.2017.01157.

https://agritech.tnau.ac.in/crop_protection/cotton_diseases/cotton_d1.html

https://doi.org/10.1007/s11042-023-14933-w

http://plantdiseas ehandbook.tamu.edu/industryspecialty/fiber-oilspecialty/cotton

https://doi.org/10.1016/*j.measen*.2022.100581.

https://agritech.tnau.ac.in/crop_protection/crop_prot_crop%20diseases_cash%20crops_cotton.html.

Lee, S. H., Chan, C. S., Wilkin, P. and Remagnino, P. 2015. Deepplant: Plant identification with convolutional neural networks. In Proceedings of the IEEE International Conference on Image Processing (ICIP). Quebec City, QC, Canada. pp. 452–456.

M. Adi *et al.*,2021. "An Overview on Plant Disease Detection Algorithm Using Deep Learning," 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom.pp.305 - 309.doi:10.1109/ICIE M51511.2021.9445336.

Mehboob-ur- Rahman ,Ali Q. Khan, Zainab Rahmat, Muhammad A. Iqbal and Yusuf Zafar. 2017. "Genetics and Genomics of Cotton Leaf Curl Disease, Its Viral Causal Agents and Whitefly Vector: A Way Forward to Sustain Cotton Fiber Security".

Vasavi, P., Punitha, A. and Venkat Narayana Rao, T. 2022." Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: a review," *International J. Electrical and Computer Engineering (IJECE).* 12(2): pp. 2079-2086. DOI: 10.11591/ijece.v12i2.

Rai, C. K. and Pahuja, R. 2023. Classification of Diseased Cotton Leaves and Plants Using Improved Deep Convolutional Neural Network. Multimed Tools Appl 82.PP. 25307–25325.

Sachin Dahiya, Tarun Gulati and Dushyant Gupta.2022.Performance analysis of deep learning architectures for plant leaves disease detection,Measurement:Sensors,Volume24,100581,ISS N26659 174.

Suriya, D. S. and Navina, N. 2023. Development and Analysis of CNN based Disease Detection in Cotton Plants. *J. Innovative Image Processing.* **5(2):** 140-160. doi:10.36548/jiip.2023.2.006

Tugrul, B., Elhoucine, E. and Eryigit, R. 2022. Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review. *Agriculture.* **12:** 1192

Udawant,P. and Srinath,P. 2019. Diseased Portion .Cassification and Recognition of Cotton Plants using Convolution Neural Networks, *International J. Engineering and Advanced Technology* (IJEAT) ISSN: PP. 2249–8958.

Shruthi, U., Nagaveni,V. and Raghavendra, B. K. 2019. "A Review on Machine Learning Classification Techniques for Plant Disease Detection," 2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 281-284. doi: 10.1109/ICACCS.2019.8728415.

Yadhav,S. Y., Senthilkumar,T., Jayanthy, S. and Kovilpillai, J. J. A. 2020. "Plant Disease Detection and Classification using CNN Model with Optimized Activation Function," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore,India.pp.564-569.doi:10.1109/ICESC 48915.2020.9155815.

Zeiler, M. D. and Fergus, R. 2014. Visualizing and understanding convolutional networks. In European conference on computer vision, Springer, Cham, pp. 818-833.