Grenze International Journal of Engineering and Technology, June Issue

A Comprehensive Review on Cotton Leaf Disease Detection using Machine Learning Method

Preeti Chopkar[1], Minakshi Wanjari[2], Pranjali Jumle[3], Pankaj Chandankhede[4], Sheetal Mungale[5] and Mohammad Shahnawaz Shaikh[6]
[1-6]Department of Electronics and Telecommunication Engineering, G H Raisoni College of Engineering, Nagpur, India
Email: preeti.chopkar.mtechcom@ghrce.raisoni.net, minakshi.wanjari@raisoni.net, pranjali.jumle@raisoni.net, pankaj.chandankhede@raisoni.net, sheetal.mungale@raisoni.net, mohammadshahnawaz.shaikh@raisoni.net

***Abstract*— In India, cotton holds a significant position as a key cash crop. Despite its economic importance, cotton crops are subject to a variety of illnesses that can reduce output and quality. Early detection of these diseases is crucial for minimizing damage and preserving the crop. Cotton is vulnerable to diverse ailments, includes powdery mildew, leaf curl, bacterial blight, target spot, leaf spot, and nutrient deficits. To put effective mitigation methods into place, it is imperative that one correctly recognizes these illnesses.**
The prevalence of cotton leaf diseases poses a substantial challenge for Indian farmers. This paper proposes the application of machine learning, specifically Convolutional Neural Networks (CNN), for disease detection as a promising solution. Leveraging technology for disease identification can empower farmers to take prompt actions, mitigating the risk of significant crop losses. Machine learning, with its demonstrated success in various domains, including agriculture, offers valuable tools for disease detection and classification. CNNs, known for their proficiency in image-related tasks, can learn to recognize disease-specific patterns in plant leaves.
This paper contributes by conducting a comprehensive review of prior research on crop disease detection and introducing an innovative approach that makes use of artificial intelligence techniques to detect illnesses in cotton leaves.

***Index Terms*— Convolutional Neural Networks (CNN), Deep Residual Neural Network (DRNN), Support vector machine (SVM),**

I. Introduction

The agricultural sector is essential for producing revenue and satisfying the growing demand for food in the world.[1]. With the growing population, the need for food continues to rise. Crop diseases pose a significant challenge to agriculture, leading to substantial losses in yield and compromising food quality and quantity [2]. Cotton, as a vital economic crop contributing to natural fiber production, holds importance in the textile industry, facilitating its expansion.
Traditional methods of diagnosing agricultural diseases are challenging for accurate crop assessment. The conventional approach involves hiring domain experts who visually inspect crops on-site, making it a labor-intensive and time-consuming process. Continuous crop monitoring is essential, but accessibility to professionals is a major hurdle in certain regions. Established techniques for evaluating crop diseases may not provide precise

Grenze ID: 01.GIJET.10.2.537

assessments.
Researchers have directed their efforts towards automating the detection of crop diseases in the agricultural sector. An automated approach to plant disease detection [4] is preferable for its accuracy, cost-effectiveness, and speed. Globally, plant diseases incur significant financial losses for the agricultural industry [5]. Efficient crop disease control is vital to ensure food quality and quantity. Early detection of plant diseases [6] is crucial to prevent transmission and enhance treatment efficacy.
The capacity of deep learning, one particular type of machine learning, to handle massive datasets including text, audio, and images has led to its rise in popularity. [7]. Deep learning can diagnose plant issues using leaf images [8,9] more efficiently due to its advanced models, solving complex problems quickly.
Convolutional neural networks (CNNs) stand out as highly effective algorithms that comprehend visual material and produce superior outcomes in finding objects, segmentation, retrieval, and image classification [10,11].
Utilising an already-trained model on a sizable dataset is known as transfer learning, contributing to accurate identification of various plant diseases [12].

II. Related Work

The author presented a Deep Residual Neural Network (DRNN) and a Plain CNN (PCNN) in [1] identify different cassava plant diseases, achieving 96.75% accuracy using DRNN. Another study [13] utilized CNN to distinguish healthy and unhealthy plant leaf images, reaching 99.76% accuracy through image segmentation, LBP, and grab-cut processes. Additionally, [15] described a genetic approach for identifying and categorizing leaf diseases, achieving 97.6% accuracy by employing various image processing algorithms and an SVM classifier on a feature dataset extracted from leaf images captured by a camera.
In [16], the author employed thermal imaging to detect diseases affecting crops, citing advantages such as rapid dynamic response, sensitive images, a vast detecting range, as well as exceptional precision. The investigation looked into a number of methods for diagnosing agricultural diseases. The economic outcomes for smallholder farmers in [17] are directly influenced by agriculture, plant life, and the presence of insect pests. The author suggested diagnosing diseases of plants with a CNN, which is frequently used for image classification.
An image processing technique to differentiate between healthy and diseased wheat crop stalks was proposed by the author of [18]. The author used neural networks and SVM classification algorithms to distinguish between healthy and unhealthy leaves using image processing techniques such as picture segmentation, feature extraction, texture features, shape features, and colour characteristics. Improved computational frameworks, particularly GPU incorporated processors, have been integrated into a deep residual network with a residual map in [19] for identifying pests.
In [20], a CNN-based classifier was suggested as a potential solution for cucumber diseases. Two datasets, including a variety of diseases, were used to train and validate the model. Popular algorithms for deep learning Convolutional neural networks and generative adversarial networks, for instance, and CNNs were discussed. The author gathered crop photographs from private online resources, employing the CafeNet architecture to achieve an 80 percent accuracy.

III. Methodology

Figure 1 illustrates the process of the machine learning model. Data collection is the initial step, followed by the removal of unwanted noise in the photographs. Annotation, often performed by agricultural professionals, is then carried out. Annotation is followed by data augmentation technique is utilised to expand the amount of the dataset. Initially, cotton leaf photos are obtained from real fields and subsequently cropped. The annotation process categorizes the dataset based on illness content. After annotation, data augmentation involves operations such as zooming, rotating, and horizontal and vertical shifting. The augmented data is then trained and evaluated using machine learning algorithms for various sickness detection. Accuracy, precision, recall, and f1 score are examples of assessment metrics. The goal of the model is to identify leaf illnesses early in the crop growth cycle, utilizing diverse leaf attributes from various models. A combination a variety of machine learning techniques to improve efficiency, precision, and F1 score.
There are two distinct methodologies available for the detection and the categorization of leaf diseases using both deep learning and traditional machine learning techniques.
The different parts that make up the systematic process of Leaf Disease Detection and labelling are as follows:

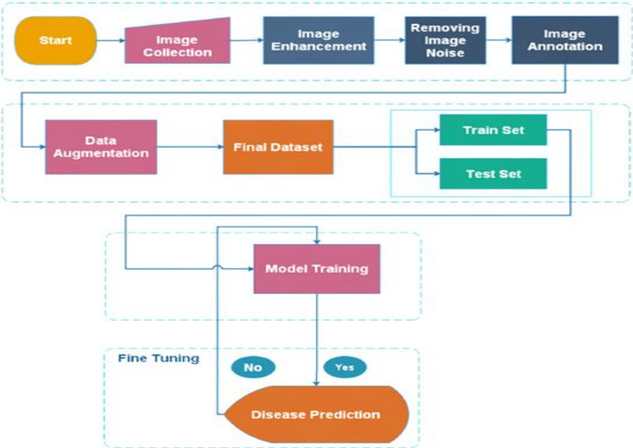

Figure 1. Workflow for Leaf Disease Detection

A. Data Selection

Cotton leaf images will be gathered and saved in either JPEG or PNG format. The dataset comprises four categories: Alternia leaf, Grey Mildew, Leaf Reddening, and Healthy leaf.

Figure2. Alternia Leaf

Figure3. Grey Mildew

Figure4. Leaf Rddening

Figure5. Healthy Leaf

B. Data Pre-Processing

Pre-processing involves a sequence of transformations applied to an original image to format the data in a predetermined analytic form.

C. Data Transformation

Images are initially converted into NumPy arrays to standardize RGB values, enabling analysis by classification models. Irregularities in the image dataset, such as varying sizes and shapes, are common. To address this, a data augmentation strategy utilizing the Keras Image Data Generator is employed. This involves altering training set images to enhance the model's classification performance. For the test set, to increase validation accuracy, batch normalization and dropout layers are applied to the CNN model in addition to augmentation.

D. Integration of Machine Learning (ML) and Transfer Learning (TL) Techniques

Various models are created and evaluated to identify the most accurate and efficient one. Two methods are used for model creation: the first involves building models from scratch, including segmentation and feature extraction for applying Random Forest and Support Vector Machine. The second method employs deep learning models (Inception v3, VGG16, and ResNet) derived from CNN using the ImageNet dataset's pre-trained weights. Retraining and analysis are limited to the final layers. Model parameters are fine-tuned through multiple tests to determine the optimal model for identifying and classifying cotton diseases.

E. Process and Evaluation Method

Model performance is evaluated scores for memories, precision, and F-1 score. The best model is selected based on these metrics.

IV. Machine Learning Technique

A. Convolution Neural Network (CNN)

A simple Convolutional Neural Network (ConvNet), a deep learning technique for image recognition, is the first model to be developed. In order to make recognition easier, it gives various regions of an input image weights and biases. This study's CNN model consists of layers such activation functions, pooling, convolutional neural networks, and fully connected. The RGB input image is interpreted by these layers, which provide a feature map that identifies low-level features like curves and edges and is used to group together cotton leaf diseases. Variable parameter filters improve the feature map and increase the efficacy of the model.

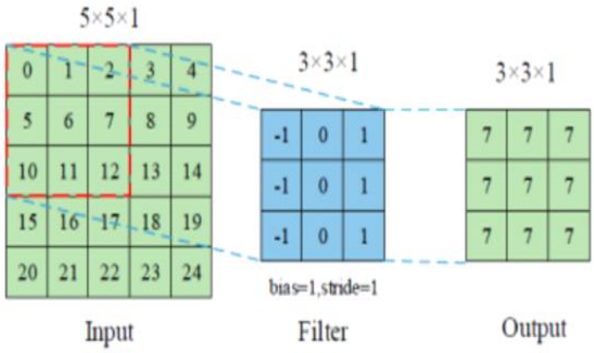

Figure 6. Convolution Process

Activation Function

To enhance the performance of a neural network, introducing non-linearity is crucial. Immediately following the first fourier layer, a non-linear activation function is used. The rectification of the linear unit function, or ReLU, was used in this study, is utilized to augment the model's non-linearity by setting all negative values to 0.

Pooling Layer

Reducing the spatial size of features is a critical function of the pooling layer, which lowers the amount of computing work and dimensionality needed for data processing. Consequently, this maintains the instruction cycle of the model's efficacy. The max-pooling layer is used in this study to hold onto the largest value, which is based on kernel size, in the covered region of photographs of cotton plants.

Completely Networked Layer

The fully connected layers in a CNN serve as the concluding stages capable of identifying features strongly associated with the output class. This culminates in a one-dimensional vector, acquired by flattening the outcomes of preceding pooling layers.

Dropout Function

Implemented to alleviate model overfitting, the dropout function selectively excludes a random subset of neurons within a layer.

Softmax Function

Serving as the classification layer, the softmax function is applied subsequent to the study's thick or fully-connected layer, especially in the instance of multiclass categorization.

TABLE I. COMPARISON AND ANALYSIS OF PREVIOUS WORK

Reference with Author name	Methodology	Crop Used	Type of Disease	Discription
[12] Chen, J.; Chen, J.; Zhang, D.; Sun, Y.; Nanehkaran, Y.	VGGNET Transfer Learning	Rice, Maize	Gray leaf spot, Common rust, Northern leaf blight	The author used VGGNET Transfer Learning algorithm for finding the rice leaves disease. By using this algorithm , the author achieved the accuracy of 92% for a dataset of 500 images of rice and 466 images of maize.
[20] Zhang, J.; Rao, Y.; Man, C.; Jiang, Z.; Li, S.	GrabCut, SVM, Deep CNN, GAN	Cucumber	anthracnose, downy mildew, and powdery mildew	The author used two stage segmentation process using GrabCut with SVM technique for extracting the lesions from leafspot. Author worked on database of 600 images of cucumber for extracting various features like Color, Texture and Border, and achieved the accuracy of 96.11%.
[21] Chopda, J.; Raveshiya, H.; Nakum, S.; Nakrani	Decision Tree Classifier	Cotton	Cotton Disease	The author used a decision tree classifier for classifying the various parameters like temperature, soil moisture. These parameters used for detecting the Cotton leaf disease.
[22] Rothe, P.; Kshirsagar, R.V	Image Segmentation, Gaussian filter, Graph cut	Cotton	Leaf Spot	Manual Image processing method is used for identifying the cotton diseases. This method is time consuming as it is manual and used Image processing techniques.
[23] Xu, R.; Li, C.; Paterson, A.H.; Jiang, Y.; Sun, S.; Robertson, J.S	CNN	Cotton	Flower species	The author used unmanned serial images for detecting cotton diseases. The author worked on cotton flower species.

V. CONCLUSION

The utilization of machine learning in agriculture has demonstrated significant potential in automating the identification of plant diseases. As machine learning methods continue to advance and become integrated into agricultural practices, it is anticipated that disease control will become more efficient, leading to higher crop yields. Applying machine learning to the detection of cotton leaf diseases holds substantial promise for substantial advancements in agriculture, manifesting in efficient disease control, cost-effectiveness, increased yields, and improved quality of the cotton crop output.

The adoption of Convolutional Neural Networks (CNNs) for detecting cotton leaf diseases marks a noteworthy stride in precision agriculture. This technology provides a dependable, efficient, and automated approach to identify and categorize diseases in a manner that contributes to the enhancement of agricultural practices.

REFERENCES

[1] Pantazi, X.; Moshou, D.; Tamouridou, A. Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers. Comput. Electron. Agric. 2018, 156, 96–104.

[2] Toda, Y.; Okura, F. How convolutional neural networks diagnose plant disease. Plant Phenomics 2019, 2019, 9237136.

[3] Dunne, R.; Desai, D.; Sadiku, R.; Jayaramudu, J. A review of natural fibres, their sustainability and automotive applications. J. Reinf. Plast. Compos. 2016, 35, 1041–1050.

[4] Iqbal, Z.; Khan, M.A.; Sharif, M.; Shah, J.H.; ur Rehman, M.H.; Javed, K. An automated detection and classification of citrus plant diseases using image processing techniques: A review. Comput. Electron. Agric. 2018, 153, 12–32.

[5] Wang, G.; Sun, Y.; Wang, J. Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. Comput. Intell. Neurosci. 2017, 2017, 2917536

[6] Saleem, M.H.; Potgieter, J.; Arif, K.M. Plant disease classification: A comparative evaluation of convolutional neural networks and deep learning optimizers. Plants 2020, 9, 1319.

[7] Hughes, D.P.; Salathe, M. An Open Access Repository of Images on Plant Health to Enable the Development of Mobile Disease Diagnostics. 2015. Available online: http://arxiv.org/abs/1511.08060 (accessed on 15 November 2021).
[8] Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. Sensors 2018, 18, 2674.
[9] Lin, K.; Gong, L.; Huang, Y.; Liu, C.; Pan, J. Deep learning-based segmentation and quantification of cucumber powdery mildew using convolutional neural network. Front. Plant Sci. 2019, 10, 155
[10] Liu, X.; Deng, Z.; Yang, Y. Recent progress in semantic image segmentation. Artif. Intell. Rev. 2018, 52, 1089–1106.
[11] Karlekar, A.; Seal, A. SoyNet: Soybean leaf diseases classification. Comput. Electron. Agric. 2020, 172, 105342.
[12] Chen, J.; Chen, J.; Zhang, D.; Sun, Y.; Nanehkaran, Y. Using deep transfer learning for image-based plant disease identification. Comput. Electron. Agric. 2020, 173, 105393.
[13] Amin, H.; Darwish, A.; Hassanien, A.E.; Soliman, M. End-to-End Deep Learning Model for Corn Leaf Disease Classification. IEEE Access 2022, 10, 31103–31115.
[14] Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. Front. Plant Sci. 2016, 7, 1419.
[15] Singh, V.; Misra, A. Detection of plant leaf diseases using image segmentation and soft computing techniques. Inf. Process. Agric. 2017, 4, 41–49.
[16] Zhu, W.; Chen, H.; Ciechanowska, I.; Spaner, D. Application of infrared thermal imaging for the rapid diagnosis of crop disease. IFAC-PapersOnLine 2018, 51, 424–430.
[17] Zhang, J.-H.; Kong, F.-T.;Wu, J.-Z.; Han, S.-Q.; Zhai, Z.-F. Automatic image segmentation method for cotton leaves with disease under natural environment. J. Integr. Agric. 2018, 17, 1800–1814.
[18] Kumar, M.; Hazra, T.; Tripathy, S.S. Wheat Leaf Disease Detection Using Image Processing. Int. J. Latest Technol. Eng. Manag. Appl. Sci. 2017, 6, 73–76.
[19] Cheng, X.; Zhang, Y.; Chen, Y.; Wu, Y.; Yue, Y. Pest identification via deep residual learning in complex background. Comput. Electron. Agric. 2017, 141, 351–356.
[20] Fujita, E.; Kawasaki, Y.; Uga, H.; Kagiwada, S.; Iyatomi, H. Basic investigation on a robust and practical plant diagnostic system. In Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016.
[21] Zhang, J.; Rao, Y.; Man, C.; Jiang, Z.; Li, S. Identification of cucumber leaf diseases using deep learning and small sample size for agricultural Internet of Things. Int. J. Distrib. Sens. Netw. 2021, 17.
[22] Chopda, J.; Raveshiya, H.; Nakum, S.; Nakrani, V. Cotton crop disease detection using decision tree classifier. In Proceedings of the 2018 International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, India, 5 January 2018; pp. 1–5.
[23] Rothe, P.; Kshirsagar, R.V. Automated extraction of digital images features of three kinds of cotton leaf diseases. In Proceedings of the 2014 International Conference on Electronics, Communication and Computational Engineering (ICECCE), Hosur, India, 17–18 November 2014; pp. 67–71.
Xu, R.; Li, C.; Paterson, A.H.; Jiang, Y.; Sun, S.; Robertson, J.S. Aerial images and convolutional neural network for cotton bloom detection. Front. Plant Sci. 2018, 8, 2235.
[24] Vidhale, Bhushan, Ganesh Khekare, Chetan Dhule, Pankaj Chandankhede, Abhijit Titarmare, and Meenal Tayade. "Multilingual Text & Handwritten Digit Recognition and Conversion of Regional languages into Universal Language Using Neural Networks." In 2021 6th International Conference for Convergence in Technology (I2CT), pp. 1-5. IEEE, 2021.
[25] Kushwaha, Vikas, and Priya Maidamwar. "BTFCNN: Design of a brain tumor classification model using fused convolutional neural networks." In 2022 10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22), pp. 1-6. IEEE, 2022.
[26] Kosarkar, Nidhi, Pallavi Basuri, Poonam Karamore, Prachi Gawali, Pradnya Badole, and Pranjali Jumle. "Disease Prediction using Machine Learning." In 2022 10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22), pp. 1-4. IEEE, 2022.
[27] Gadde, Abhishek, Abhinandan Gorde, Aditya Bopte, Nidhi Kapgate, Aashay Nevatia, and Swapna Choudhari. "Employee Alerting System Using Real Time Drowsiness Detection." In 2022 10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22), pp. 1-5. IEEE, 2022.
[28] Yamsaniwar, Sucheta, Surekha Tadse, Saikat Ranajit, and Rutvik Walde. "Glaucoma and Cataract Hybrid Classifier." In 2022 10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22), pp. 1-6. IEEE, 2022.
[29] Joshi, Sonali, and Preeti Bajaj. "Design & Development of portable vata, pitta & kapha [VPK] pulse detector to find Prakriti of an individual using artificial neural network." In 2021 6th international conference for convergence in technology (I2CT), pp. 1-6. IEEE, 2021.
[30] Mate, Nikita, Durva Akre, Gaurav Patil, Gopal Sakarkar, and Thomas Anung Basuki. "Emotion Classification of Songs Using Deep Learning." In 2022 International Conference on Green Energy, Computing and Sustainable Technology (GECOST), pp. 303-308. IEEE, 2022.

[31] Choudhary, Swapna, and Sanjay Dorle. "A quality of service-aware high-security architecture design for software-defined network powered vehicular ad-hoc network s using machine learning-based blockchain routing." Concurrency and Computation: Practice and Experience 34, no. 17 (2022): e6993.
[32] Kotadi, Chinnaiah, K. Mithun Chakravarthi, Srihari Chintha, and Kapil Gupta. "Analysis of COVID-19 Data Using Machine Learning Algorithm." Object Detection by Stereo Vision Images (2022): 147-157.
[33] Yelne, Ajinkya, and Archana Raut. "Digital Health-Care System for Smart IPD Booking." Grenze International Journal of Engineering & Technology (GIJET) 8, no. 2 (2022).
[34] Kshirsagar, Pravin R., Hariprasath Manoharan, S. Shitharth, Abdulrhman M. Alshareef, Nabeel Albishry, and Praveen Kumar Balachandran. "Deep learning approaches for prognosis of automated skin disease." Life 12, no. 3 (2022): 426.
[35] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955.