International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Learning Models for Cotton Leaf Disease Detection with VGG-16

Awad Bin Naeem*[1], Biswaranjan Senapati[2], Alok Singh Chauhan[3], Sumit Kumar[4], Juan Carlos Orosco Gavilan[5], Wael M. F. Abdel-Rehim[6]

Submitted: 10/11/2022 **Accepted**: 13/02/2023

Abstract: Cotton, potatoes, tomatoes, and chilies are the principal crops exported from Pakistan. Cotton is produced in greater quantities in India, China, and the US than in Pakistan. Eighty percent of all public oil is made up of cotton. Like other crops, it is being attacked by viruses, which lowers production and reduces economic revenue. There is misinformation on how illness diagnosis and care affect a region's ability to produce more. In this research, deep convolutional neural networks (CNN) were trained to recognize three forms of cotton leaf disease using a transform learning approach (Cotton leaf curl virus, fusarium wilt, bacterial blight). This study's main goal was to develop a single framework that could handle the challenging process of finding, identifying, and diagnosing cotton leaf disease. Additionally, to improve their performance on datasets of healthy and allergic cotton leaves, bigger weight parameter optimization using the Adam and RMSProp optimizers was explored. When compared to the other DL meta-architectures, Inception-VGG-16 trained with the feature extractor showed the greatest mean average accuracy. The recommended method was found to be new since it distinguished between leaf types that were healthy and those that were unhealthy. Using the DL approach to accurately identify cotton leaf disease would help to avoid the adverse effects of dietary management issues. On photos of cotton leaves, the trained model recognizes and labels the four classes (Cotton leaf curl virus, bacterial blight, fusarium wilt, healthy). CNN was 98% accurate overall.

Keywords: *Leaf Disease, Adam Optimizers, VGG-16, CNN.*

1. Introduction

The recent sharp drop in agricultural production puts the world's food security in jeopardy. In the globe, Pakistan is rated sixth in terms of cotton output. Pakistan's GDP is around 0.6% cotton-related. Its production has decreased gradually in recent years, by as much as 22%, and if it continues to decline at this pace, productivity will soon suffer [1]. The whitefly, or Bemisia Tabaco, is one of the common pests that attack plants and carries several viruses, including cotton leaf curl disease (CLCUD). One of the top 100 worst alien invasive species is the whitefly. A cotton field and nearby crops may get infected by whiteflies, which may reduce plant growth by up to 50%. As a result, this tiny bug harms local and international food crops as well as cash crops [2].

Field inspection is steadily becoming more automated thanks to advancements in technology, including spatial drones for detection and prediction and cutting-edge automated disease detection algorithms. Before there were no effective methods for identifying plant diseases, therefore one had to personally examine each plant and provide the proper treatment, which was very time-consuming, labour-intensive, and needed superior professional knowledge and abilities [3]. Since they are the major source of human energy generation and provide nutritional, therapeutic, and other benefits, plants are acknowledged as being crucial. Plant diseases may harm the leaf at any point during crop cultivation, causing significant losses in crop productivity and financial value. Therefore, identifying leaf disease is important in the agricultural business [4].

The cotton crop is vulnerable to several illnesses as a result of weather fluctuations, insects like the pink bollworm, and a variety of other causes. These illnesses reduce agricultural output, and farmers now identify the illnesses based on their personal experience. However, across vast plantation lands, these sorts of eye-only inspections do not provide reliable data [5]. Thus, it is essential to create an automated, precise, and cost-effective technique for identifying leaf diseases. Using a deep learning technology called Convolutional Neural Network (ConvNet/CNN), this research aims to identify diseased cotton leaves. Numerous diseases, such as leaf spot, target spot, bacterial blight, nutrient deficit, powdery mildew, leaf curl, etc. may affect cotton. For appropriate action to be taken, illness detection must be done accurately [6]. Deep learning is crucial for correctly diagnosing plant diseases. The suggested model built on meta–Deep Learning is utilized to precisely detect various

[1] *Ph.D Student in Computer Science NCB&E, Multan, Pakistan*
[2] *Doctor in Computer and Data Science Parker Hannifin Corp*
[3] *ABES Engineering College, Ghaziabad, India*
[4] *Integrated Academy of Management and Technology, Ghaziabad, India*
[5] *Universidad Privada del Norte, Campus Virtual, Lima, Peru*
[6] *Department of Computer Science, Suez University, Suez, Egypt*
** Corresponding Author Email: Awadbinnaeem@gmail.com*

illnesses of cotton leaves. For this research, we collected pictures of cotton leaves from the field. The collection includes 2385 pictures of both healthy and ill leaves. With the aid of the data augmentation technique, the dataset's size was expanded. The dataset was trained using our suggested model—the meta deep learn leaf disease diagnosis model—as well as Custom CNN, VGG16 Transfer Learning, ResNet50, and ResNet50[7].

In this study, we examined transfer learning methods for identifying diseases and developed a sequential deep convolutional neural network for distinguishing between healthy and unhealthy plants. The suggested model outperforms pre-trained models like VGG16, ResNet50, and ResNetl52V2 in terms of accuracy and performance. Therefore, the suggested approach may be used to monitor huge fields for quicker diagnosis and treatment for higher productivity.

With greater acreage than any other crop, cotton is Pakistan's most significant crop. 80% of the nation's oil and meal are made from cottonseed. Cotton harvests account for the lion's share of the country's export revenues. Cotton leaf disease, caused by several viruses, has had a considerable influence on cotton productivity and economic earnings. The main purpose is to find leaf disease. It also reveals the kind of ailment that the leaf is suffering from at the time.

A completely automated technique has been suggested to handle this problem, removing the danger of human mistakes and shortening the time required to establish the severity of the illness. The proposed approach will benefit farmers and researchers by making cotton leaf disease diagnosis and management easier.

The structure of this work is as follows: The review of literature is included in Section 2. The methodology of the investigation is presented in Section 3. Section 4 explains the results. Section 5 finally describes the conclusion and future work.

2. Literature Review

Plant diseases are significant problems in agricultural production; if they are not discovered promptly, they will have a detrimental effect on crop output and quality. As is well knowledge, early diagnosis and warning being the cornerstones of successful plant disease prevention and control, which are vital in management and decision-making [8]. However, in many nations and regions, the main method for identifying plant diseases still relies on the visual observations of professionals or seasoned farmers [9]. This conventional method has numerous drawbacks, such as the time-consuming manual observations required in big farms and the prohibitive cost of frequent expert consultations [10]. Therefore, the automated diagnosis of plant diseases, which attempts to identify the signs of plant diseases as soon as they occur on leaves, is of considerable actual value. In India, one of the main drivers of economic development has been agriculture [11]. The soil type, local weather, and crop's economic worth are taken into consideration when the farmer chooses a crop. The agricultural sectors started looking for innovative ways to boost food production as a consequence of growing populations, changing weather patterns, and political unpredictability [12]. A farmer uses pesticides to reduce pests, prevent illnesses, and boost crop output. Due to industrial agriculture, poor yield, economic losses, and crop diseases, farmers are having issues [13]. Consequently, the emphasis of illness identification and severity is on the need to be characterized as suitable.

The primary source of food, money and employment for rural residents in economically developing nations is agriculture. Crop loss caused by plant diseases, which reduces output by 20 to 30%, is the main factor affecting agricultural productivity [14]. Conventional methods have been used to diagnose illnesses in an attempt to prevent such losses, but they are inaccurate [15]. To reduce losses brought on by such illnesses, early and accurate detection of plant diseases is crucial. But sometimes such harvests and grains suffer a significant amount of damage [16], if not destruction, owing to a lack of suitable cultivating knowledge, expertise, and a sense of disease prediction. This results in severe hardship for the farmers as well as for the nation's economic growth [17]. So, to lessen the loss caused by diseases of plant leaves, this research attempts to integrate a portion of agriculture with the use of artificial intelligence [18]. We employed transfer learning models built with several CNN architectures, such as ResNet50, VGG19, InceptionV3, and ResNet152V2, to address this issue. To determine which technique performs best at diagnosing cotton leaf illnesses, we conducted trials using all four methods on the standard dataset of cotton leaves [19].

Numerous plant diseases have a substantial impact on the crop's quality and yield. Early identification of these disorders is thus very beneficial. One of the significant crops that is produced in great amounts and has a high economic value is the tomato [20]. The pace at which several tomato illnesses are affecting the crop is worrying. In this study, we implemented two convolutions neural network (CNN)-based models, Google Net and VGG16, for the classification of tomato leaf diseases [21]. With the use of deep learning, the proposed effort seeks to identify the most effective method for the issue of tomato leaf disease detection. Complex deep-learning models have shown acceptable performance in diagnosing plant illnesses and leaf diseases [22]. But training the model takes a long time and a lot of processing power because of the algorithm's intricacy and the vanishing gradient issue. In this study, we investigated how well deep neural networks predicted healthy and ill cotton plants and leaves [23]. Three distinct models, including VGG16, ResNet50, and Mobile Net, were compared. VGG16 and Mobile Net models provide the greatest outcomes on the train set, validation set, and test set, according to our findings [24]. Accuracy, loss, and the number of accurate predictions provided by the model are the metrics taken into account while analyzing the models [25]. The identification and detection of leaf diseases have received an increasing amount of study and attention since the creation and widespread use of intelligent agricultural systems [26]. To explore the detection and categorization of apple leaf illnesses, we employed data sets of healthy leaves, apple grey-spot disease, black star disease, cedar rust disease, and healthy leaves [27]. ResNet and VGG convolutional neural network models, as well as SVM classifiers for image segmentation, were employed for comparison and enhancement [28].

Plants are recognized as being essential because, because of their nutritional, medicinal, and other characteristics, they are the primary source of human energy. At any time during crop farming, plant diseases may affect the leaf, resulting in significant losses in crop output and market value [29]. As a result, in the agricultural business, identifying leaf disease is critical. It does, however, need considerable work, more preparatory time, and extensive plant pathogen expertise [30].

3. Material and Methods

This section contains the methodology of the study

3.1 Proposed Framework

The next three parts in a row will implement the categorization and detection of cotton leaf diseases.

- Building the database of images.
- Finding and extracting features.
- Find and classify the illnesses affecting cotton leaves.

Phase-1 involves gathering the data set and preparing it. In phase-2, VGG-16 is used to identify illnesses in cotton leaves. Creating a model for standardized non-destructive analysis to detect cotton leaf diseases is phase-3. Figure 1 presents the suggested work's technique.

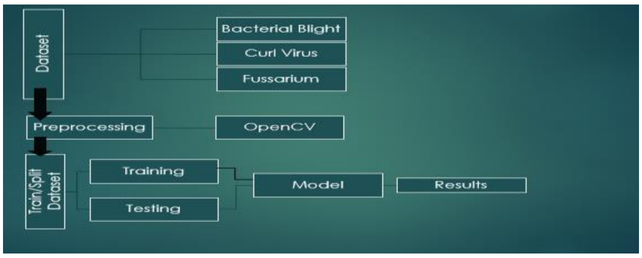

Fig. 1. Flow Diagram

3.2 Dataset Collection

Images of cotton leaves were carefully gathered from several cotton fields to create the dataset. Using a Canon D3D camera, more than seven hundred high-resolution pictures of cotton leaf diseases were captured. The dataset for cotton leaf disease was divided at a ratio of seventy-five percent to twenty-five percent into training and validation datasets.

3.3 Pre-processing and Feature Extraction

Preprocessing of input photographs often takes place in the layer labelled "input." It's conceivable that further preprocessing steps like size and normalization will be required. VGG-16 CNNs, on the other hand, needed substantially less preprocessing than other neural networks. However, to eliminate pointless differences, a simple preprocessing step is necessary. To identify features and preserve the associations between pixels in image processing, convolution is a crucial transformation method. This is performed using convolution in VGG-16 CNN. It does this by recognizing a certain feature using the proper convolution kernel. A single convolution operation may be carried out in both the horizontal and vertical dimensions. This feature extraction approach is effective and efficient when dimensions are appropriately lowered and a data set known as a feature map is generated as a consequence. To find and eliminate certain features from the original picture, the kernel of each feature identifier is inspected. In the end, a map illustrating the distribution of these characteristics is produced using elevation data. The photographs must be processed using specialized generic feature characterizing techniques, such as convolution, rather than directly analyzing them to correctly deconstruct the symbolic information present in the photos. The basic purpose of convolution in CNN is to automatically extract useful features. There aren't many convolution kernels, and they stay in the same settings throughout the identification of an individual item in the picture. Each receptive neuron has a weight-sharing characteristic as a result.

3.4 Recognition Process

A highly layered structural neural network, VGG-16 is a convolutional neural network created by Google. In a CNN's fundamental function, there are three different kinds of layers: coevolutionary, pooling, and entirely linked. The symmetric and asymmetric building blocks of Inception v3 include convolutions, average pooling, maximum pooling, dropouts, and completely connected layers. Due to the great efficiency of these neurons, a bottleneck layer of eleven convolutions significantly minimizes the amount of computation that must be performed. The Inception architecture was created for this purpose: to extract characteristics from various layers of abstraction. To do this, the same network module was used for all three convolutions.

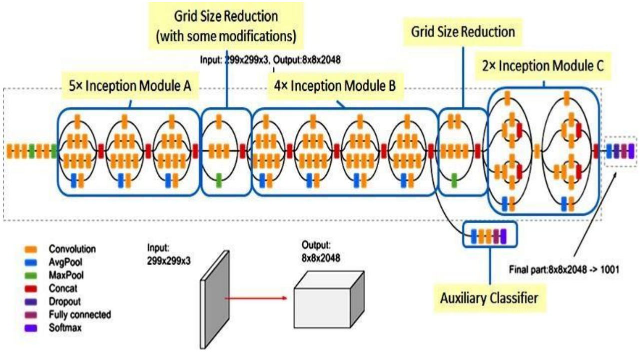

Fig. 2. CNN's Recognition Process

To conduct grade classification for the most common cotton leaf disease in Pakistan, Inception v3 was used in figure 2. The VGG-16 model takes a picture of the various cotton leaf disease classes, processes it, and then categorizes it according to the relevant cotton leaf disease class. Here, the CNN model receives as input a picture of cotton leaf disease. After going through many convolution and pooling layer processes, it serves as the input for fully connected layers before being subjected to the SoftMax algorithm to provide results for multiclass classification.

3.5 Comparative Study

It's possible to find a CNN algorithm architecture that's right for your data and your application by experimenting with a different number of layers and structures. On this dataset, we evaluate VGG16, Inspection, and Resnet for both speed and accuracy.

3.6 Inception Module

On the ImageNet dataset, the image recognition model Inception v3 has been proven to achieve higher than 78.1% accuracy. This is the culmination of the efforts of numerous researchers over several decades. With Inception V3, you get convolutions, pooling, dropouts, and fully connected layers in symmetric and non-symmetric building blocks. The batch norm is applied to activation inputs in Inception architecture. Softmax is used to determine the loss using a large data set. Convolution filters with some pooling are applied to the same input in the inception module and the output is concatenated. This architecture has the advantage of improving network speed by extracting multi-level characteristics in this manner. The (5x5) filter, for example, collects both global and local data.

3.6.1 VGG16:

The CNN architecture VGG16 is regarded as the best vision model architecture. The enormous number of hyper-parameters makes VGG16 stand out among other CNN models. It is common practice to utilize layer kernels of 3x3 sizes with a stride of 1 in convolution. Maximum pooling for a 2x2 kernel with stride 2. Max-pooling and convolutional layers are positioned continuously throughout the VGG16 design. After the CNN design and before the softmax, two completely linked layers are used for output. The neat setup of the VGG 16 is shown in Figure 3.

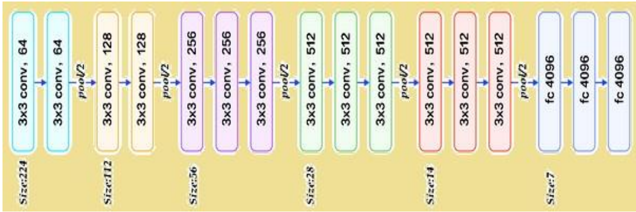

Fig. 3. Flow diagram of VGG16

3.6.2 Resnet -50:

ResNet's design, in contrast to AlexNet and VGGNet, is built on building-block modules (network-in-network architectures). Convolution, ReLU, and pooling are utilized as building blocks to build the network that results in the macro-architecture. This innovation reduced the model size by up to 102 MB by replacing entirely connected layers with global average pooling in ResNet. The 150 layers of this architecture stack on top of one another at regular intervals, and before the entire connected layer is utilized to output the classes, stride two is used to double the number of filters. One of the hyperparameters used in this approach is batch normalization. Each convolution layer experiences it. ResNet uses a 0.1 learning rate, a 1e-5 weight decay, and a 256-batch size. The absence of dropouts is another important aspect to emphasize.

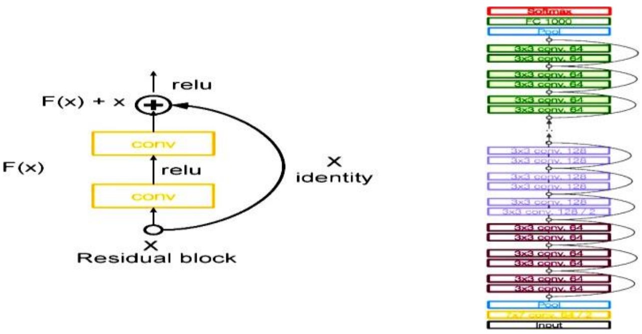

Fig. 4. Show Architecture of ResNet-50

The main idea behind the development of ResNet-50 is introducing an identity shortcut connection to skip the unnecessary layer as shown at right in the above figure 4. ResNet covers the drawbacks of the VGG module by skipping unnecessary layers. It reduces the computation cost and improves performance.

4. Results & Discussions

To get the best outcomes, the suggested technique is put into practice. After conducting a comparison study utilizing several models, the findings are gathered and described using the suggested technique with the use of graphs and tables. Datasets are divided into a training dataset and a testing dataset with a percentage of seventy percent and thirty percent, respectively, to accomplish results. Larger training datasets of higher quality provide greater performance in deep learning, according to the general rule of thumb. The above-mentioned dataset distribution is employed in the proposed work to prevent the model from being underfit owing to a limited training dataset. Here are some key terms that need to be understood to comprehend the outcome analysis. Figure 5 displays the cotton leaf disease training and validation loss using VGG-16. The results demonstrate that the model's training loss is continually reducing as the training goes on, and we halt the training when the loss is ideal.

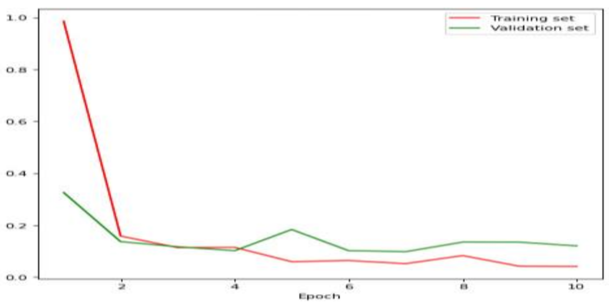

Fig. 5. VGG-16 Training and Validation loss accuracy

The primary model parameters were modified to 0.98, which was determined to be CNN's accuracy, after parameter optimization. The generalization ability was very strong, as shown by the F1 score of 0.9674, recall of 0.9648, and accuracy of 0.9563. Table 1 also displays the ROC curve and related AUC values. The 0.98 accuracy of our suggested technique, which is shown in Table 1, is evidence of its better performance.

Table 1. Result

Model	Feature	Recall	Precision	F1 Score	Accuracy
CNN	Resnet-50	96.48	95.63	96.74	98%

4.1 Accuracy of Model

The percentage of all occurrences that have been properly categorised may be used to measure accuracy. The percentage of all occurrences that have been properly categorized may be used to measure accuracy. Predictive model accuracy is crucial to accurately assessing forecasts, which is required to create scientific evidence for use in decision- and policy-making. Making differences between observed and predicted values is necessary for the quantitative measurement of accuracy in predictive modelling. Predicted values often refer to the values that were modelled based on previous data used for training. Prediction accuracy is based on the difference between anticipated and actual values for fresh samples. Figure 6 displays the cotton leaf disease detection model's accuracy on training and validation sets.

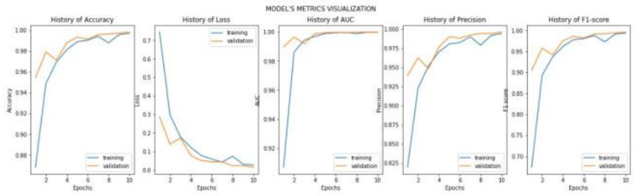

Fig. 6. Reliability of the training and validation set

4.3 Confusion Matrix

A popular tool for evaluating the effectiveness of classification models is a confusion matrix. The matrix (table) displays the number of cases that were correctly and incorrectly recognized based on the test data's actual results. When the model confuses two distinct classes, an assessment tool like a confusion matrix, for instance, provides a more in-depth analysis than a simple proportion of properly classified samples (accuracy), which might result in inaccurate results if the dataset is imbalanced (i.e., when there are huge differences in several between different classes). In this instance, there are n classes in the matrix. Binary classifiers have simply two categories: yes and no, positive and negative, and male and female. The predicted and observed events are cross-tabulated into four categories in a confusion matrix: The confusion matrix for the VGG-16 is shown in Figure 7.

- ❖ True Positive (TP): Predicting a label accurately.
- ❖ True Negative (TN): Predicting the other label with accuracy.
- ❖ False Positive (FP): Predicting a label incorrectly.
- ❖ False Negative (FN): Labels that are missing or incoming.

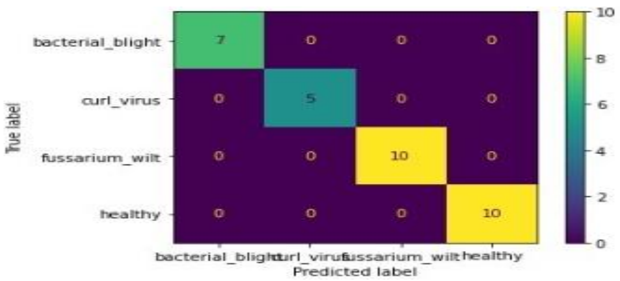

Fig. 7. Confusion Matrix of VGG16

4.3 Recognizing Cotton Leaf Disease

The experimental results show how the model predicts a set of test images. The results illustrate the model's functionality in figure 8. In photos of cotton leaves, the trained model recognizes the four classes.

Fig. 8. Model prediction using a series of test photos.

5. Conclusion and Future work

The acquired findings have led to the conclusion that deep learning VGG-16 has offered a superb answer to the picture classification issue. There is no need to manually extract the features for image classification in CNN since it automatically learns crucial picture aspects including form, texture, colour, size, and flaws that drastically minimize error. To extract unique features, CNN employs kernel matrices as local feature extractors. The model has to be tuned by adjusting the hyperparameter to identify a set of kernels from which to extract a useful discriminative feature. However, Deep Learning is the most effective method for classifying fruit images and determining their nutritional content. Using this method, results are 98% accurate. This research offers a methodical, non-destructive technique for treating cotton leaf disease. Farmers would benefit from detection to avoid loss. Farmers would benefit from the ability to detect cotton leaf disease to avoid loss. In the future, we may create mobile applications with end users who have been educated to identify cotton leaf disease. Farmers would benefit from detection to avoid loss.

Author contributions

Awad bin Naeem 1: Conceptualization, Methodology, Software, Writing-Reviewing. **Biswaranjan Senapati 2:** Data curation, Writing-Original draft preparation, Field study. **Alok Singh Chauhan 3:** Visualization, Investigation. **Sumit Kumar 4**: Writing-Reviewing and Editing. **Juan Carlos Orosco Gavilan 5**: Field study. **Wael M. F. Abdel-Rehim 6**: Software.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] J. Zhi, T. Qiu, X. Bai, M. Xia, Z. Chen, and J. Zhou, "Effects of Nitrogen Conservation Measures on the Nitrogen Uptake by Cotton Plants and Nitrogen Residual in Soil Profile in Extremely Arid Areas of Xinjiang, China," Processes, vol. 10, no. 2, p. 353, 2022. [Online]. Available: https://www.mdpi.com/2227-9717/10/2/353.

[2] T. Li, H. Zhao, S. Wang, C. Yang, and B. Huang, "Attack and Defense Strategy of Distribution Network Cyber-Physical System Considering EV Source-Charge Bidirectionality," Electronics, vol. 10, no. 23, p. 2973, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/23/2973.

[3] V. S. Dhaka et al., "A Survey of Deep Convolutional Neural Networks Applied for Prediction of Plant Leaf Diseases," Sensors, vol. 21, no. 14, p. 4749, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/14/4749.

[4] A. Arshad et al., "Impact of Climate Warming on Cotton Growth and Yields in China and Pakistan: A Regional Perspective," Agriculture, vol. 11, no. 2, p. 97, 2021. [Online]. Available: https://www.mdpi.com/2077-0472/11/2/97.

[5] S. S. Zaidi et al., "Molecular insight into cotton leaf curl geminivirus disease resistance in cultivated cotton (Gossypium hirsutum)," (in eng), Plant Biotechnol J, vol. 18, no. 3, pp. 691-706, Mar 2020, doi: 10.1111/pbi.13236.

[6] A. Rehman, A. A. Chandio, I. Hussain, and L. Jingdong, "Fertilizer consumption, water availability and credit distribution: Major factors affecting agricultural productivity in Pakistan," Journal of the Saudi Society of Agricultural Sciences, vol. 18, no. 3, pp. 269-274, 2019/07/01/ 2019, doi: https://doi.org/10.1016/j.jssas.2017.08.002.

[7] A. Iqbal et al., "Genotypic Variation in Cotton Genotypes for Phosphorus-Use Efficiency," Agronomy, vol. 9, no. 11, p. 689, 2019. [Online]. Available: https://www.mdpi.com/2073-4395/9/11/689.

[8] R. Koech and P. Langat, "Improving Irrigation Water Use Efficiency: A Review of Advances, Challenges and Opportunities in the Australian Context," Water, vol. 10, no. 12, p. 1771, 2018. [Online]. Available: https://www.mdpi.com/2073-4441/10/12/1771.

[9] Z. Chen and D. R. Gallie, "The ascorbic acid redox state controls guard cell signaling and stomatal movement," (in eng), Plant Cell, vol. 16, no. 5, pp. 1143-62, May 2004, doi: 10.1105/tpc.021584.

[10] D. R. Jones, "Plant Viruses Transmitted by Whiteflies," European Journal of Plant Pathology, vol. 109, no. 3, pp. 195-219, 2003/03/01 2003, doi: 10.1023/A:1022846630513.

[11] M. Akila and P. Deepan, "Detection and Classification of Plant Leaf Diseases by using Deep Learning Algorithm," International journal of engineering research and technology, vol. 6, 2018.

[12] R. W. Briddon, "Cotton leaf curl disease, a multicomponent begomovirus complex," Molecular Plant Pathology, vol. 4, no. 6, pp. 427-434, 2003, doi: https://doi.org/10.1046/j.1364-3703.2003.00188.x.

[13] M. N. Sattar, Z. Iqbal, M. N. Tahir, and S. Ullah, "The Prediction of a New CLCuD Epidemic in the Old World," (in eng), Front Microbiol, vol. 8, p. 631, 2017, doi: 10.3389/fmicb.2017.00631.

[14] J. Jenifer and D. J. B. Raja, "Skeleton Based Leaf Identification and Detection of Grape Plant Leaf Disease by Using K-Means Clustering Algorithm," 2017.

[15] A. Ahmad et al., "Engineered Disease Resistance in Cotton Using RNA-Interference to Knock down Cotton leaf curl Kokhran virus-Burewala and Cotton leaf curl Multan betasatellite Expression," Viruses, vol. 9, no. 9, p. 257, 2017. [Online]. Available: https://www.mdpi.com/1999-4915/9/9/257.

[16] K. L. Adams and J. F. Wendel, "Allele-specific, bidirectional silencing of an alcohol dehydrogenase gene in different organs of interspecific diploid cotton hybrids," (in eng), Genetics, vol. 171, no. 4, pp. 2139-42, Dec 2005, doi: 10.1534/genetics.105.047357.

[17] L. M. Abualigah, A. T. Khader, M. A. Al-Betar, and O. A. Alomari, "Text feature selection with a robust weight scheme and dynamic dimension reduction to text document clustering," Expert Systems with Applications, vol. 84, pp. 24-36, 2017/10/30/ 2017, doi: https://doi.org/10.1016/j.eswa.2017.05.002.

[18] V. S. Rana et al., "A Bemisia tabaci midgut protein interacts with begomoviruses and plays a role in virus transmission," (in eng), Cell Microbiol, vol. 18, no. 5, pp. 663-78, May 2016, doi: 10.1111/cmi.12538.

[19] I. Korkontzelos, A. Nikfarjam, M. Shardlow, A. Sarker, S. Ananiadou, and G. H. Gonzalez, "Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts," (in eng), J Biomed Inform, vol. 62, pp. 148-58, Aug 2016, doi: 10.1016/j.jbi.2016.06.007.

[20] P. R. Rothe and R. V. Kshirsagar, "Cotton leaf disease identification using pattern recognition techniques," 2015 International Conference on Pervasive Computing (ICPC), pp. 1-6, 2015.

[21] D. Mondal, D. K. Kole, A. Chakraborty, and D. D. Majumder, "Detection and classification technique of Yellow Vein Mosaic Virus disease in okra leaf images using leaf vein extraction and Naive Bayesian classifier," 2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI), pp. 166-171, 2015.

[22] A. H. Paterson et al., "Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres," (in eng), Nature, vol. 492, no. 7429, pp. 423-7, Dec 20 2012, doi: 10.1038/nature11798.

[23] M. Amlekar, A. T. Gaikwad, R. R. Manza, and P. L. Yannawar, "Leaf shape extraction for plant classification," 2015 International Conference on Pervasive Computing (ICPC), pp. 1-4, 2015.

[24] M. S. Shahid, S. Mansoor, and R. W. Briddon, "Complete nucleotide sequences of cotton leaf curl Rajasthan virus and its associated DNA β molecule infecting tomato," Archives of Virology, vol. 152, no. 11, pp. 2131-2134, 2007/10/01 2007, doi: 10.1007/s00705-007-1043-9.

[25] S. S. Sohrab, E. I. Azhar, M. A. Kamal, P. S. Bhattacharya, and D. Rana, "Genetic variability of Cotton leaf curl betasatellite in Northern India," Saudi Journal of Biological Sciences, vol. 21, no. 6, pp. 626-631, 2014/12/01/ 2014, doi: https://doi.org/10.1016/j.sjbs.2014.11.006.

[26] R. Guy, "Water-use efficiency and productivity trends in Australian irrigated cotton: a review," (in English), Crop & pasture science, vol. v. 64, no. no. 12, pp. pp. 1033-1048-2013 v.64 no.12, 0000 2013, doi: 10.1071/CP13315.

[27] Z. Jian and Z. Wei, "Support vector machine for recognition of cucumber leaf diseases," 2010 2nd International Conference on Advanced Computer Control, vol. 5, pp. 264-266, 2010.

[28] W. Nazeer, A. L. Tipu, S. Ahmad, K. Mahmood, A. Mahmood, and B. Zhou, "Evaluation of cotton leaf curl virus resistance in BC1, BC2, and BC3 progenies from an interspecific cross between Gossypium arboreum and Gossypium hirsutum," (in eng), PLoS One, vol. 9, no. 11, p. e111861, 2014, doi: 10.1371/journal.pone.0111861.

[29] W. Malik et al., "Molecular markers and cotton genetic improvement: current status and future prospects," (in eng), ScientificWorldJournal, vol. 2014, p. 607091, 2014, doi: 10.1155/2014/607091.

[30] S. Akhtar et al., "Regional Changes in the Sequence of Cotton Leaf Curl Multan Betasatellite," Viruses, vol. 6, no. 5, pp. 2186-2203, 2014. [Online]. Available: https://www.mdpi.com/1999-4915/6/5/2186.