Open Access
Full Length Article

ORIGINAL ARTICLE

A Novel-Based Deep Learning Approach for Early Diagnosis of Cotton Crop Disease

Muhammad Faisal Idrees[1], Mubshra Qadir[2], Mujahid Alam[3], Shahid Farid[4]

[1,4]Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan
[2]Department of Information Security, The Islamia University of Bahawalpur, Pakistan
[3]ICCC, Informatics Complex, H-8, Islamabad, Pakistan

ABSTRACT

Pakistan is among those countries where most of the GDP of a country depends on agriculture. As agriculture plays a vital role in the economic stability of Pakistan. The farmers are encountering various challenges in diagnosing diseases in cotton crops for the last decade. The attack of disease ultimately lessens the cotton yield and as a result, the economy of the country and the income of the farmers also go down. The detection of disease in crops at an early stage is crucial which can ultimately save farmers from loss. Manual detection is difficult and time-consuming an automated system to detect disease in plants with high accuracy will be handy for farmers. A deep learning approach is used in the proposed work to detect the disease in the crop plant. The Deep Convolutional Neural Networks model has been utilized to detect cotton plant disease using the image dataset of the cotton plant leaves. The model achieved a training accuracy of 99.08% validation accuracy of 100%. The proposed methodology will allow farmers to correctly detect diseases in the early stage. The diagnosis of plant disease helps to cure it in time. If the disease is controlled in time, its ultimate effect will be on the yield of the cotton crop. This will improve farmer income and the GDP of a country.

Keywords:

Author`s Contribution	Address of Correspondence	Article info.
[1,2] Data analysis, interpretation and manuscript writing, Active participation in data collection,[3] Conception, synthesis, planning of research. *[4.] Interpretation and discussion*	*Shahid Farid* *Email: shahidfarid@bzu.edu.pk*	*Received: August 16, 2022* *Accepted: December 07, 2022* *Published: December 30,2022*

Cite this article: *Idrees FM, Qadir M, Alam M, Farid S. A Novel Based Deep Learning Approach for early Diagnosis of Cotton Crop Disease. J. inf. commun. technol. robot. appl.2022; 13(1):1-07.*

Funding Source: *Nil*
Conflict of Interest: *Nil*

INTRODUCTION

The economy of Pakistan depends on agriculture. Farmers use traditional techniques for agriculture, which affects crop yield. The use of old methods results in low production due to disease attacks on crops. There are other reasons also for less yield of crops, among those crop disease is one of the major reasons for low crop yield. The best way to improve crop yield is by preventing and diagnosing the disease in its early stages. Most of the farmers are less educated and they use traditional ways to detect the disease which takes time to detect the disease correctly. The diagnosis of disease in a plant is very important as the early and correct detection of disease can save farmers from damage or loss. Timely detection of disease helps in curing the disease in its early stage. The correct detection of disease is also a big concern as there are similarities in leaf images of different diseases. The timely and correct detection of disease can be done with the help of precise agricultural knowledge and a fast responsive system.

With the revolution in Artificial Intelligence, there is remarkable progress is achieved in machine learning-based plant disease detection. In the last decade, a remarkable effort is done for plant disease detection using computer-based techniques. Much research has already been proposed for the diagnosis of disease in plants using computer vision and machine learning techniques. No doubt there are challenges to extracting and selecting the right features from image-based data. There may be noise in images due to the field environment. It takes a lot of effort to filter all these problems. Deep learning has an important role in the recognition of diseases in humans and plants. Convolutional Neural Network is one of the popular techniques to date. CNNs are built using the same principles as human neural networks, and they are capable to select the correct features to form the image data. The feature selection is performed automatically by CNNs instead of by manual selection. This has made a great impact on correct disease detection of plant disease. This technique gives remarkable results in the diagnosis of a single plant disease from image data but if multiple diseases are present on the image, it becomes challenging to handle. Hence, the objective of this study is to develop a mechanism that takes an input in the form of an image input from the user and diagnoses the cotton plant disease with high accuracy. The proposed CNN is trained on an image dataset of a cotton plant and then used for disease prediction with the testing dataset. The proposed work will ultimately increase the income of farmers and the economy of the country also gets positive effects.

LITERATURE REVIEW

Artificial intelligence-based techniques are popular in almost every field of life to solve various problems. The deep learning technique is proven to be the best one from others. It is used for image classification or disease detection. It very responsive and high accuracy generating approach. A deep learning approach has been utilized in[1] to detect a disease from cucumber leaf images. An augmented dataset of leaf images was used for training a deep convolutional neural network model. The result of their model was to generate high accuracy as compared to other comparative models. Due to the use of a large, augmented dataset, their model requires high computational and complexity costs. In[2] Nikhil et. al. proposed Artificial Neural Network based model. The ANN-based approach is deployed on an image dataset of cotton leaves. Before the training of ANN, the preprocessing of the image dataset is done. The cancelation of noise from images is done by the team to reduce the noise from the image dataset. The background removal enhances the image quality. The proposed technique detects the type of disease from the preprocessed enhanced image dataset. The proposed technique lacks accurate measurements.

A deep Convolutional Neural Network is used to detect lesion disease in cotton leaves Rafael et. al.[3] They proposed a DCNN-based approach to observe the cotton plants' health and assist in the detection of lesion disease in cotton crops. This helps to detect the disease in the early stage which results into maintain better health of the cotton crop. The outcome of the DCNN is matched with some traditionally used techniques of Machine learning algorithms. The proposed mechanism performs better as compared to its counterpart. A deep learning model has been deployed by[4] on the internet of things base device. Their work is based on disease detection from cotton plants. The IoT device takes an image of a crop plant and generates results on whether the plant is infected by disease or not. The supervised learning approach is used for training of DCNN model. The image dataset consists of healthy and infected cotton leaf images. The IoT-based devices give better results up to a possible extent. Their technique also urges researchers to work on it and produce more precise and better results in the future.

Jenifa et. al.[5] implements the DCNN model to detect a cotton disease from an image dataset. The dataset is based on multiple disease images of cotton crop leaves. It was claimed that a deep convolutional neural network-based approach has been deployed for the multi-classification of cotton disease that was never used before in this context. The proposed model identified the cotton disease automatically. Their experimentation results are remarkable. The model gives high accuracy. The image dataset is small and based on 600 images of both nutritious and diseased leaves. An effort was made in[6] utilizing Deep Convolutional Neural Network to reveal the disease of different guava plant species. They used the color histogram technique. Preprocessing and augmentation of the image dataset are also done to make the dataset larger. A transfer learning technique is used in the proposed work five different pre-trained models are used and one model gives the highest accuracy. Liu et. al.[7] used a deep convolutional neural network to identify apple plant disease from apple leaf images. This proposed model improved the disease detection accuracy and also strengthen the CNN model's stability. The dataset is big enough and a pre-trained model is used to get results. There is multi-disease detection is performed on apple leaf images. Arnab et. al.[8] deployed a shallow CNN approach to detect disease in different plants. An image dataset is used and consists of leaf images of different plants. The real-time testing dataset is used and the model gives exceptional accuracy measures. Navina et. al.[9] used deep learning to identify diseases from cotton leaf images. A real-time dataset is used and preprocessing is performed to get better results. Features are extracted on basis of the color and texture of the image. They proposed an application-based system that takes an image as input and generates results as an infected or healthy leaf.

Meta Deep Leaning approach is used by Suleman et. al.[10] for the detection of disease in cotton crops. The model is proposed to identify the disease from the input leaf image of cotton. They aim to detect multiple diseases using the meta DCNN. High accuracy is achieved by the model. A machine learning approach is also used for cotton crop disease detection. Pechuho et. al.[11] used machine learning to detect cotton disease. Arnal[12] used deep learning to detect diseases from different plants. Lesions and spots are taken as a feature from the image dataset of different plants and each plant has multiple disease leaf images in the data. The model gave high accuracy but the dataset was not enough to cover the complete area of plant disease but it can be achieved by making the dataset larger. In[13] Azmath et. al. used a deep learning approach to detect disease in cotton and also worked on the diagnosis of pests in cotton plants. Preprocessing of the image dataset is focused to improve the classifier results. K-fold cross-validation is used for dataset splitting. The boosted generalized convolutional neural network model is used for classification. The model gives a 96.4% accuracy score in the classification of cotton diseases and pests in cotton. Shriya et. al.[14] implemented a deep-learning approach for disease prediction in plants. The author used a deep CNN model to predict disease in the leaves of 13 different plants taken from the plant-village leaf disease dataset. The performance of the model is also validated on a cotton leaf dataset taken in real-time. Md. Ashiqul et. al.[15] proposed a deep-learning-based CNN model which takes a paddy leaf image as input and classifies whether that leaf is diseased or healthy. The proposed work is for the multi-classification of disease in paddy plant leaves. A novel-based deep-learning approach is proposed for disease identification in rice by Santosh et. al.[16] The proposed work is to predict three rice diseases from the rice plant image dataset that comprises more than 4000 images of both healthy and diseased rice pant leaves each. The author claims that prediction accuracy is higher than in the previous proposed work. Various approaches have been identified from the state of the literature for the detection of cotton crop disease. Besides all the cited work, this study contributed in a way to propose a sequential model based on a deep convolutional neural network and give higher accuracy for disease detection in cotton crops which is never been done in this context.

RESEARCH METHODOLOGY

Automated disease detection in plants is replacing the traditional methods of disease detection which are based on manual observation. The researchers proposed automated systems centered on both deep and machine learning approaches. They utilized various datasets some are publicly available, and some have used real-time datasets. Deep learning-based systems performed better than machine learning-based systems. In this proposed work we are using deep learning sequential model designed by using scratch. The block diagram of the suggested methodology is illustrated in Fig. 1 whereas, Fig. 2 represents the flow diagram of the intended architecture respectively.

Image Dataset:

The dataset is taken from Kaggle which is an open-source cotton disease dataset available publicly. The dataset consists of 2310 images of the cotton plant and leaf of the cotton plant. The dataset is divided into the train, test, and validation datasets.

Table 2. Training Dataset Regeneration								
Parameter	Rescale	Rotation Range	Width Shift Range	Height Shift Range	Shear Range	Zoom Range	Horizontal Flip	Fill Mode
Value	1.0/255	45	0.25	0.25	0.25	0.25	True	Nearest

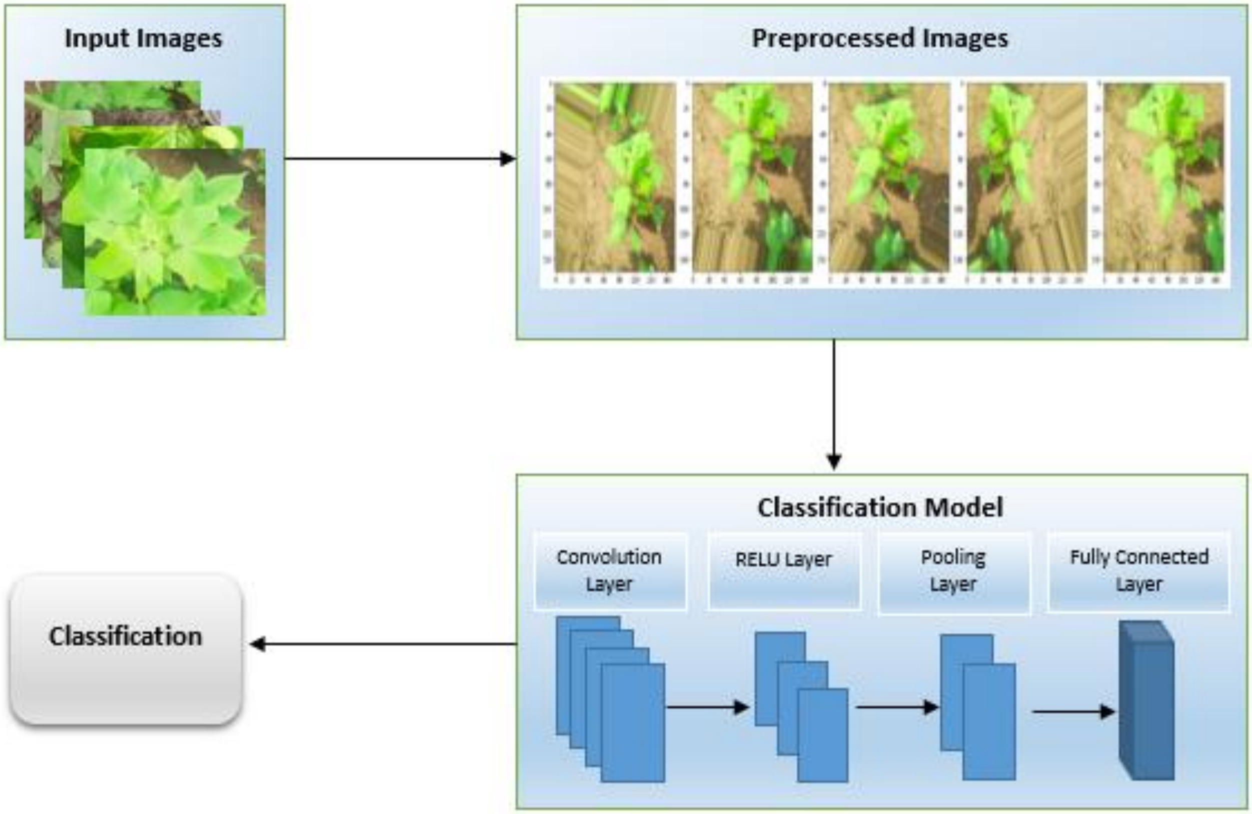

Figure 1. Block diagram of the research methodology

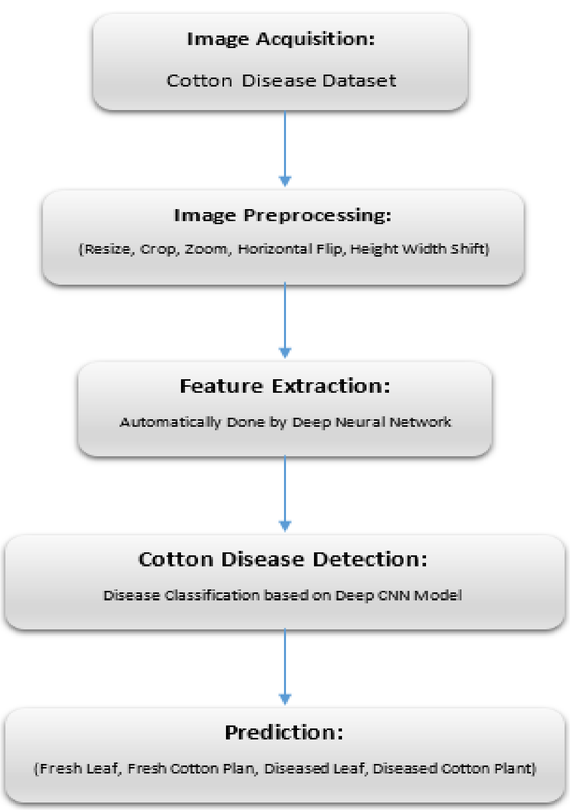

Figure 2. Flow Diagram of Proposed Architecture

Then each dataset consists of four classes named as disease cotton plant, disease cotton leaf, fresh cotton plant, and fresh cotton leaf.

Each class consists of relevant images of the cotton plant. The number of images in each class is illustrated in Table 1.

Table 1. Number of Images of Each Class in Dataset

Cotton Disease Class	Train Dataset	Test Dataset	Validation Dataset
Disease Cotton Plant	815	28	78
Disease Cotton Leaf	288	25	43
Fresh Cotton Plant	421	27	66
Fresh Cotton Leaf	427	26	66

Preprocessing of Dataset

The image dataset consists of images of different sizes. The dataset is preprocessed before using it for training and testing of the CNN model. An image regenerator is used for preprocessing of the dataset. The following parameter values (as depicted in Table 2) are implemented on the image dataset. Train and Test data sets are preprocessed before using them for model training and testing.

The image dataset is resized to 128 x 128 targeted size. Batch size is taken as 32 and class mode is taken as categorical. The dataset is normalized and labeled. The testing and validation dataset is only rescaled to 1.0/255.

CNN Model

The convolutional neural network consists of at least one convolutional layer in the input layer of the model. In proposed word model is a sequential model and it consists of five convolutional layers followed by a max pooling layer after each convolutional layer in the model. The mathematical equation for convolutional operation is given as

$$Conv_i^l = Bias_i^1 + \sum_{i=1}^{l-1} w_{i,j}^{(l-1)} * C_i^l$$

In the above equation $Conv_i^l$ is the output of the convolutional operation and $Bias_i^1$ is the bias matrix with the ith iterative region of operation , w is convolved window and i, j is number of rows or column in w . C_i^l is multiplied with pixel of image where area is defined by convolved window.

Pooling is used in a convolutional neural network to calculate min, max, and average pools to get a single output value from given kernels. The formula for pooling is as follows

$$D_{output} = x_h * x_w + x_d$$

In the above equation D_{output} is the output obtained after pooling, where x is the input and h, w are the height and width respectively and d is represents the color channels.

ReLU is rectified activation unit and it gives an output value of 1 if the input value is greater than 0, otherwise, it gives 0 as output. ReLU is calculated as follows.

$$ReLU = max(0, x)$$

$ReLU$ is representing the output of the activation unit and x is the input and if the input is greater than 0 then the output will be max otherwise $ReLU$ gives output 0.

The Softmax function is used to get probability between 0-1. It working is like multilinear regression where multiple class prediction is done on basis of internal value.

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}$$

In the above equation $\vec{z}$, is the input vector and softmax calculation is done for that input vector. e is an exponential function and it is used to generate values greater than zero or in simple it will give positive values.

RESULT AND DISCUSSION

The proposed work is to automate the disease detection process in cotton plants. The proposed DCNN sequential model is trained on the cotton disease dataset. The dataset is preprocessed before being used it for the training of the model. The "model.fit" is used to train the model. Parameters for model training are given in Table 3.

In machine learning and deep learning model, the accuracy of the model is calculated to identify the performance of model for classification problem. The equation for accuracy measure is given as

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

The accuracy equation is based on ratio of sum of TP and TN over sum ofTP, TN, FP, FN. The loss of proposed model is measured the knowing the number of false prediction by the proposed model. The following equation is used to measure the error rate or miss classification done by the classification model.

$$Loss = \frac{(FP + FN)}{(TP + TN + FP + FN)}$$

The loss equation is based on ratio of sum of TF and FN over sum of TP, TN, FP, FN.

The model achieves training accuracy of 99% and testing accuracy of 99.06%. Their graphical representation is in given below Fig. 3.

Table 3. Model Training Parameters

Parameters	Optimizer	Learning Rate	Loss	Metrics	Epochs	Verbose
Values	Adam	0.0001	Categorical Crossentropy	Accuracy	300	1

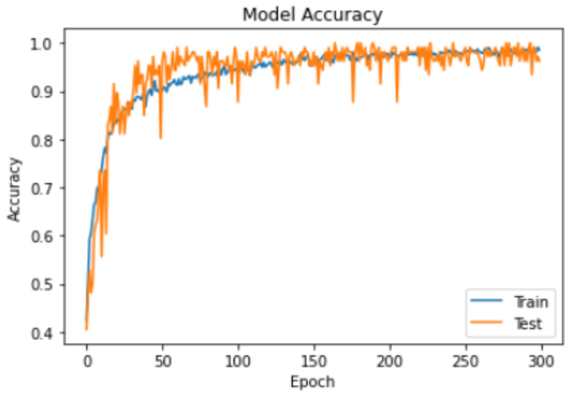

Figure 3. Proposed Model Accuracy

The model gives training loss 0.0341 of and testing loss of 0.0353 and the graphical representation is illustrated in Fig. 4.

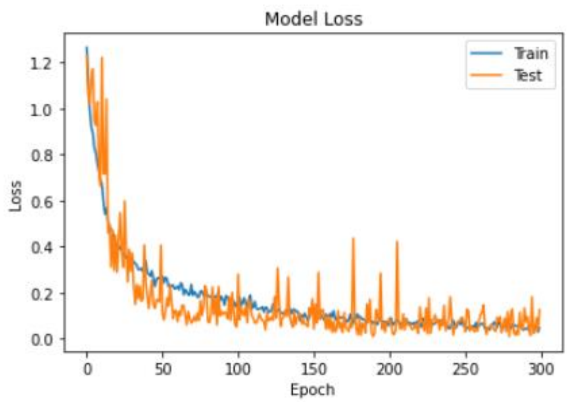

Figure 4. Proposed Model Loss.

The proposed model performed better as compared to previous models. The model achieves excellent results given in Table 4.

Table 4. Model Performance

Parameters	Training Accuracy	Training Loss	Validation Accuracy	Testing Loss
Values	0.9908	0.0271	1.0000	0.0078

The proposed model is validated on validation data set and it performed with high accuracy and give 100% accuracy on validation dataset image. The model is tested with different image of validation dataset and proposed model predicted correctly. Results are in Fig. 5.

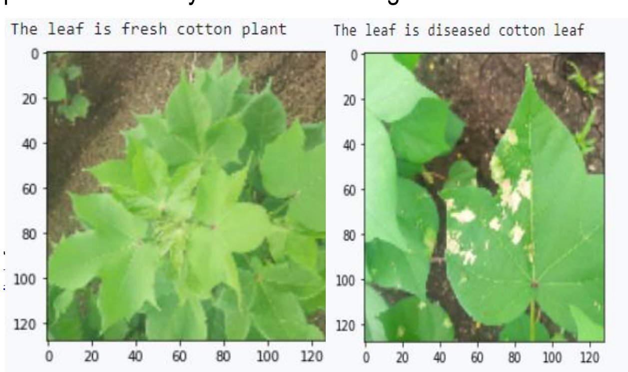

Figure. 5 Prediction Result of the proposed model

CONCLUSION

For most of the world's nations, agriculture is a significant industry that has a substantial impact on their GDP. Pakistan is among the list of those countries where a large part of the economy relies on agriculture. Cotton is considered one of the most significant and high-yielding crops in agriculture. In the modern world, there are multiple new diseases in the cotton crop and its correct detection in the early stage is difficult. The proposed model is aimed to give a mechanism that will automatically detect cotton crop disease. A cotton crop disease dataset based on images has been utilized in this study. The dataset comprises four classes named fresh cotton plant, fresh cotton plant leaf, diseased cotton plant and disease cotton plant leaf. The dataset has been categorized into three categories including train dataset, test dataset, validation dataset. The dataset is preprocessed before use it for training and testing of model. The image dataset is deployed to train the model after the dataset has been preprocessed. Sequential deep CNN model is designed in proposed work. The deep learning model perform automatic feature selection form the dataset. The model give training and testing accuracy of 99.08% and 100% respectively. Minimum loss for model training and testing is 0.0271 and 0.0078 respectively. The trained model gives 100% validation accuracy on validation dataset and perform correct prediction. In future other plants and crops shall be taken into consideration for the detection of diseases. Moreover, the model will be tested with dataset taken on real time. A user-friendly environment can also be developed using this model which can be more helpful to farmers. It can be effective while diagnosing disease in rice crop or wheat crop.

REFERENCES

1. Ma, J., Du, K., Zheng, F., Zhang, L., Gong, Z., & Sun, Z. (2018). A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. *Computers and electronics in agriculture*, *154*, 18-24. https://doi.org/10.1016/j.compag.2018.08.048
2. Shah, N., & Jain, S. (2019, February). Detection of disease in cotton leaf using artificial neural network. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 473-476). IEEE.https://doi.org/10.1109/AICAI.2019.8701311
3. Caldeira, R. F., Santiago, W. E., & Teruel, B. (2021). Identification of cotton leaf lesions using deep learning techniques. *Sensors*, *21*(9), 3169.https://doi.org/10.3390/s21093169
4. Patil BV, Patil PS. Computational method for Cotton Plant Patil BV, Patil PS. Computational method for Cotton Plant disease detection of crop management using deep learning and internet of things platforms. InEvolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2020 2021 (pp. 875-885). Springer Singapore.. https://doi.org/10.1007/978-981-15-5258-8_81
5. Jenifa, A., Ramalakshmi, R., & Ramachandran, V. (2019, December). Cotton leaf disease classification using deep convolution neural network for sustainable cotton production. In *2019 IEEE international conference on clean energy and energy efficient electronics circuit for sustainable development (INCCES)* (pp. 1-3). IEEE.3).https://doi.org/10.1109/INCCES47820.2019.9167715
6. Mostafa, A. M., Kumar, S. A., Meraj, T., Rauf, H. T., Alnuaim, A. A., & Alkhayyal, M. A. (2021). Guava disease detection using deep convolutional neural networks: A case study of guava plants. *Applied Sciences*, *12*(1), 239.. https://doi.org/10.3390/app12010239
7. Liu, P., Mahmood, T., & Khan, Q. (2017). Multi-attribute decision-making based on prioritized aggregation operator under hesitant intuitionistic fuzzy linguistic environment. *Symmetry*, *9*(11), 270..https://doi.org/10.3390/sym10010011
8. Wang, L., Zhao, J., & Mortier, R. (2022). *OCaml Scientific Computing*. Springer Cham.https://doi.org/10.1007/978-3-030-97645-3_11
9. Johri P, Parashar N, Saxena U, Pushpanjali K. Plant Disease Detection Using Deep Learning Model. In2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON) 2022;1:456-462. https://doi.org/10.1109/COM-IT-CON54601.2022.9850667
10. Memon, M. S., Kumar, P., & Iqbal, R. (2022). Meta Deep Learn Leaf Disease Identification Model for Cotton Crop. *Computers*, *11*(7), 102. https://doi.org/10.3390/computers11070102
11. Pechuho, N., Khan, Q., & Kalwar, S. (2020). Cotton Crop Disease Detection using Machine Learning via Tensorflow. *Pakistan Journal of Engineering and Technology*, *3*(2), 126-130.. https://theskyjournal.net/journals/pakjet/article/view/417
12. Barbedo JG. Plant disease identification from individual lesions and spots using deep learning. Biosystems Engineering. 2019 Apr 1;180:96-107.https://doi.org/10.1016/j.biosystemseng.2019.02.002
13. Zekiwos, M., & Bruck, A. (2021). Deep learning-based image processing for cotton leaf disease and pest diagnosis. *Journal of Electrical and Computer Engineering*, *2021*, 1-10.0.https://doi.org/10.1155/2021/9981437
14. Jadhav, S., & Lal, A. M. (2022). Multi-Class Plant Leaf Disease Detection Using a Deep Convolutional Neural Network. *International Journal of Information System Modeling and Design (IJISMD)*, *13*(1), 1-14. https://doi.org/10.4018/IJISMD.315126
15. Islam, M. A., Shuvo, M. N. R., Shamsojjaman, M., Hasan, S., Hossain, M. S., & Khatun, T. (2021). An automated convolutional neural network based approach for paddy leaf disease detection. *International Journal of Advanced Computer Science and Applications*, *12*(1). https://doi.org/10.14569/IJACSA.2021.0120134
16. Upadhyay, S. K., & Kumar, A. (2021). A novel approach for rice plant diseases classification with deep convolutional neural network. *International Journal of Information Technology*, 1-15.. https://doi.org/10.1007/s41870-021-00817-5