Classification of Cotton Leaf Diseases Using Transfer Learning-DenseNet-121

B. Arathi and Uma N. Dulhare

Abstract Farmers growing cotton will have a great help if the cotton yield is predicted accurately and helps in making decisions such as crop insurance, how much to store, investments, requirement of fertilizers, water, etc. Generally, yield is measured by means of sample surveys using destructive sampling of cotton fields and will take enormous time, cost for the labor is high. As we know that these cotton plants are affected by various bacterial and fungal diseases based on the climate conditions resulting in the decline of cotton productivity. Plants are prone to numerous diseases. In the cotton plants, the mostly affected part is the leaf that damages the plant resulting in the damage of the entire crop. In order to detect the diseases of the cotton leaf, image processing and machine learning techniques are employed. In the existing work DEEP learning technique, CNN is employed for feature extraction which is used to detect plant diseases. There is an issue with the accuracy of these traditional CNN algorithms. The Experimental results achieved showed that the proposed model i.e., DenseNet-121 pre-trained Model is capable of classifying different leaf images in the dataset with higher classification accuracy of 91%. This transfer learning technique uses ImageNet weights to detect the diseases of cotton plant accurately. The abstract should summarize the contents of the paper in short terms, i.e. 150–250 words.

Keywords Cotton leaf disease · DenseNet-121 · Convolutional neural network

B. Arathi (✉)
Computer Science and Engineering, Kamala Institute of Technology and Science, Huzurabad, Karimnagar, Telangana, India
e-mail: artibairi@kitssingapuram.ac.in

U. N. Dulhare
Computer Science and Artificial Intelligence, Muffakham Jah College of Engineering and Technology, Hyderabad, Telangana, India

A. B. Reddy et al. (eds.), *Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems*, Lecture Notes in Networks and Systems 612, https://doi.org/10.1007/978-981-19-9228-5_33

1 Introduction

Among all the fiber crops cotton accounts for 35% of the global total fiber and can be used in the production of biofuel, staple, and as a raw material for manufacturing clothes (Satraj Sohrab et al. 2014). India's economy is affected by its yield (Ashourloo et al. 2016). It is the main crop that is available economically but it is affected by various pathogens that limits its production. Our country ranks first in the cotton yield and there is a necessity to detect the diseases priorily to still further enhance the productivity (Dulhare and Gouse 2022).

According to yield, India tops rank one in area wise and rank three in the production. For India, Cotton is very vital in its economy and in its industrial activities (Bashish et al. 2011).

Temperature and humidity are vital factors which strengthens the yield of the crop and the root exudates of it improve the nutrients of the soil (Tijare et al. 2015). The plant growth and yield are affected by the plant illness and have an influence on agriculture both socio-biological and financial aspects (Zheng et al. 2019).

At various stages of plant growth, plant infections influence its development, collection, acquire, and analyze of cotton diseases manually may not be efficient and identifying the diseases becomes not that easy (Lv et al. 2012). These diseases occur because of various microorganisms, nematodes, and other agents.

Keeping all this in mind, effective methods for the management of diseases is necessary. Earlier work shows that these diseases will have an adverse affect on the country's economy.

Various cotton leaf diseases, symptoms and methods. Imbalances are caused in cotton plants because of low nutrients, environmental stress, and chemical factors which leads to low productivity of cotton. Diseases like verticillium wilt, cotton leaf curl limits the growth of cotton (Kalbande and Pati 2016) (Table 1).

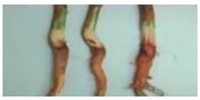
2 Literature Survey

Plant diseases can be detected and classified by various algorithms and accordingly research is taking place in this area and few of them are working on the classification of cotton diseases and increase the yield of cotton. Following are the work carried out by various researchers in this area.

Rothe and Kshirsagar (2014) made use of picture preparing strategies to find the diseases and extracted the main portions of the diseased leaf images by using image enhancement and segmentation.

Li et al. (2011) developed an image processing procedure alongside edge detection that takes computerized pictures and different strategies are applied to remove RGB highlights which helps in identifying the diseased part.

Table 1 Various cotton leaf diseases, symptoms and methods

Disease Name	Image	Symptoms	Suitable methods
Angular leaf spot or black arm disease		• Small spots appear below the cotyledons, which may also dry and fall off • Elongated dark brown lesions appears on the steam and branches	Principal component analysis K-nearest neighbor (KNN)
Vascular Wilt disease		• The yellowing of cotyledons sauteing of petioles and filling of dried leaves occur at the seedling stage • In the young and grown-up plants deficiency of bloat, leaf hanging and demise of the plants takes place	Edge detection
Grey Mildew		• Nonuniform rakish, pale and lackluster spots which leads the leaf color to yellowish brown	Decision tree algorithm and K-nearest neighbor
Anthracn ose disease		• Seedling stage minor, reddish round spots appear on the cotyledons and primary leaves • Bolls are circular, slightly sunken reddish-brown spots turns to black	Fuzzy feature and decision tree
Root Rot disease		• Brown sports in color on the cotyledons • At the collar area, there is dark shading, extends towards the lower parts • Rotting and shredding appear at the bark of the roots	Neural network
Boll Rot disease		• Small spots disease's which are brown or black in color which expands encloses the completed bolls	Neural network

(continued)

Table 1 (continued)

Disease Name	Image	Symptoms	Suitable methods
Leaf spot or blight disease		• Irregular round spots with tiny pale brown exits on the leaves so several spots together form blighted areas	Image processing and support vector machine

Brugger et al. (2011) employed the method of Agricultural Growers Resource Organization to find the disease of the crop in the initial stages to bridge a gap between farmers and agricultural expert.

Hannunal et al. (2011) worked on quick and exact recognition of cotton diseases using K means and KNN with a detection accuracy of 94%.

Kranthi et al. (2002) worked on cotton leaf diseases and predicted them using machine learning methods.

Wani and Ashtankar (2017) used the random forests algorithm with decision trees using MLP classifier to characterize and found from their work that multi yield regressor performed well by using a random forest system.

Rothe and Kshirsagar (2014) used preprocessing image technology to smooth the images in the detection of cotton plant diseases such as Myrothecium, Alternaria, etc.

Gulhane and Gurjar (2011) worked on the identification of cotton plants where in their leaf's are affected with diseases known as Leaf spot with the help of Support vector machine classification. It uses GLCM technique for disease detection and HSV algorithm in the identification of leaf disease.

Prashar et al. (2019) presented a paper to recognize the leaf diseases by using multilayered perceptron with overlapping pooling to find the contaminated leaves. KNN and SVM are used for morphological division, design coordinating restricting the diseased area with 96% accuracy.

Li et al. (2016) implemented CNN for segmenting the cotton bolls using cotton boll segmentation algorithm.

Images can be analyzed by using semantic segmentation. Xu et al. (2021) implemented the tool ENVI deep learning module i.e. U-Net Network model that makes use of Tensor Flow to train the model and identifies the feature in the image according to the spatial dimensions. The UAV remote sensing data images is given as input to the model.

Bargoti and Underwood (2017) and Rahnemoonfar and Sheppard (2017) implemented CNN for the extraction of the given images and finds applications in agriculture for detecting the objects.

A CNN based models needs labeled data as a training dataset, but it takes a longer time for processing if the set is bigger and sometimes the model misses the cotton bolls even they are present in the images. Therefore, its challenging for CNN to avoid misclassification.

Liu et al. (2016) used the CNN, a deep learning technique for recognizing the flower species effectively.

Dulhare et al. (2019) proposed Image fusion to produce a very high-resolution muti spectral picture by adjoining two or more images which helps in reducing the redundancy of the output.

Pérez-Zavala et al. (2018) discussed in their work regarding the detection of grape bunch using Image-based methods. Yin et al. (2018) proposed plant leaf segmentation and tracking and Stein et al. (2016) discussed mango fruit detection.

3 Methodology

The methodology for identifying the leaf diseases with its classification methods are presented in Fig. 1. The work discusses on the cotton plant leaf diseases and their classification to detect the contaminated leaves accurately.

Image Segmentation and Feature Extraction: In this classification, identification of diseases utilizing picture preparing strategies which extracts the important aspects like region, major, and minor axes, direction, etc. from the images of diseased leaves. For this purpose, two strategies image enhancement and image segmentation are used.

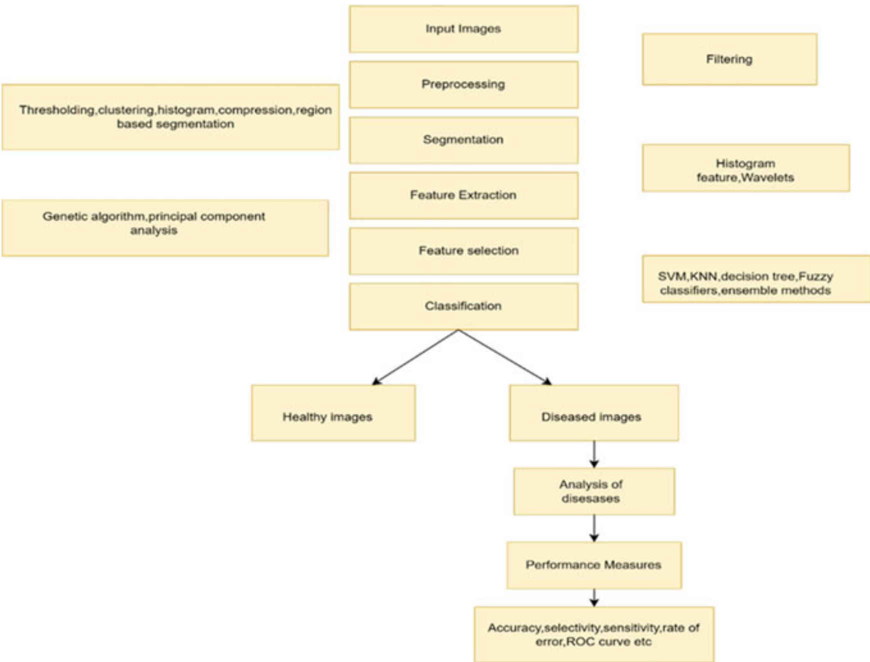


Fig. 1 Structure of cotton leaf diseases detection

The proposed method uses the CNN image classification technique to classify the leaf picture. Based on retrieved information at each convolution layer, it can detect and recognize diseases automatically. For disease identification, the system uses the image processing techniques. The image of a cotton plant leaf should be uploaded first. The algorithm will pre-process the image before the CNN algorithm is applied. The network is built using a combination of pretrained on ImageNet and the DenseNet-121 component, and it performs the other state-of-the-art techniques. Every convolution layer within dense block is tiny, so each convolution kernel is still in charge of learning the tiniest details.

From many years, humans have grown plants for various domestic needs, and the monitoring of diseases on these plants is very important to avoid shortage of food and raw material for various other needs.

The following Fig. 1 shows general structure of cotton leaf disease detection and list the various detection analysis methods.

There are various ways to detect cotton leaf disease detection:

Step 1: Leaf images are given as input for pre-processing and using filtering techniques, the noises in the images are eliminated.
Step 2: The filtered images are segmented using thresholding, clustering, histogram, compression, region-based segmentation.
Step 3: In the segmented images, the features are extracted as histogram feature, wavelets.
Step 4: The extracted features are selected using genetic algorithm and principal component analysis.
Step 5: The selected features are then classified using various classification techniques like K-nearest neighbor Support Vector Machine, decision tree, fuzzy classifiers and ensemble methods.
Step 6: The classification techniques help us in segregating healthy images and diseases images.
Step 7: The diseased images are analyzed for the presence of diseases using Support vector machine, KNN, Principal Component Analysis, Ensemble Method (Random Forest) and Decision Tree, Neural Network.
Step 8: After analyzing the diseases the various performance measures are determined to find the accuracy of the selected Transfer Learning technique.

4 Proposed Model

DenseNet-121 was employed to minimize error and maximize the performance. DenseNet-121 is proposed as one of the best performers in image classification of popular datasets such as CIFAR-10, ImageNet, etc. DenseNet-121 uses a simple pattern for connecting layers to each other directly in a feed-forward pattern, that is each layer consists of several additional inputs from previous layers and transmits its own feature maps to subsequent layers.

Fig. 2 Healthy and Unhealthy cotton plant leaves

In order to perform classification, we used the New Plant village dataset which is a hugely popular dataset available from Kaggle are downloaded which is a crowd-sourced and open-source platform in Lu et al. (2017) and Kaur et al. (2019) where they get trained, challenged to solve numerous problems. The New Plant village dataset consists of 1951 images with both healthy and diseased plant leaves in Fig. 2 Adedoja et al. (2022). The images are then split in the ratio as train and test sets. Every image is resized into 100 × 100 pixels to perform model predictions on these images. To split the data into training, testing and validation 40% data goes to testing and remaining 60% goes to training.

Table 2 lists the various parameters that are used in training DenseNet-121 from the dataset of plant village. Initially, the dataset images are loaded with the help of augmented image data store. Then, pretrained DenseNet-121 network is loaded, and the graphs from the trained network are extracted. Finally with the help of augmented training data, Dulhare and Ali (2021).The network is trained, and by validation data the classification accuracy is found out.

DenseNet-121 which is a pre- trained model which consists of 121 layers by a process known as Transfer Learning, which is re-trained by the DenseNet121 model by augmenting images, resized them to 100 × 100 size. The results were much more promising. Our main objective is to find out the higher classification accuracy.

From Tables 3 and 4 represents Transfer Learning DenseNet-121 model is trained over 20 epochs with each batch size of 32 so each epoch tells us the number of times model will be trained and achieved high Training and validation accuracy of 91% is achieved at 20 epochs. After training, the dataset consists of 781 validation images of healthy and unhealthy plant leaves.

Table 2 DenseNet-121 training validation parameters

Parameters	Value
Parameters	Value
Input layer size	100 × 100
Epochs	20
learning rate	0.001
Image augmentation	Image flipping, translations, rotations, shearing, zooming
Hardware resources	Single GPU
Batch size	32

Table 3 Performance comparison of training/validation loss for the dataset used

Training	Training Loss	Validation loss
1	1.3012	0.4892
2	0.5020	0.4098
3	0.4144	0.3534
4	0.3710	0.3762
5	0.4028	0.3717
6	0.3467	0.3108
7	0.3264	0.2777
8	0.3086	0.3112
9	0.3256	0.4516
10	0.2880	0.2641
11	0.2734	0.2673
12	0.2876	0.2823
13	0.2880	0.3221
14	0.2778	0.3294
15	0.2900	0.3515
16	0.2542	0.2750
17	0.2802	0.2369
18	0.2162	0.3127
19	0.2292	0.3209
20	0.2246	0.2629

5 Results and Discussion

The matrix shown below is the confusion matrix between predictions and test images. The matrix illustrates the four classes. The confusion matrix is used to make the summary of the prediction in a graphical way. It compares actual test images with predictions made by the model after training and running the test images through

Table 4 Performance comparison of training/validation accuracy for the dataset used

Training	Training accuracy	Validation accuracy
1	0.5940	0.8041
2	0.8051	0.8464
3	0.8453	0.8784
4	0.8701	0.8528
5	0.8479	0.8604
6	0.8632	0.8860
7	0.8769	0.9052
8	0.8795	0.8848
9	0.8744	0.8284
10	0.8872	0.9052
11	0.8974	0.9078
12	0.8863	0.9078
13	0.8957	0.8860
14	0.8966	0.8784
15	0.8855	0.8758
16	0.8991	0.9065
17	0.8991	0.9193
18	0.9265	0.8848
19	0.9103	0.8796
20	0.9060	0.9065

the DenseNet-121. Rows represent the Predicted labels, while columns represent the actual True labels.

In the below Fig. 3, Confusion Matrix can analyze the model as four class labels namely Class '0'-unhealthy cotton leaf, Class '1'-unhealthy cotton plant, Class '2'-healthy cotton leaf, Class '3'-healthy cotton plant.

From the Table 5, it is evident that model predictions are almost accurate. The experimental results achieved showed that the proposed model is capable of classifying healthy and unhealthy plant leaf images contains 781 validation images in

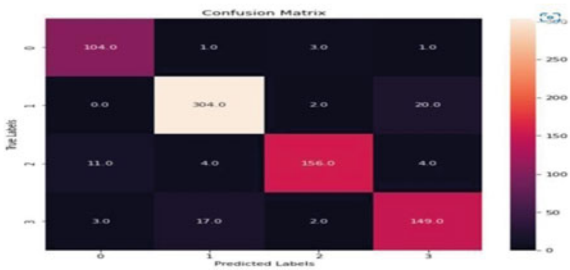


Fig. 3 DenseNet-121 Confusion Matrix

the dataset with higher classification accuracy 91% using DensNet-121 pretrained model with four classes '0', '1', '2', '3'. From Fig. 3 Transfer learning DenseNet-121 Model Layer is trained and represented Figs. 4 and 5.

In Fig. 4, the graph illustrates the plot between loss and validation loss of DenseNet-121 model training. In the initial iteration 20 epochs were used, the loss is found to be very high around 1.3, validation loss is 0.48, during training, after each epoch, both loss and validation loss comes down. During the 5th epoch, there was a slight increase in Val loss, but it comes down again in the next epoch, as the training still continued.

In Fig. 5, the graph represents a validation accuracy of 91% obtained over 20 epochs of training, while a training accuracy of 89% was reported. This is an effective measure of the classification made by Transfer Learning-DenseNet-121 model illustrates the plot between Training and Validation accuracy of DenseNet-121 model training. In the initial iteration 20 epochs were used and the accuracy is found to be very low around 0.5, validation accuracy is 0.8, during training, after each epoch,

Table 5 Classification results with various metrics

Classes	Precision	Recall	F_1-score	Support
Class 0 Unhealthy cotton leaf	0.88	0.95	0.92	109
Class 1 Unhealthy cotton plant	0.93	0.93	0.93	326
Class 2 Healthy cotton leaf	0.96	0.89	0.92	175
Class 3 Healthy cotton plant	0.86	0.87	0.86	171
Macro avg	0.91	0.91	0.91	781
Weighted accuracy	0.91	0.91	0.91	781
Accuracy			0.91	781

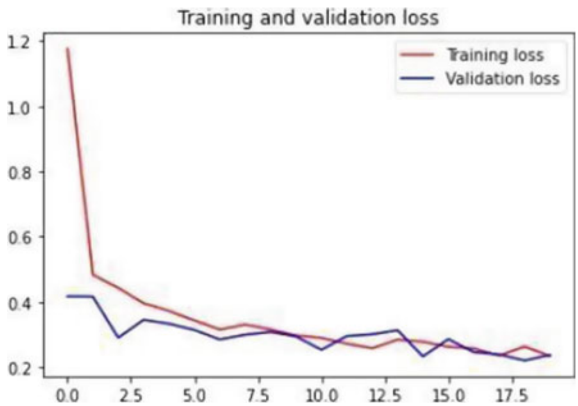


Fig. 4 Graphical representation of training and validation loss of the model

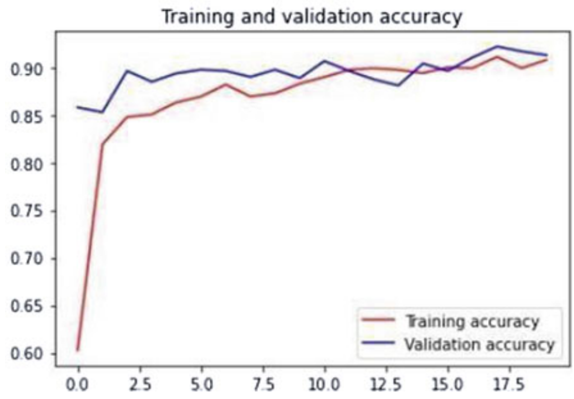


Fig. 5 Graphical representation of training and validation accuracy of the model

both accuracy and validation accuracy goes high. During the 5th epoch, there was a slight decrease in accuracy but it comes high in the next epoch, as the training still continued. After the 11th epoch the validation accuracy begins to rise.

6 Conclusion

The proposed model analyzes and provides plant village dataset of various type plant leaf images using Transfer Learning DenseNet-121 which performs the classification of healthy and unhealthy plant leaf images with an accuracy of 91%. In this system, the images are captured by the aerial vehicle are given and compared with the predefined dataset images to classify healthy and unhealthy which helps for early prediction is made so as to give prior information to farmers for the improvement of cotton yield.

7 Future Scope

In the future work, the diseases that are causing the cotton leaves unhealthy can be identified and classified according to the disease of the cotton leaf so that measures can be taken to protect the cotton crop from the disease thus enhancing the cotton yield.

References

A.O. Adedoja, P.A. Owolawi, T. Mapayi, C. Tu, Intelligent mobile plant disease diagnostic system using NASNet-Mobile deep learning. Int. J. Comput. Sci. (IJCS) (2022)

D. Ashourloo, H. Aghighi, A.A. Matkan, M.R. Mobasheri, A.M. Rad, An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. IEEE J. Sel. Top. Appl. Earth Observations Remote Sens. 1–7 (2016)

S. Bargoti, J.P. Underwood, Image segmentation for fruit detection and yield estimation in apple orchards. J. Field Rob. (2017). https://doi.org/10.1002/rob.21699

D. Al Bashish, M. Braik, S. Bani-Ahmad, Detection and classification of leaf diseases using k-means based segmentation and neural network-based classification. Inf. Technol. J. 267–275(2011). ISSN 1812-5638

F. Brugger, Mobile Applications in Agriculture. Syngenta Foundation, Basel, Switzerland (2011)

U.N. Dulhare, M.H. Ali, Underwater human detection using faster R-CNN with data augmentation. Mater. Today Proc. (2021)

U.N. Dulhare, S. Gouse, Automation of rice cultivation from ploughing-harvesting with diseases, pests and weeds to increase the yield using AI, in *Proceedings of the 4th International Conference on Communications and Cyber Physical Engineering (ICCCE)*, vol. 828, pp. 505–513 (2022)

U.N. Dulhare, A. Mohammed, A.M. Hussam. A review on diversified mechanisms for multi focus image fusion, in *International Conference on Communication and Information Processing (ICCIP)* (2019)

V.A. Gulhane, A.A. Gurjar, Detection of diseases on cotton leaves and its possible diagnosis. Int. J. Image Process. (IJIP) **5**(5), 590–598 (2011)

S. Hannunal, N. Anantrasirichi, S. Subrmanian, S. Prashant, A. Jhunjhunwala, C.N. Canagarajah, Agriculture disease mitigation system. ICTACT J. Commun. Technol. Spec. Issue Next Gener. Wirel. Netw. Appl. **2**(2), 364–369 (2011)

H. Lv, W. Song, S. Sun, Service oriented architecture for mobile geospatial information data collection. Comput. Syst. Appl. **21**(7), 43–46 (2012)

B.B. Kalbande, A.S. Pati, Plant tissue culture independent agro bacterium tumefaciens medicated in-plant transformation strategy for upland cotton. J. Genet. Eng. Technol. **14**(1), 9–18 (2016)

S. Kaur, G. Joshi, R. Vig, Plant disease classification using deep learning google net model. Int. J. Innovative Technol. Exploring Eng. (IJITEE) **8** (2019)

K. Kranthi, D. Jadhav, S. Kranthi, R. Wanjari, S. Ali, D. Russell, Insecticide resistance in five major insect pests of cotton in India. Crop Prot. **21**(6), 449–460 (2002)

H. Li, R. Ji, J. Zhang, X. Yuan, H. Kaiqun, L. Qi, WEB- Based intelligent diagnosis system for cotton diseases control. IFIP Adv. Inf. Commun. Technol. **346**, 483–490 (2011)

Y. Li, Z. Cao, H. Lu, Y. Xiao, Y. Zhu, A.B. Cremers, In-field cotton detection via region-based semantic image segmentation. Comput. Electron. Agric. **127**, 475–486 (2016)

Y. Liu, F. Tang, D. Zhou, Y. Meng, W. Dong, Flower classification via convolutional neural network, in *International Conference on Functional- Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA) (Qingdao)*. https://doi.org/10.1109/FSPMA.2016.7818296 (2016)

Y. Lu, S. Yi, N. Zeng, Y. Liu, Y. Zhang, Identification of rice diseases using deep convolutional neural networks. Neurocomputing **267**, 378–384 (2017)

R. Pérez-Zavala, M. Torres-Torriti, F.A. Cheein, G. Troni, A pattern recognition strategy for visual grape bunch detection in vineyards. Comput. Electro. Agric. **151**, 136–149 (2018). https://doi.org/10.1016/j.compag.2018.05.019

K. Prashar, R. Talwar, C. Kant, CNN based on over-lapping pooling method and multi-layered learning with SVM and KNN for American cotton leaf diseases recognition, in *Proceedings of the International Conference on Automation, Computational and Technology Management (ICACTM), London, UK* (2019)

M. Rahnemoonfar, C. Sheppard, Deep count: fruit counting based on deep simulated learning. Sensors (basel). (2017). https://doi.org/10.3390/s17040905

P.R. Rothe, R.V. Kshirsagar, A study and implementation of active contour model for feature extraction: with diseased cotton leaf as example. Int. J. Curr. Eng. Technol. **4**(2) (2014a)

P.R. Rothe, R.V. Kshirsagar, A study and implementation of active contour model for feature extraction with diseased cotton leaf as example. Int. J. Current Eng. Technol. **4**(2), 812–816 (2014b)

S. Satraj Sohrab, E.I. Azhar, M.A. Kamal, P.S. Bhattacharya, D. Rana, Genetic variability of cotton leaf curl beta satellite in Northern India. Saudi J. Biol. Sci. **21**(6), 626–631(2014)

M. Stein, S. Bargoti, J. Underwood, Image based mango fruit detection, localisation and yield estimation using multiple view geometry. Sensors **16**(11), 1915 (2016)

R. Tijare, P. Khade, R. Jain, The survey of disease identification of cotton leaf. Int. J. Innovative Res. Comput. Commun. Eng. (2015)

H.K. Wani, N. Ashtankar, An appropriate model predicting pest/diseases of crops using machine learning algorithms, in *4th International Conference on Advanced Computing and Communication Systems (ICACCS)* (2017)

W. Xu, P. Chen, Y. Zhan, S. Chen, L. Zhang, Y. Lan, Cotton yield estimation model based on machine learning using time series UAV remote sensing data (2021)

X. Yin, X. Liu, J. Chen, D.M. Kramer, Joint multi-leaf segmentation, alignment, and tracking for fluorescence plant videos. IEEE Trans. (2018)

Q. Zheng, W.J. Huang, X.M. Cui, Y.Y. Dong, Y. Shi, H.Q. Ma, L.Y. Liu, Identification of wheat yellow rust sing optimal three-band spectral indices in different growth stages. Sensors **19**, 35 (2019)