Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 2
February 2024

Groundnuts Leaf Disease Recognition Using Neural Network with Progressive Resizing

T. Kosalairaman[1], M. Sathyanarayanan[2], V. Kumaresan[3] K. Sivasankari[4]

[1,2,3]Asst. Prof , Department of Information Technology

V.S.B Engineering College, Karur, Tamilnadu, India.

[4]Asst. Prof , Department of Computer Science, E.S Arts and Science College,Villupuram, Tamilnadu, India.

Abstract
Groundnut is an important oilseed crop worldwide and India is the second largest producer of groundnut. This crop is susceptible to many diseases, which is one of the main reasons for the decline in productivity and quality. Both ultimately lead to a downturn in the agricultural economy. Therefore, a better and more reliable automatic solution for detecting groundnut leaf diseases is needed. In this paper, we propose a deep learning-based progressive resizing model for peanut leaf disease detection and classification tasks. There are five main categories of groundnut leaf diseases: leaf spot, armyworm, wilt, yellow leaf, and healthy leaf. The proposed model was trained with and without stepwise resizing and simultaneously validated using cross-entropy loss.The first dataset of its kind used for training and validation purposes was manually created in the Saurashtra region of Gujarat, India. The resulting dataset was unbalanced because each category had a different number of samples. An advanced focus loss feature was used to solve the dataset imbalance problem. To evaluate the performance of the proposed model, various performance metrics such as accuracy, sensitivity, F1 score, and accuracy were applied. The proposed model achieved an accuracy of 96.bythe state-of-the-art standard12.The progressive resizing model performed better than the traditional cross-entropy loss-based core neural network model.

Keywords – Detection of groundnut leaf diseases, Gradual resizing, Deep Learning, Neural Networks.

I. Introduction

Groundnut cultivation plays an important role in India's agricultural exports and edible oilseed economy. In 2019 alone, the total cultivated area and production of groundnut in India was 3.4,444,931 million hectares or 6,444,862 million tons (IOPEPC, 2019) [1], but the average yield is still low. Disease attack is a major factor contributing not only to early leaf blight but also to productivity and quality decline (Konate et.al., 2020) [2].Therefore, steps must be taken to develop rapid and accurate methods for detecting groundnut leaf diseases in order to sustainably increase productivity. This is very important to various stakeholders. To date, few commercially available tools are available for accurate detection of groundnut leaf diseases, and few high-quality research papers have been published on this topic. One of the main reasons for this could be the lack of benchmark datasets available for research and experiments on peanut leaf disease detection.In this study, a peanut leaf dataset was manually created from Saurashtra, Gujarat. First, all images were captured in square background format with size 3000x3000 pixels (3 color channels), and then the final dataset was created with different image sizes such as 32x32, 64x64, 128x128, 256x256.

A comprehensive review of (Ngugi et al., 2020; Chouhan et al., 2020; Kaur et al., 2019) [3-5] on methods used for leaf disease detection. Image processing, machine learning, and deep learning techniques. Current research shows that only a few researchers are interested in detecting diseases on groundnut leaves. Chen et al.(2019) [6] used spectral index and disease index based on the correlation of the leaf spectral range from 325 nm to 1075 nm, and as a result, the near-infrared spectral reflectance of the tree canopy was It was shown to decrease with increasing disease index. In the regression model, the normalized difference spectral index was R938 and the R2 value was up to R761 0.68 for peanut spot disease detection. Based on the high agreement between the estimated and observed values of the exponential model, they concluded that this model could be used to detect peanut spot disease. An improved random algorithm was used for peanut disease classification (Chaudhary et al.,2016)[7]. It uses instance filter re sampling and attribute scoring methods to balance the class distribution of a multiclass dataset. We proposed a forest classifier.

The proposed method was also applied to five different datasets such as diabetes, soy, hearing, voting, and breast cancer, yielding F1 measurements ranging from 0.89 to 0.97.They argued that the results of their method are effective when the dataset is unbalanced with respect to different numbers of samples between different classes. Ramakrishnan and Sahaya (2015) [8] used back propagation algorithm for detection and classification of peanut leaf diseases. First, RGB is converted to its HSV, then layer separation and color feature extraction steps are performed Dong et al.(2019)[9] applied capsule networks to peanut leaf disease detection and used dynamic routing to overcome rotational invariance and spatial relationship issues. Their empirical observations show that the detection accuracy of the capsule network is 82.17%, which is better than the corresponding value of 81.14% for the convolutional neural network. At this point, it is worth mentioning that capsule networks were originally proposed by Sabour et al.(2017) [10].Vaishnave et al.(2020)[11] used convolutional networks to classify peanut diseases and claimed high accuracy for training and testing on the Plant Village dataset, However, infact, no peanut disease benchmark image dataset has been published or published by plant Village to date (Plant Village, 2020) [12].

Since the introduction of deep learning, many state-of-the-art benchmarks such as DenseNet this architecture have been used (Huang et al., 2017) [13], Deep Residual Learning (He et al., 2016) [14], Inception-v4 (Szegedy et al., 2017) [15], GoogLeNet (Szegedy et al.,2015) VGG (Simonyan and Zisserman, 2014) [17] and AlexNet (Krizhevsky et al., 2012) [18] have been found to provide

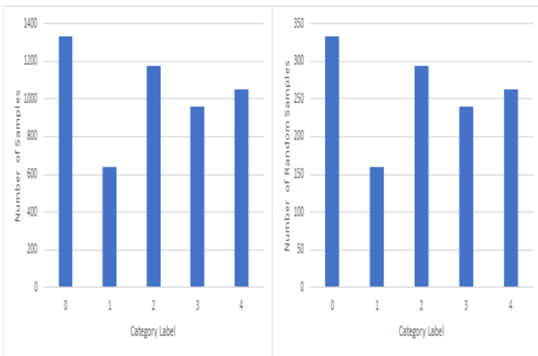
TABLE I GROUNDNUT LEAF DISEASE CLASSES AND CORRESPONDING LABEL

Class Name	Class Label
Groundnut Yellow Leaf	0
Groundnut Wilts	1
Groundnut Leaf Spot	2
Groundnut Healthy Leaf	3
Groundnuts Army worms Effect	4

impressive performance in object detection and various computer vision tasks.Many researchers also use these architectures for transfer learning to detect plant diseases.(Fuentes et al., 2018) [19] Using R-CNN, AlexNet, and GoogLeNet to identify tomato leaf diseases (Liu et al.,2018)[20] Using AlexNet We detected diseases on apple leaves (Kaya et al., 2019; Barbedo 2019; Brahimi et al., 2018; Mohanty et al., 2016)[21-24] used several benchmark architectures for the detection of multiple plant diseases. Recently, methods based on convolutional neural networks are in great demand for detecting plant leaf diseases through automatic deep feature extraction.

An optimized dense convolutional neural network model was proposed to detect and classify corn leaf diseases (Waheed et al., 2020) [25].Gee et al.(2020) [26] proposed a convolutional neural network-based architecture for multi-label learning for plant leaf disease detection and severity estimation. To overcome the problem of unbalanced datasets (Zhong and Zhao, 2020), [27] proposed DenseNet-121 as the backbone network and used regression, multi-label classification, and loss functions to solve apple disease. Has been identified. We focused on three methods they achieved test accuracies of 93, 51%, 93, 31%, and 93.or71%.This was better than 92.29 Accuracy achieved with traditional multiple classification methods using cross-entropy loss functions. Sethy et al.(2020) [28] used different CNN architectures based on deep features using support vector machines to identify leaf diseases in rice.

A self-attention-based CNN architecture was proposed by (Zeng and Li, 2020) [29] to detect leaf diseases in plants Zhang et al.(2019) [30] used a global pooling augmented convolutional neural network to identify cucumber leaf diseases.Karlekar and Seal (2020) [31] proposed his CNN-based Soya Net for soybean leaf disease classification. Your proposed network achieved higher precision, recall, and F1 score compared to other networks.



(a) Training Dataset (b)Testing Dataset

Fig. 1. Data Distribution for Training and Testing

Nearly all existing CNN-based models planned for

Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 2
February 2024

foliar illness location and classification have a few degree of execution in terms of exactness, review, F1 score, and exactness. Be that as it may, when preparing and optimizing CNN-based models on expansive sets of expansive and little pictures or in real-time situations, the time complexity and demonstrate generalization ended up major challenges due to the number of include sets and layers within the engineering. In this paper, we propose a CNN-based dynamic his resizing engineering for demonstrate generalization, optimization, and execution advancement.

II Materials and Methods

A Dataset

There are very few publicly available benchmark datasets for groundnut leaf disease research, based on the best knowledge available from related literature. The five main types of groundnut leaves used in this study were leaf spot, armyworm effects, wilts, yellow leaf, and healthy leaf. The dataset for this study was assembled by hand. To take the pictures, all of the main varieties of symptomatic leaves were physically removed from the plants and placed against a stable background. All of the photos are first taken in square format, measuring 3000 by 3000 with three colour channels. Later, they are all enlarged using progressive resizing to fit into various sizes, including 32 by 32, 64 by 64, 128 by 128, and 256 by 256, for the purpose of developing models. The generated dataset was split into training and testing groups in an 80:20 ratio. Table I lists every class and label that were taken into consideration for this study.Fig. 1 shows the distribution of the training and testing datasets, which further suggests that the dataset is unbalanced because the distribution varies by class. The Focal loss function was used to solve this, and a detailed discussion of it is given in section 2.4. With the exception of healthy leaves, which make up the majority of the dataset, spot disease is the most prevalent foliar disease in groundnut crops, as illustrated by the data presented in Figure 1. Likewise, the dataset's lowest percentage is made up of wilted peanuts. Notable is the fact that it is also a more harmful disease.

B Proposed Network Architecture

To avoid duplication, well-known terms from CNNs including convolutional layers, pooling layers, filters, and fully connected layers are not covered here. The model functions best when the input image size is fixed and is based on a conventional convolutional neural network. When fed huge photos, this model performs well, but the training process is time-consuming and computationally demanding. In order to train, I have to zoom in on extremely small input images and out on enormous input images. In order to increase recognition rate, we took into consideration the gradual resizing strategy (first suggested by Jeremy Howard in his 2018) [32] for model training and generalisation. The following is the suggested working progressive resizing workflow:

Phase 1: A 32 × 32-pixel picture with three colour channels was used to train the initial model.
Phase 2: Upscaled 64x64 photos were used to train the model that follows. During training, the layers and weights of the earlier, smaller model were combined.

Phase 3: 128 × 128 photos were used to train the third model. Phase 3 received the output of Phase 2 as its input.

It is the responsibility of each model to uncover some new patterns and features that were obscured by earlier, smaller models.

Phase 4: 256x256 images were used to build and train the final design. Every large model incorporates older, smaller models into its design.
Fig.2.Proposed CNN based Architecture [Image by Author]

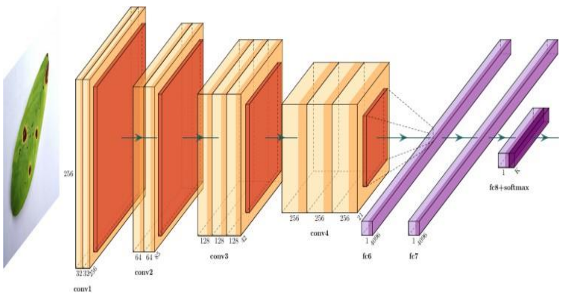


TABLE II TRAINING PARAMETERS

TABLE II. TRAINING PARAMETERS

Parameter	Settings
Final Input size	(256,256,3)
Batch size	32
Learning Rate	1e-3
Optimizer	Adam with decay of 1e-5

TABLE III DISTRIBUTION OF TEST DATASET

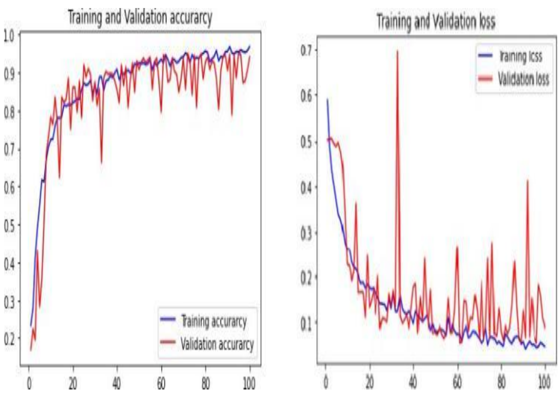
Class	Number of Samples
Groundnuts Healthy Leaf	333
Groundnuts Army worms	263

Effect	
Groundnuts Wilts	160
Groundnuts Leaf Spot	293
Groundnut Yellow Leaf	240

TABLE IV. Cross Entropy Loss without Progressive Resizing

	Precision (%)	Sensitivity (%)	F1 score (%)
Groundnuts Healthy Leaf	0.929509	0.919712	0.924585
Groundnuts Armywomrs Effect	0.929019	0.937031	0.933008
Groundnuts Wilts	0.857722	0.849856	0.853771
Groundnuts Leaf Spot	0.938732	0.930847	0.934773
Groundnut Yellow Leaf	0.919686	0.928934	0.924287
Weighted avg	0.920766	0.918823	0.919774

The suggested model was first constructed at 32 by 32 dimensions before being scaled up by a factor of 4. In this case, X represents the input image's initial size. A number of features were extracted from each phase of the proposed network, which was trained on images of a specific size. Weights were saved with the trained model,



and they were not modified for further training. Phases of the network models based on small images generalise better to larger input sizes and require less processing time; each subsequent phase served to extract additional detectable features that were hidden and not found in previous phases (Howard, 2018) [33].In Figure 2, the suggested combination architecture is displayed. To introduce nonlinearity into the model, a Rectified Linear Unit (ReLU) was used in each convolution operation. The ReLU function F(x) = max(0, x) returns x for all values x> 0 and 0 for all values x ≤ 0.

C Softmax Loss

A Rectified Linear Unit (ReLU) was used in each convolution operation in the model to introduce nonlinearity. The ReLU function, F(x) = max(0, x), yields x for all values x > 0 and 0 for all values x ≤ 0.Here, the categorical cross-entropy loss, or softmax loss, is calculated using the cross-entropy loss function based on the class probabilities generated by softmax activation. $(s)i$ = $\sum Ces$j is the definition of this, where si is the network value for each class i in C. The symbol f(si) is used to symbolise it. In this study on the classification of diseases, labels are considered one-hot. This suggests that the cross-entropy loss, which has the following definition: Only positive classes Cp are taken into account by cross entropy = $-\sum(f(s))$. Target vector t contains non-element zero. Once the sum elements that are zero due to the target label are removed, the cross-entropy loss equation, ti=tp, can be expressed.

D. Lack of concentration

Unbalanced datasets was another issue with the proposed classification. While some classes have nearly twice as many samples as others, some classes have very few samples. A network that has been trained on an unbalanced dataset will be biassed to learn classes with a high sample count preference, underestimating the remaining classes. Utilising a focus loss function that multiplies the cross-entropy loss function by a modulation factor, the problem of class imbalance was resolved. By including focusing and balance parameters in the cross-entropy loss function, the focal local loss function is an enhanced version of the latter. This improves misclassified observations and boosts the network's efficiency. Facebook AI Research (Lin et al., 2017) [33] first suggested the focus loss function for binary classification in object recognition tasks.

The focal loss function has the following general form:

A = -(1 − $pt)\gamma log(pt$)

This is the derivation of the focal loss for multi-class classification.$FocalLoss$ = $-(1 - f\ (s)i)\ \gamma log(f(s)i)$

The magnetic flux parameter in this case is gamma (y) f(s)i, and f(s)I are The entropy cross For multiclass classification, Softmax is employed. This expression is the same as the mutual entropy loss expression if y = 0. It is necessary for the focus parameter y to be ≥ 0.Gamma (y) determines how the curve is shaped. At a given level, the loss of well-classified observations is decreased by increasing the gamma value (y). 1 − f(s)i)y is regarded as the modulation factor in this equation. Lastly, for unbalanced data, the focal loss function can be defined as follows using the balanced parameters: where alpha (*a*) is the equilibrium parameter.Here, alpha refers to the network's weights, which are relatively light.

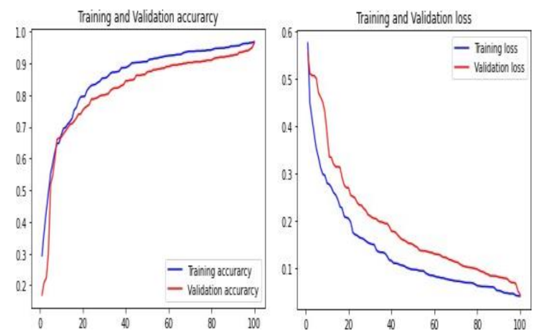
TABLE V.CROSS ENTROPY LOSS WITH PROGRESSIVE RESIZING

	Precision(%)	Sensitivity(%)	F1score(%)
Groundnuts Healthy Leaf	0.963758	0.958712	0.961228
Groundnuts Armyworms Effect	0.961034	0.960001	0.960517
Groundnuts Wilts	0.913827	0.901834	0.907791
Groundnuts Leaf Spot	0.962093	0.959823	0.960957
Groundnut Yellow Leaf	0.942003	0.943748	0.942875
Weighted avg	0.952575	0.949381	0.950971

TABLE VI. FOCAL LOSS(Γ=2)WITHOUT PROGRESSIVE RESIZING

Categories	Precision (%)	Sensitivity (%)	F1 (%
Groundnuts Healthy Leaf	0.937483	0.929594	0.
Groundnuts Armyworms Effect	0.932743	0.942032	0.
Groundnuts Wilts	0.896029	0.88	0.
Groundnuts Leaf Spot	0.9488	0.948921	0.
Groundnut Yellow Leaf	0.926023	0.929782	0.
Weighted avg	0.931809	0.930404	0.

(a)Training and Validation Accuracy. (b)Training and Validation Loss.Fig.3.Training and Validation of Core CNN Model with FocalLoss.



(a) Training and Validation Accuracy. (b) Training and Validation Loss.Fig.4.Training and Validation with Progressive Resizing and Focal Loss.

III. Results and Discussion

Table II displays the training parameters for the suggested model. The final model was trained on images with dimensions of 256x256x3. The model was trained on a range of image sizes beginning at 32x32x3.Following a series of trials, 32 was chosen as the final batch size and le-3 as the learning rate. Utilising Optimizer Adam, LE-5 was brought down. Table VI displays the outcomes of the suggested model with focal loss (η=2) and no stepwise resizing. The reported average precision was 0.930404. This outperforms the accuracy of 0.918823 obtained with cross-entropy loss. Figure. 3(a) or Figure.3(b) displays the training validation accuracy and training validation loss using the focus loss function. Here, a lack of focus led to an uneven training programme that prioritised more challenging observations over simpler ones. Table VII displays the evaluation results of the suggested CNN-based incremental resizing model with focal loss function.

With an average precision of 0.961229, it outperforms all other cases in terms of precision. In essence, lowering the relative loss of well-classified observations is achieved by setting the value $\gamma > 0$. When $\gamma = 2$ was fixed, high accuracy was achieved in the proposed model. The class "Groundnuts Wilts" received the lowest F1 score, and the class "Groundnuts Armyworms Effect" received the highest. The CNN-based progressive resizing model with focal loss function training and validation accuracy and loss are displayed, Fig 4(a) and 4(b), respectively.

IV. Summary

The most prevalent disease on groundnut leaves, spot, is classified into classes along with healthy leaves, armyworm, groundnut wilt, and yellow leaves using CNN-based progressive resizing architecture. Cross-entropy loss functions and focus loss functions were used to assess the suggested architectures with and without progressive resizing. The average accuracy of the proposed model with progressive resizing with focus loss was 96.12%, while the achieved results of the model without progressive resizing are 91.88% and 93.04% using cross-entropy loss and focus loss, respectively. The progressive resizing-based model, which is trained at different scales beginning with small images of size 32x32 and ending with images of size 256x256, is a more general model based on empirical results. The focus loss function aids in addressing the issue of unbalanced data sets, and the results obtained in various scenarios demonstrate that the CNN-based progressive resizing model performs better than the CNN-based core architecture. We intend to use a combination of computer vision techniques to apply the suggested real-time disease detection concept in agriculture in the future.

REFERENCES

[1] Indian Oilseed and Produce Export Promotion Council(IOPEPC,2019),Kharif-2019 Groundnut Crop Survey, retrieved on Feb 4, 2021, from http://www.iopepc.org/misc/2019_20/Kharif%202019%20 Groundnut%20Crop%20 Survey.pdf.

[2] Konate, M., Sanou, J., Miningou, A., Okello, D. K., Desmae, H., Janila,P., & Mumm, R. H. (2020). Past, present and future perspectives on groundnut breeding in Burkina Faso. Agronomy, 10(5), 704.

[3] Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2020). Recent advances in image processing techniques for automated leaf pest anddisease recognition-a review. Information Processing in Agriculture.

[4] Chouhan, S. S., Singh, U. P., & Jain, S. (2020). Applications of computer vision in plant pathology: a survey. Archives of computational methods in engineering, 27(2), 611-632.

[5] Kaur, S.,Pandey, S.,&Goel, S. (2019). Plants disease identification and classification through leaf images: A survey. Archives of Computational Methods in Engineering, 26(2), 507-530.

[6] Chen, T., Zhang, J., Chen, Y., Wan, S., & Zhang, L. (2019). Detection of peanut leaf spots disease using canopyhy perspectral reflectance. Computers and electronics in agriculture, 156, 677-683.

[7] Chaudhary, A., Kolhe, S., & Kamal, R. (2016). An improved random forest classifier for multi-class classification. Information Processing in Agriculture, 3(4), 215-222.

[8] Ramakrishnan, M. (2015, April). Groundnut leaf disease detection and classification by using back probagation algorithm.In 2015International Conference on Communications and Signal Processing (ICCSP) (pp.0964-0968). IEEE.

[9] Dong,M.,Mu,S.,Su,T.,& Sun,W.(2019,August).Image Recognition of Peanut Leaf Diseases Based on Capsule Networks. In International CCF Conference on Artificial Intelligence (pp. 43-52). Springer, Singapore.

[10] Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic routing between capsules. In Advances in neural information processing systems (pp. 3856-3866).

[11] Vaishnnave, M. P., Devi, K. S., &Ganesh kumar, P. (2020). Automatic method for classification of groundnut diseases using deep convolutional neural network. SOFT COMPUTING.

[12] Plantvillage (2020), Topics, retrieved on August 08, 2020, fromhttps://plantvillage.psu.edu/topics.

[13] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017).Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[15] Szegedy, C., Ioffe, S., Vanhoucke, V., &Alemi, A. (2017, February).Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 31, No. 1).

[16] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ...& Rabino vich, A.(2015).Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[17] Simonyan, K., &Zisserman, A. (2014). Very deep convolutionalnetworks for large-scale image recognition. arXiv preprintarXiv:1409.1556.

[18] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances inneural information processing systems, 25, 1097-1105.

[19] Fuentes, A. F., Yoon, S., Lee, J., & Park, D. S. (2018). High-performance deep neural network-based tomato plant diseases and pests diagnos is system with refinement filter bank. Frontiers in plant science, 9, 1162.

[20] Liu, B., Zhang, Y., He, D., & Li, Y. (2018). Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry,10(1),

[21] Kaya, A., Keceli, A. S., Catal, C., Yalic, H. Y., Temucin, H.,&Tekinerdogan, B. (2019). Analysis of transfer learning for deep neural network based plant classification models. Computers and electronics inagriculture, 158, 20-29.

[22] Barbedo, J. G. A. (2016). A review on the main challenges in automaticplant disease

identification based on visible range images. Bio systems engineering, 144, 52-60.

[23] Brahimi,M.,Arsenovic,M.,Laraba,S.,Sladojevic, S.,Boukhalfa,K.,&Moussaoui, A. (2018). Deep learning for plant diseases: detection and saliency map visualisation. In Human and machine learning(pp. 93-117). Springer, Cham.

[24] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in plant science, 7, 1419.

[25] Waheed, A., Goyal, M., Gupta, D., Khanna, A., Hassanien, A. E., & Pandey, H. M. (2020). An optimized dense convolutional neural network model for disease recognition and classification in corn leaf. Computers and Electronics in Agriculture, 175, 105456.

[26] Ji, M., Zhang, K., Wu, Q., & Deng, Z. (2020). Multi-label learning for crop leaf diseases recognition and severity estimation based on convolutional neural networks. Soft Computing, 24(20), 15327-15340.

[27] Zhong, Y., & Zhao, M. (2020). Research on deep learning in apple leafdisease recognition.Computers and Electronics in Agriculture, 168,105146.

[28] Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020).Deep feature based rice leaf disease identification using support vector machine. Computers and Electronics in Agriculture, 175, 105527.

[29] Zeng, W., & Li, M. (2020). Crop leaf disease recognition based on Self-Attention convolutional neural network. Computers and Electronics in Agriculture, 172, 105341.

[30] Zhang, S., Zhang, S., Zhang, C., Wang, X., &Shi, Y. (2019).Cucumber leaf disease identification with global pooling dilated convolutionalneural network.Computers and Electronics in Agriculture,162, 422-430.

[31] Karlekar, A., & Seal, A. (2020). SoyNet: Soybean leaf diseases classification. Computers and Electronics in Agriculture, 172, 105342.

[32] Jeremy Howard (2018), Now anyone can train Image net in 18 minutes,retrievedonJan1,2021,fromhttps://www.fast.ai/2018/08/10/fastai-diu-imagenet/.

[33] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).