Transfer Learning Approach for Classification of Apple Leaf Disease

Bhavani R and Priyanka V

Research Journal of Agricultural Sciences
An International Journal

P- ISSN: 0976-1675
E- ISSN: 2249-4538

Volume: 13
Issue: 04

Res. Jr. of Agril. Sci. (2022) 13: 1191–1195

CARAS

Res. Jr. of Agril. Sci. (Jul-Aug 2022)
13(4): 1191–1195

www.rjas.org

ISSN: 0976-1675 (P)
ISSN: 2249-4538 (E)

Full Length Research Article

Transfer Learning Approach for Classification of Apple Leaf Disease

Bhavani R[1] and Priyanka V*[2]

Received: 24 May 2022 | Revised accepted: 29 Jul 2022 | Published online: 03 Aug 2022

A B S T R A C T

Economy of most of the nations in the world depends on agriculture. One of the serious causes of less productivity in agriculture is leaf diseases in plants. Apple leaves suffer from three common diseases namely Apple scab, Black rot and Cedar rust. This paper aims to develop a deep learning model using transfer learning approach to classify the apple leaf disease. Four pretrained models namely Inceptionv2, EfficientNetV2, ResNetV2 and MobileNet were analysed for this purpose. The top layers of the pretrained models were frozen and a dense layer was introduced for classifying the three-leaf disease category. The experimented results show that MobileNet is well suited for apple leaf disease classification with less time and an accuracy of 99.21%.

Key words: Transfer learning, Leaf disease, Image classification, Convolutional neural network, Deep learning

Agriculture is the main source of livelihood for 70% of population in India. One of the major struggles faced by farmers in agriculture is the plant disease [1]. Early detection and controlling infections may increase the agricultural productivity. Current research in this field aims at developing automatic disease detection algorithms. In early days it was with only naked eyes experts identify plant leaf disease. It requires lot of expert knowledge about the particular type of plant. This work aims at developing image-based detection of leaf disease in plants. Image processing is the area of research that takes image as input and provides useful information as output. The most common steps followed in image processing are given in (Fig 1). Images of infected leaves are gathered first. They are then preprocessed. Image preprocessing step include resizing, reshaping, rescaling, etc. Next preprocessed images are segmented to locate the lesion portion of the leaf. After segmentation, features are extracted from the lesion image [2]. Most commonly extracted features from lesion image are color features, shape features and texture features. The final step is to classify the diseased leaves using the extracted features.

Classification of plant leaf disease is carried out using machine learning approaches. Here, a classifier model is built from the features extracted from images. The built model is used to classify a new sample with the classes learnt earlier by the machine learning model. It is a time-consuming task to segment the images and then extracting the features to build a classification model [3]. Deep learning-based classifiers build the model with the features learned automatically from the input images. There is no need for separate feature extraction step in deep learning. The accuracy of the deep learning model is high compared to machine learning models as they learn from the whole input space unlike machine learning models that learn from the lesion portion of the image. Transfer learning is one of the key features of deep learning models. It uses the knowledge gained for solving a problem to solve another different but related problem. It does not need large amount of data to train the network [4]. In this work, transfer learning is used for apple leaf disease classification. The comparative analysis of the performance of four different deep learning Convolutional Neural Network (CNN) models viz., InceptionV2, EfficientNetV2, ResNetV2, and MobileNet for apple leaf disease classification is carried out in this paper.

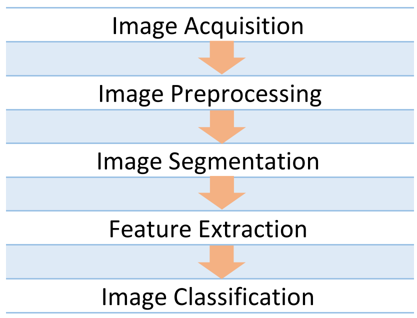

Fig 1 Image classification steps

Related work

Over several years, many techniques were proposed by researchers for plant leaf disease classification. The images in RGB color space is converted to HSV color space. Next images are segmented using k-means clustering and the green pixels are

* **Priyanka V.**
priyanka07vv@gmail.com

[1-2] Department of Computer Science and Engineering, Government College of Technology, Coimbatore - 641 013, Tamil Nadu, India

masked. Additionally, the pixels with zero reds, green and blue values and the pixels on the boundaries of the infected clusters are masked using Ostu's method. Next color co-occurrence texture analysis is developed using spatial gray-level dependence matrices. In addition to that, texture features like angular moment, product moment, entropy, correlation, contrast are extracted Matlab based implementation of feed forward neural network is used for classification. The number of neurons in output layer is 6 to represent 5 disease class and one healthy class. The classification accuracy of 94% is achieved [5].

Gabor wavelet transform technique is used to extract relevant features from diseased tomato leaf. Gabor filters are linear filters that are used to analyze the texture features of images. Support vector machine (SVM) classifier with kernel functions namely Cauchy kernel, Invmult kernel and Laplacian kernel are used for classification. Since SVM is a binary classifier, only two classes of tomato leaf disease are considered. It yields an accuracy of 95% [6].

Potato leaf diseases are classified using neural network approach [7]. Color, texture and shape features are extracted. Eighteen color features viz., Red Mean, Red Variance, Red Range, Green Mean, Green Variance, Green Range, Blue Mean, Blue Variance, Blue Range, Hue Mean, Hue Variance, Hue Range, Saturation Mean, Saturation Variance, Saturation Range, Value Mean, Value Variance, Value Range are considered. Different texture features extracted from lesion images are energy, entropy, contrast, homogeneity and co-relation. Only one shape feature considered in this paper is Area. Totally twenty-four features are extracted and given as input to Backpropagation neural network and categorized into four classes. The system is able to get a classification accuracy of 92%. Classification of leaf diseases in grapes namely Anthracnose, Powdery Mildew and Downy Mildew are considered [8]. They have used real field data for classification. Their approach starts with removal of background colors followed by image enhancement using decorrelation stretching and removal of sun spots. Next the green pixels are masked using thresholding method. Location Binary Pattern (LBP-HF) is extracted. The features are classified using machine learning algorithms like Artificial Neural Network, SVM and Random Forest (RF). Their results show that RF outperforms ANN and SVM with the classification accuracy of 86% for LBP-HF features.

Texture feature-based image classification is carried out [9]. Leaf images of various diseased features are extracted using Gray level co-occurrence matrix (GLCM). In order to do this, the RGB images are converted to gray scale images and the GLCM is applied to get texture features. K-Means clustering algorithm is used to cluster the images to groups according to their similarity. The lazy learning algorithm, K-nearest neighbor is applied to classify the image features into multiple classes and classification accuracy of 92% is achieved.

Major part of the plant diseases is brought about by the assault of microorganisms, development and contamination [10]. Images are captured with digital camera and a database is built. RCB images are converted to HSV format as it is a human perception color. The appearance of the images is enhanced by highlighting the important features in it. Images are segmented to retrieve the constituent parts of interest using Principal Component Analysis (PCA). PCA method reduces the dimensionality of features while maintaining the information without affecting the data. Color, texture and morphological features are extracted and images classification is carried out using Naive Bayes classifier. The classification accuracy of 97% is achieved.

Cucumber leaf diseases are classified using three machine learning algorithms SVM, K-NN and Decision tree (DT) [11]. Matlab is used for feature extraction and classification. RGB images of cucumber leaves are first downloaded and converted to L*a*b color model. Images are segmented using K-means clustering and SIFT (Scale-invariant Function) features are extracted. The features were classified using SVM, K-NN and DT. 91% of classification accuracy is achieved for SVM and it outperforms K-NN and DT in classification.

Ensemble learning not only considers a single classifier but a set of classifiers [12]. The combined prediction of classification based on voting is done. Tomato leaf disease classification is carried out using texture features extracted from GLCM. Multiple classifiers namely SVM, Multilayer perceptron (MLP) and RF are used for classification. A soft voting classifier predicts the output based on highest probability of chosen class as output. The individual classification accuracy is 88.74%, 89.84% and 92.86% for RF, MLP and SVM classifier. Soft voting-based ensemble learning yielded an accuracy of 93.13%.

The above literature study is about the leaf disease classification using machine learning approaches. It is learnt that the machine learning algorithm performs well even if the amount of data is small. When amount of data is large machine learning algorithms cannot perform well. It may lead to overfitting. Also in machine learning, features need to be identified by an expert system and then hand-coded as per domain and datatype. But deep learning algorithms try to learn high level features from data, reducing the task of developing features for every problem [13]. The following paragraph explains how CNN is applied for various applications.

A new approach of using deep learning method was explored in order to automatically classify and detect plant diseases from leaf image. CNN with varying number of convolutional layer and pooling layer are proposed for plant leaf disease classification [14-18]. The accuracy of models proposed is in the range of 96% to 98%. If the size of the dataset increases then the time taken for training the model also increases. Tensorflow hub and Keras application has a repository of the trained CNN models on different applications. Using transfer learning approaches the pretrained models were used for splicing and copy-move image tampering detection [19], brain tumour classification [20], tuberculosis culture diagnosis [21], Anomaly detection [22], speech recognition [23]. This paper aims at applying transfer learning approach for classification of apple leaf diseases.

MATERIALS AND METHODS

The steps followed in the proposed approach for classifying plant leaf diseases are shown in (Fig 2).

Table 1 Summary of diseased leaves taken for study

Leaf disease name	Number of images
Apple Scab	630
Apple Blackrot	621
Apple Cedar Rust	275

Input dataset

In this work, images were downloaded from plant village dataset. The dataset consists of 54303 diseases leaf images classified into 38 categories.

For experimentation only 3 categories are considered in this paper. The categories and the number of leaf images in each category are given in (Table 1).

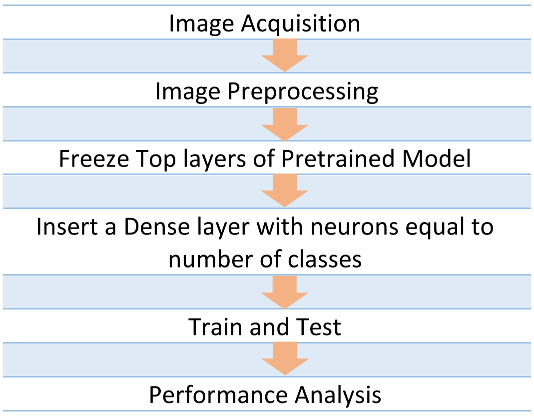

Fig 2 Steps followed in transfer learning approach

Data preprocessing

Pretrained weights of several CNN models are available in Tensorflow hub which could be used for solving other similar problem. Each CNN model accepts input of its own size. The raw image downloaded is of size 256*256. In preprocessing step, all the images are resized to a value that is accepted by the pretrained model. Next all the images are converted to an array and the values of the array are converted in the range of 0 and 1. Input to the CNN models is the RGB images of diseases leaf. So, the images are resized to 224*224 with three channels.

Model building

The pretrained models of InceptionV2, EfficientNetV2, ResNetV2 and MobileNet are considered for experimentation in this paper. In general, CNN have many convolutional layer, maxpooling layer and final fully connected layer. The number of convolutional layer, maxpooling layer and final fully connected layer differ for each model.

Convolutional layers are the core part of CNN where lot of computation occurs. A filter moves across the receptive fields of the image checking whether the feature is present [24]. Next the filter shifts by a stride, repeating the process until the filter has slid across the entire image. The final output is the feature map which is the dot product of input image pixels and the filter. (Fig 3) illustrates the convolution operation of input image size 5*5 and filter size 3*3 and stride 1.

Pooling layer perform down sampling by reducing the number of parameters in the inputs [24]. In the pooling layer the filter does not have any weights. The kernel applies an aggregation function to the values within the receptive field. (Fig 4) illustrates the maxpooling with filter 2*2 and stride 1.

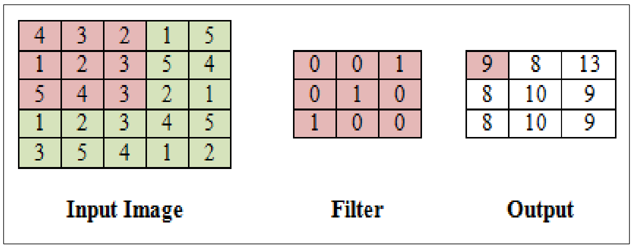

Fig 3 Convolution process

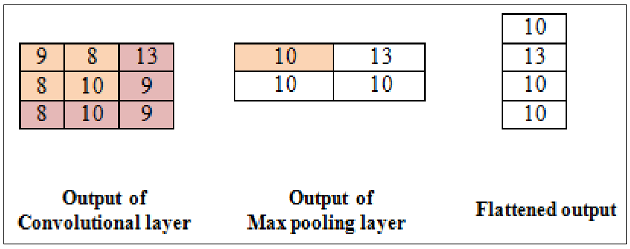

Fig 4 Maxpooling process

In fully connected layer, each node in the output layer connects directly to a node in the previous layer. This layer performs classification based on the features learned from previous layer [24]. Convolutional and pooling layer use ReLu function but fully connected layer use softmax activation function which gives the output between 0 and1. Parameters are the weights that are learnt during training. (Table 2) summarizes the total number parameters and the number of trainable parameters in the models considered for study. In each model, weights of the layer before the fully connected layer are frozen and layers after that are removed. A dense layer with 3 neurons, i.e., the number of categories considered in the study is added. Since the top layers of CNN models are frozen, their weights are not updated. Only the weights of the newly added layers are updated. So, the number of trainable parameters is less while using transfer learning approach.

Table 2 Parameters in CNN models

Model name	Total parameters	Trainable parameters
InceptionV2	10176187	3075
EfficientNetV2	6934967	3843
ResNetV2	58337795	6147
MobileNet	3231939	3075

Train and test

For training and testing the model built, the dataset is split in the ratio of 80:20. 80% of the images were used for training and the remaining 20% of the images were used for testing the model. The models were trained and the accuracy was recorded for different number of epochs.

Performance metrics

The metrics used to evaluate the model of classification is accuracy and is given in Eq. (1) respectively [25].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \ldots\ldots\ (1)$$

Where TP, TN, FP, FN refers to True Positives, True Negatives, False Positives, False Negatives respectively.

RESULTS AND DISCUSSION

The classification of diseased apple leaf images were carried out using Python and Tensorflow framework in Google Colab with GPU. The dataset is split in 80:20 ratio for training and testing. Input shape of the image is (222, 224, 3). The CNN models were trained for 10 epochs with batch size of 36. The average time taken for each epoch is tabulated in (Table 3). It is observed that depending upon the number of trainable parameters, the time taken for each epoch in the CNN models varies accordingly. Time taken by MobileNet (92ms) is minimum compared to other three models. During training and testing process, the accuracy of the models recorded is given in (Table 4). The same is graphically represented in (Fig 5).

Table 3 Time analysis

Model name	Average time taken for 1 Epoch
InceptionV2	165 ms
EfficientNetV2	170 ms
ResNetV2	635 ms
MobileNet	92 ms

From the graphs in (Fig 5), it is clearly understood that MobileNet performs well for apple leaf disease classification compared to Inceptionv2, EfficientNetV2 and ResNetV2. It is also observed that transfer learning yields better results with small amount of data and in very minimum epochs. The difference in testing accuracy and training accuracy is 2.35, 1.39, 1.31 and 0.70 for Inceptionv2, EfficientNetV2, ResNetV2 and MobileNet respectively. This deviation shows that in 10 epochs the first three models are over fitted. The difference in train accuracy and testing accuracy is only 0.70 for MobileNet. This shows that MobileNet is suitable for classifying the apple leaf diseases.

Table 4 Accuracy of the models

Model name	Training accuracy	Testing accuracy
InceptionV2	97.90	95.55
EfficientNetV2	98.51	97.12
ResNetV2	98.69	97.38
MobileNet	99.91	99.21

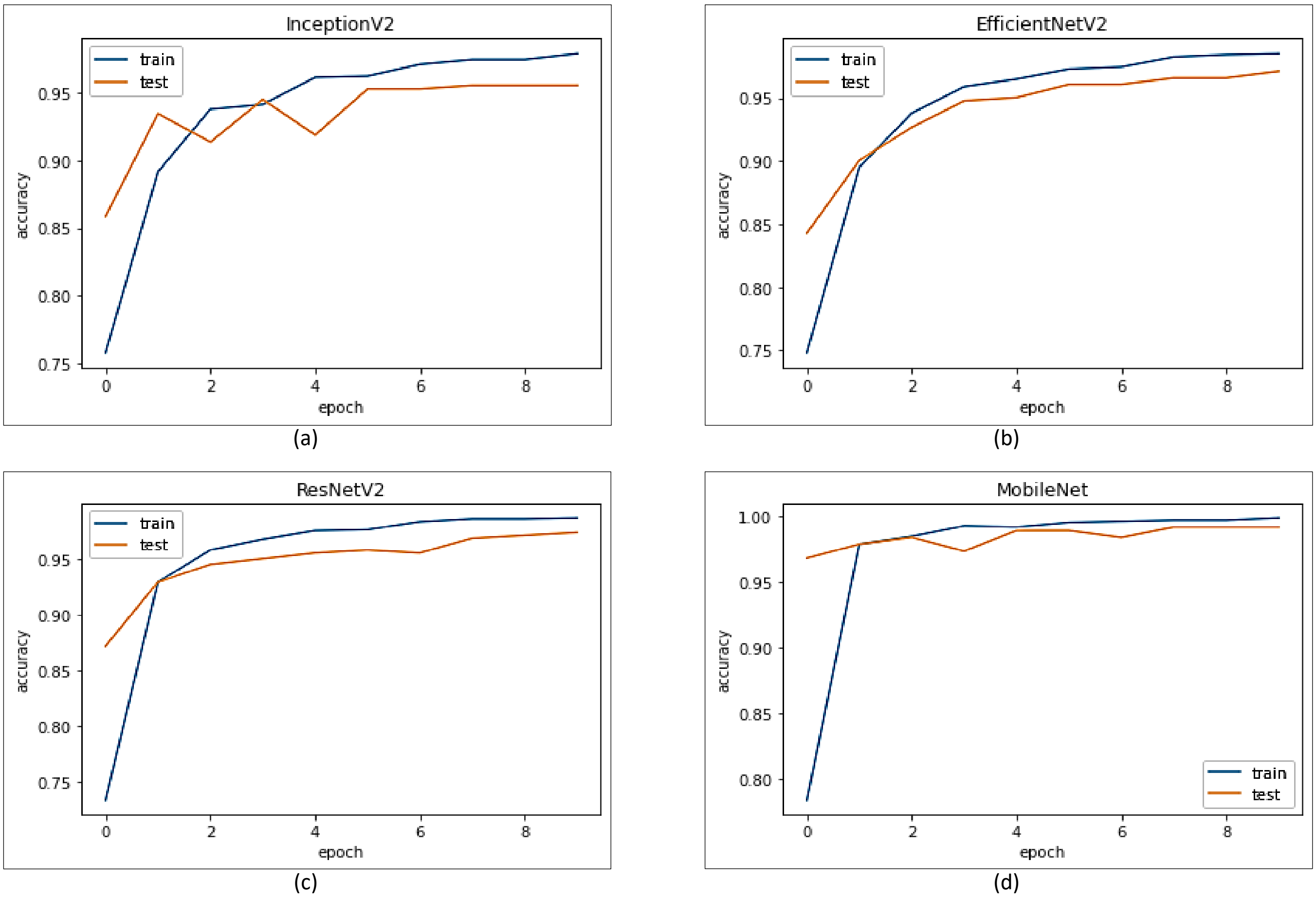

Fig 5 Training and testing accuracy with respect to number of epochs while fine tuning the models (a) InceptionV2 (b) EfficientNetV2 (c) ResNetV2 (d) MobileNet

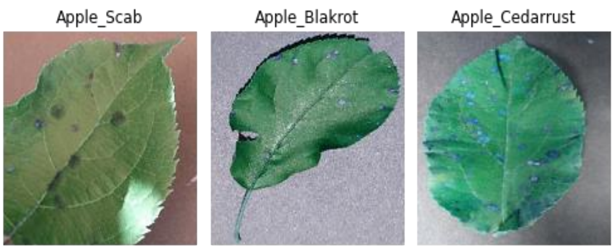

Fig 6 Sample leaf images

The (Fig 6) shows the sample leaf image for each category and (Fig 7) is the screenshot of the web-based application developed to classify the input leaf image using the developed model. First, the user has to upload an image of the diseased leaf and when the user clicks the predict button, the leaf image with its predicted class label is displayed. In (Fig 7), the uploaded image is correctly classified as Apple Scab disease.

Fig 7 Screenshot of the web-based application

CONCLUSION

Leaf diseases in plants cause a great loss to farmers. This paper aimed to develop an efficient deep learning model to identify the type of leaf disease in less time and more accuracy. Transfer learning was followed to build new models from the pretrained models. Four pretrained deep learning models namely Inceptionv2, EfficientNetV2, ResNetV2 and MobileNet were considered for this experimentation with minimum number of image data. Classification models gave accuracy in the range of 95.5% to 99.21%. The performance of MobileNet was better for apple leaf disease classification compared to other three models in terms of time and accuracy. In future the same model could be used to classify more number of leaf diseases. Also, a mobile app could be developed and deployed to help the farmers know about the infection of plant with a single photograph of the image.

LITERATURE CITED

1. Ayaz M, Ammad-Uddin M, Sharif Z, Mansour A, Aggoune EHM. 2019. Internet-of-things (IoT)-based smart agriculture: Toward Making the fields talk. *IEEE Access* 7: 129551-129583.
2. Khirade SD, Patil AB. 2015. Plant disease detection using image processing. *In*: *Proceedings of International Conference on Computing Communication Control and Automation*. pp 768-771.
3. Sharma P, Hans P, Gupta SC. 2000. Classification of plant leaf diseases using machine learning and image preprocessing techniques. *In*: *Proceedings of International Conference on Cloud Computing, Data Science and Engineering*. pp 480-484.
4. Tsiakmaki M, Kostopoulos G, Kotsiantis S, Ragos O. 2020. Transfer learning from deep neural networks for predicting student performance. *Applied Sciences* 10(6): 2145.
5. Al-Hiary H, Sulieman BA, Reyalat M, Braik M, Alrahamneh Z. 2011. Fast and accurate detection and classification of plant diseases. *International Journal of Computer Applications* 17(1): 31-38.
6. Mokhtar, Usama, Mona AS Ali, Hassenian AE, Hefny H. 2015. Tomato leaves diseases detection approach based on support vector machines. *In*: *Proceedings of International Computer Engineering Conference*. pp 246-250.
7. Athanikar G, Priti B. 2016. Potato leaf diseases detection and classification system. *International Journal of Computer Science and Mobile Computing* 5(2): 76-88.
8. Biswas S, Avil S, Sanat S, Srinivasu P. 2016. Random forest-based classification of diseases in grapes from images captured in uncontrolled environments. *In*: *Proceedings of International Conference on Signal Processing*. pp 1775-1780.
9. Gautam K, Bala R. 2017. GLCM and KNN based algorithm for plant disease detection. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* 6(7): 5845-5852.
10. Mohanapriya K, Balasubramani M. 2019. Recognition of unhealthy plant leaves using naive bayes classifier. *In*: *IOP Conference Series: Materials Science and Engineering* 561(1): 012094.
11. Nandhini N, Bhavani R. 2020. Feature extraction for diseased leaf image classification using machine learning. *In*: *Proceedings of International Conference on Computer Communication and Informatics*. pp 1-4.
12. Jeyalakshmi S, Radha R. Classification of tomato diseases using ensemble learning. *ICTACT Journal on Soft Computing* 11(4): 2408-2415.
13. Krishna ST, Kalluri HK. 2019. Deep learning and transfer learning approaches for image classification. *International Journal of Recent Technology and Engineering* 7(584): 427-432.
14. Srdjan S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. 2016. Deep neural networks-based recognition of plant diseases by leaf image classification. *Computational Intelligence and Neuroscience* 2016: Article ID:3289801.
15. Prabavathi S, Kanmani P. 2021. Plant leaf disease detection and classification using optimized CNN Model. *International Journal of Recent Technology and Engineering* 9(6): 233-238.
16. Prajwala T, Pranathi A, Ashritha KS, Chittaragi NB, Koolagudi SG. 2018. Tomato leaf disease detection using convolutional neural networks. *In*: *Proceedings of International Conference on Contemporary Computing*. pp 1-5.
17. Zhang S, Zhang S, Zhang C, Wang X, Yun S. 2019. Cucumber leaf disease identification with global pooling dilated convolutional neural network. *Computers and Electronics in Agriculture* 162: 422-430.
18. Liu B, Zefeng D, Liangliang T, Dongjian H, Li S, Wang H. 2020. Grape leaf disease identification using improved deep convolutional neural networks. *Frontiers in Plant Science* 11: Article ID: 1082.
19. Hebbar NK, Kunte AS. 2021. Transfer learning approach for splicing and copy-move image tampering detection. *ICTACT Journal on Image and Video Processing* 11(4): 2447-2452.
20. Deepak S, Ameer PM. 2019. Brain tumor classification using deep CNN features via transfer learning. *Computers in Biology and Medicine* 111: Article ID:103345.
21. Chang RI, Yu-Hsuan C, Jeng-Wei L. 2020. Two-stage classification of tuberculosis culture diagnosis using convolutional neural network with transfer learning. *The Journal of Supercomputing* 76: 8641-8656.
22. Bansod S, Abhijeet N. 2019. Transfer learning for video anomaly detection. *Journal of Intelligent and Fuzzy Systems* 36(3): 1967-1975.
23. Matassoni M, Roberto G, Daniele F, Diego G. 2018. Non-native children speech recognition through transfer learning. *In*: *Proceedings of International Conference on Acoustics, Speech and Signal*. pp 6229-6233.
24. Michael AN. 2015. *Neural Networks and Deep Learning*. Determination Press.
25. Han J, Jian P, Micheline K. 2011. *Data Mining: Concepts and Techniques*. Elsevier 2011.