

# Accurate plant species analysis for plant classification using convolutional neural network architecture

Savitha Patil<sup>1</sup>, Mungamuri Sasikala<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, Sharnbasva University, Kalaburagi, India

<sup>2</sup>Department of Electrical and Electronics Engineering, Sharnbasva University, Kalaburagi, India

## Article Info

### Article history:

Received Aug 22, 2022

Revised Jul 7, 2023

Accepted Aug 5, 2023

### Keywords:

Architecture

Convolutional neural network

Feature extraction

Plant classification

Plant identification

Vietnam plant-200 dataset

## ABSTRACT

Recently, plant identification has become an active trend due to encouraging results achieved in plant species detection and plant classification fields among numerous available plants using deep learning methods. Therefore, plant classification analysis is performed in this work to address the problem of accurate plant species detection in the presence of multiple leaves together, flowers, and noise. Thus, a convolutional neural network based deep feature learning and classification (CNN-DFLC) model is designed to analyze patterns of plant leaves and perform classification using generated fine-grained feature weights. The proposed CNN-DFLC model precisely estimates which the given image belongs to which plant species. Several layers and blocks are utilized to design the proposed CNN-DFLC model. Fine-grained feature weights are obtained using convolutional and pooling layers. The obtained feature maps in training are utilized to predict labels and model performance is tested on the Vietnam plant image (VPN-200) dataset. This dataset consists of a total number of 20,000 images and testing results are achieved in terms of classification accuracy, precision, recall, and other performance metrics. The mean classification accuracy obtained using the proposed CNN-DFLC model is 96.42% considering all 200 classes from the VPN-200 dataset.

*This is an open access article under the [CC BY-SA](#) license.*



## Corresponding Author:

Savitha Patil

Department of Computer Science and Engineering, Sharnbasva University

Kalaburagi, Karnataka, India

Email: SavithaPatil123456789@gmail.com

## 1. INTRODUCTION

There are numerous environmental resources on the planet and one of the most essential and advantageous environmental resources is plants. Plants are the most essential element for the survival of humans and a key resource of all the available ecological resources. Plants can be of different varieties such as green plants, mossy plants, flowering plants, grass, wine plants, and seed plants (angiosperms and gymnosperms). The plant is extremely important to human society because they contribute massively to providing human food and they generate synthetic starch with the help of the photosynthetic process. Further, plants absorb carbon-di-oxide ( $CO_2$ ) gas and exhibit oxygen ( $O_2$ ) gas, which is the most essential element for human survival. It also controls ecological conditions like temperature, global warming, and humidity. According to research conducted by the food and agriculture organization (FAO) in the United Nations of America (USA), the world population will grow up to 9.1 billion by the year 2050. Thus, the nutrition production rate needs to be increased by 70% to provide nutrition to such a huge number of people by the year 2050 [1]. However, multiple factors can heavily affect the growth of nutrition production rates such as limited clean water and the absence of large areas for cultivation.

Furthermore, diseases in crops certainly do not help in increasing the production rate of nutrition as they massively attack the quality as well as quantity of crops. The existence of diseases in plants hurts the food production rate. These diseases in plants can be of various types but plant disease can be identified by precisely detecting the types of marks or lesions that occurred on the leaves, flowers, fruits, or stems. Usually, plant disease starts from leaves and can be controllable if identified early. Every disease on the plant leaves has some unique patterns which are also called abnormalities. By identifying these abnormalities, plant disease identification, and analysis of their symptoms can be possible [2]. If diseases do not identify in the initial stages of crops production, then food insecurity will enhance, and in these types of cases, crops become wasted more often [3]. The most effective solution to avoid these types of cases is early detection of diseases in plants so that they can be prevented from disease and proper disease control ideas and precautions always play a key role in the management or decision-making of plant production. Furthermore, image analysis and classification of plant species have gained massive attention in the last few years, especially in the field of machine learning and computer vision. The main objective of computer vision and machine learning techniques is used to analyze and identify images belonging to numerous categories or meta-categories. These categories can be varied kinds of plants, animals, vehicles, retail products, and medicines. The primary objective and challenge to understanding these images are analyzing fine-grained visual variations so that objects can be distinguished efficiently among all the objects with similar appearances. However, all the objects have different characteristics. The identified discriminative region generates high-quality features which carry the most significant and distinctive information about an image. Based on these distinctive features, the classification of plant leaf species can be achieved successfully. However, the extraction of discriminative features from plant leaf species requires a strong feature extraction technique. Thus, deep learning methods can be a powerful tool to extract discriminative features from plant leaf species. Recently, deep learning methods have found several breakthroughs in the analysis of discriminant features and learning of fine-grained characteristics of plant leaf images [4]–[7].

However, there are a few problems associated with the traditional deep learning-based discriminant feature extraction methods through deep learning methods such as high-class variance, object similarities, complex backgrounds, and poor fine-grained analysis. Therefore, a convolutional neural network based deep feature learning and classification (CNN-DFLC) model is employed to identify plant leaf species and classify plant images belonging to exactly which class. The proposed CNN-DFLC model distinguishes plant species among several classes. The proposed CNN-DFLC model obtains the most significant information from discriminative image regions so that efficient training is performed and improved classification accuracy is obtained. The proposed CNN-DFLC model is tested on the Vietnam dataset and classification performance can be measured on the testing dataset using obtained fine-grained discriminative features. The proposed CNN-DFLC model comparably improves the identification efficiency of plant leaf images.

## 2. LITERATURE SURVEY

In this world, there is an abundant amount of plants present and the leaves of these plants are the same in color, appearance, and shape. As a result, the classification of plant leaf species becomes a challenging and complex process. To distinguish between medicinal and non-medicinal plants, extraction of fine-grained discriminative features is quite important which can be achieved using deep learning methods. Recently, many deep learning methods are presented by different researchers to identify medicinal plants among several plant categories. One of the best deep learning methods for plant leaf identification among several categories can be CNN architecture. Some of the research works are presented in the next paragraph regarding the classification of plant leaves through CNN architecture.

A detection and classification method for the analysis of plant species and diseases is reviewed using deep learning methods [8]. The deep learning method is utilized for handling challenges and learning essential features of plant leaf images. The latest and advanced imaging techniques can be utilized to improve efficiency and obtain discriminative features. Plant type classification [9] is performed for feature filtering and fine-grained features. Here, Adaboost.M1 and LogitBoost algorithms are utilized to improve plant classification efficiency. Here, the classification of plant species is obtained using four types of classifiers such as k-nearest neighbors (kNN), random forest (RF), support vector machine (SVM), and multi-layer perceptron (MLP). A deep learning method [10] is presented to detect and classify plant diseases. Here, low-intensity information is obtained from the background and foreground of the image. Further, to acquire information related to the images such as image structure, chrominance, and image positions, deep learning methods are utilized. Here, a disease classification system of plants is enabled to get the information related to the plant and to handle plant diseases. Mathulapragasan and Lanthong [11], a leaf disease detection system is utilized to classify cassava leaves based on CNN architecture. Here, testing results are obtained using the DenseNet121 model, and obtained classification accuracy using this DenseNet121 model is 94.32% and the F1-score at 92.13%. A deep residual dense network [12] is presented to identify tomato leaf diseases. A hybrid deep learning technique is

adopted to improve the efficiency of the deep residual dense network. This technique significantly reduces several training parameters to enhance classification accuracy. Haider *et al.* [13], disease classification and verification mechanisms are presented to improve knowledge-based decisions. Jin *et al.* [14], deep learning methods are utilized to identify weed plant species and a training image dataset is adopted using image processing techniques and reduces Bayesian classification errors. The center-net model is utilized to achieve precision and recall of 95.6% and 95%, respectively. This model significantly reduces the computational cost. A fine-grained-generative adversarial network (GAN) method is adopted to identify leaf spot diseases that occurred in grape leaves [15]–[17]. Therefore, the CNN-DFLC model is presented to identify plant leaf classes among several classes. The next section discusses the method related to the proposed CNN-DFLC model.

### 3. MODELLING FOR MODEL

This section discusses the method regarding the proposed CNN-DFLC model for quality features extraction from the given plant input images so that efficient classification is performed and evaluates among several classes which image belongs to which class. The successful implementation of the proposed CNN-DFLC model can provide efficient plant analysis and classification. Most of the research works are focused on the identification of plant diseases (type of diseases). However, very few methods are focused on the detailed study of plant classification, and can efficiently classify which image belongs to which class among available several classes. Plant classification is a complex and challenging process and has been given very little attention, especially for the classification of around 200 classes by using an advanced deep learning architecture CNN-DFLC model. There are numerous species present across the world related to plants and the identification of which plant belongs to which species, is a challenging process. Therefore, in this research work, a deep learning-based plant classification process is performed to identify accurate classes of plant images using the proposed CNN-DFLC model. Based on the efficient training of the proposed CNN-DFLC model, classification accuracy can be massively improved. The focus of this research work is better optimization of training weights of neural networks. The first step of plant identification and classification is the selection of the large dataset and the second step is pre-processing of dataset images available in different classes and performing tuning of hyper-parameters. The next step is an analysis of this plant dataset to get pre-trained weights. In the next step, the obtained pre-trained weights are utilized to perform efficient deep training. The final step is testing the proposed CNN-DFLC model based on the obtained fine-grained discriminative features and performing classification. The testing results will provide several performance metrics using the testing dataset and class prediction-related results. The proposed CNN-DFLC model efficiently estimates which image belongs to which plant class. So that efficient plant identification of different species can be achieved. Here, the real outputs were compared with the predicted outputs to detect errors. Moreover, individual and overall accuracy, precision, recall, and other performance metrics are measured to evaluate the efficiency of the proposed CNN-DFLC model. With the help of certain training parameters and optimizers, efficiency improvement is achieved. Finally, the successful classification and identification of plant species are achieved.

The proposed training framework consists of a plant image dataset with varied classes and these images are fed as input to the proposed CNN-DFLC model. The proposed CNN architecture consists of varied sequential layers, soft-max activation, and dense blocks. In addition, varied optimizers are utilized to improve performance and perform model fitting for plant detection and classification. Customization of the proposed CNN-DFLC model is achieved with the help of convolutional layers, max-pooling layers, batch normalization layers, dropout layers, and dense blocks. There are different stride sizes of convolutional layers, and max-pooling layers are utilized. Varied types of optimizers are adopted such as Adam, RMS-prop, and AMS-grad to improve analysis and classification efficiency. The visualization of performance metrics is analysed using the Loss curves, training accuracy, validation accuracy, and confusion matrix. Moreover, the best hyper-parameters for proposed CNN architecture are achieved using a cross-validation approach. Here Figure 1 provide details of plant classification process using the proposed CNN-DFLC model from the data acquisition stage to the final classification performance enhancement stage.

#### 3.1. Model pre-processing

The proposed CNN-DFLC model consists of varied layers such as sequential layers, dropout layers, max-pooling layers, fully linked layers, and soft-max layers. It also consists of a few dense blocks. In this work, Vietnam dataset is selected for the training of the proposed CNN-DFLC model. This dataset is a large plant image dataset that contains several images of 200 classes. Sometimes, noise or distortions are not visible or visualization is not possible from the naked eyes. Thus, in the proposed model, pre-processing is an essential step in which the dataset images are filtered from noise and unwanted distortions so that pre-trained features can be fine-grained. Generally, deep learning or CNN-based classification models require a large number of dataset images to avoid over-fitting. Therefore, the dataset images are transformed into varied shapes like

horizontal, rotational, vertical, and zooming of certain regions in different epochs for each image, and images can be transformed into several different orientations. The regions of each image are transformed into each step of an epoch. Therefore, all the regions of each image are covered and accurate training is performed. Most of the plant species are symmetric in nature, so more training images can be obtained by mirroring and rotating the given dataset images using transformation and augmentation methods. Moreover, histogram equalization improves contrast values and colour augmentation efficiency. All the training images must be of the same size for efficient network modelling. Padding and scaling can be performed to analyse images precisely as the images are gathered at varying heights and angles. Thus, after pre-processing, pre-trained features can be generated from the model analysis and efficient training can be performed. Furthermore, computational complexity reduction, dataset uniformity, image smoothening, and feature learning enhancement can be achieved using pre-processing in the proposed CNN-DFLC model.

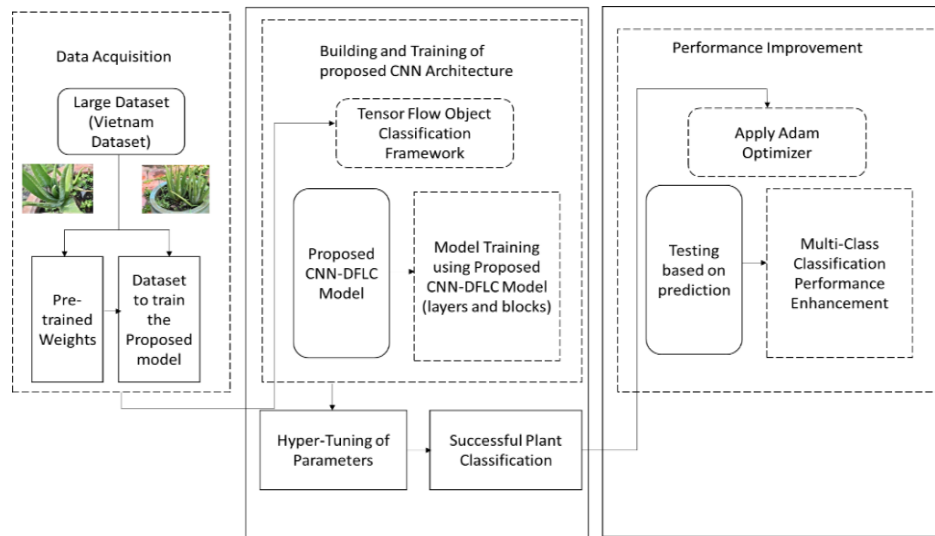


Figure 1. Plant classification process using proposed CNN-DFLC model

### 3.2. Model architecture

The main objective of the proposed CNN-DFLC model is to design an accurate learning and computationally compact model. The generated pre-trained features can be utilized for model training to get the efficient classification of plant species. Three sets of layers are presented in the proposed CNN-DFLC model. In the first set of varied convolutional layers, a batch normalization layer is present followed by a rectified linear units (ReLU) activation functional layer. In the second set, two different max-pooling layers, and the third set of layers consists of a soft-max layer, a classification layer, and a fully linked layer.

#### 3.2.1. Convolutional layers

Convolutional layers are the key building blocks of the proposed CNN architecture. The convolutional layers consist of several feature detectors, which are utilized to generate feature maps. These layers contain multiple filters like blur, sharpen, edge detect, edge enhancement, and emboss. The main objective of the proposed CNN-DFLC model is the extraction of unique fine-grained discriminative features. The size of convolutional filters is modified from a higher dimension convolutional filter to a smaller dimension convolutional filter and the number of filters is reduced to minimize computational complexity. The feature extraction from multiple convolutional filters is obtained using the (1):

$$L_j = \Psi(R_{j-1} * K_j + y_j) \quad (1)$$

where input image is given by  $K_j$  and features weights are expressed by  $R_j$ . Here, the ReLU activation function is represented by  $\Psi$  and  $y_j$  is the bias value. The output feature map is given by  $L_j$ . The convolutional operator is represented by an operator  $(*)$ . Each convolutional layer in the proposed CNN-DFLC model analyses different attributes or characteristics to gather discriminative fine-grained features from input images to differentiate between various classes of plant species. The training parameters are constantly updated in these layers and so the data distribution also updates regularly and feature weights vary for each image. Thus, this

parameter variation has a massive impact on the proposed CNN-DFLC model in terms of training speed. The reduction of filter size minimizes computational cost and generates quality weights. The overall loss in the feature extraction process is evaluated by (2):

$$K(j, e, t, i) = D^{-1} [K_{cnfl}(j, e) + R_j K_{lcz}(j, t, i)] \quad (2)$$

where  $K_{lcz}$  is represented as pixel localization loss in an input image,  $K_{cnfl}$  and  $R_j$  are expressed as validation loss and feature weights, respectively. The number of training iterations is given by  $D$ .

### 3.2.2. Batch normalization and ReLU activation functional layer

As discussed before, layers are updated regularly so the input to the layers can be changed. Thus, the batch normalization layer is employed for deep training of neural networks and used for the normalization of layer contributions in each mini-batch. The proposed CNN-DFLC model minimizes the number of training epochs. This layer is utilized to parametrize the proposed neural network model. Moreover, this layer minimizes the number of iterations used in training significantly without compromising performance efficiency. The batch normalization layer is employed to normalize outputs of a given layer in terms of standard deviation normalization.

$$a = [(c - b)(\varepsilon^2 + \lambda)^{1/2}] \cdot \beta + \mu \quad (3)$$

Where mean and standard deviation is given by  $b$  and  $\varepsilon$ , respectively for the present epoch  $c$ . Trainable parameters  $\beta$  and  $\mu$  get updated regularly after each epoch. A small constant is added to the variance and represented by  $\lambda$  so that zero-division could be avoided. Moreover, the mean and standard deviation are evaluated only for the training dataset, not for the testing dataset to avoid problems. Finally, average mean and standard deviation statistics are used in the training dataset. After the batch normalization layer, a ReLU activation layer is employed to enhance the nonlinearity of the proposed CNN-DFLC model or to improve non-linear decision boundaries so that over-fitting can be avoided. The ReLU activation layer is mostly utilized for object identification using deep learning and CNN models. Thus, training speed is enhanced to get better classification results. Then, the ReLU activation function is given by (4):

$$f(K_j) = \begin{cases} K_j, & \text{if } K_j > 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

in (5) can be rewritten as,

$$f(K_j) = \max(0, K_j) \quad (5)$$

then, the final representation of the ReLU activation function is given by (6).

$$act(K_j) = \max(0, [R_j \cdot K_j + y_j]) \quad (6)$$

The main objective of the ReLU activation function is to retain all the positive pixel values of the input image  $K_j$  and convert all the negative pixel values to zero. The input image is fed to the convolutional layers and the weights generated from the information related to the input image are utilized in terms of tensor values. The element-wise multiplication is performed between weighted kernels and input tensor values for each region of an image. Finally, all the output values are summed to obtain the final output tensor.

### 3.2.3. Pooling layers and drop out layers

Pooling layers are the most important part of the proposed CNN-DFLC model and these layers are mainly utilized to encode the dimensions and size of convoluted features. The height and width of feature maps are compressed while the number of channels remains constant. This layer is essential to minimize the required computational resources for an image processing approach. Pooling layers can be divided into two categories such as max pooling and average pooling. Max-pooling gives the maximum pixel values of an image whereas average pooling gives the average pixel values of an image. The pooling layer introduces translational invariance and reduces spatial resolution. This layer is employed for capturing different mean and max values within a particular image region from a convoluted image. Then, the output feature map is updated for the  $j^{th}$  pooling layer by (7).

$$L_{jwz} = \max_{g,h \in K_j} (N_{jgh}) \quad (7)$$

Where  $N_{jgh}$  represents elements of a particular region  $(g, h)$  of an image using the pooling layer and  $L_{jwz}$  is the output pooled feature map. Drop-out layers are utilized to improve the training capabilities of the proposed CNN-DFLC model and avoid over-fitting by pixel regularization and are also utilized for scaling. The proposed CNN-DFLC model supports multinomial probability distribution.

### 3.2.4. Flatten layers and fully linked layers

The flattened layers are utilized to obtain feature vectors from the pooled feature maps and fed as input to the fully linked layers. This layer adds an extra layer to the dimensions. Here, all the input layers are linked to the previous output's layers. Fully linked layers are employed to obtain classification features for the respective purpose. This layer maps obtained feature vectors to the predicted labels and a soft-max layer is also utilized as a classifier for multi-class classification and is used as the activation layer for the output. This layer predicts labels based on the obtained image attributes and features. The predicted labels  $q$  can be compared with the ground truth labels  $p$  to evaluate classification performance. So, the architecture of the proposed CNN-DFLC model is summarized as follows. First of all, specific features are obtained from an image using a convolutional layer and can be down-sampled using pooling layers. Then, the flattened layer can be utilized to obtain feature vectors and fed to the fully linked layers to get the final output. In (8) provides a distribution probability and the summation of the probability should be 1 and the class with the highest probability is considered as a final class for the respective image. Non-linear mapping is performed for all the nodes of fully linked layers and the probability distribution is given by (8):

$$g(L = z | \Phi^{(j)}) = \exp \Phi^{(j)} \cdot \left( \sum_{z=0}^w \exp \Phi_w^{(j)} \right)^{-1} \quad (8)$$

where  $g(L = z)$  is the probability of belonging to the  $z^{th}$  class among all the available  $w$  classes. Moreover, total training loss is evaluated by (9):

$$M(p, q) = \frac{1}{S} \sum_{o=1}^S (p_o - q_o)^2 \quad (9)$$

where  $M(p, q)$  is the square difference between ground truth labels and predicted labels and is termed as the loss function. The total number of training images is given by  $S$  and  $p_o$  represents ground truth labels and  $q_o$  represents the predicted class labels. Furthermore, categorical cross-validation and hyper-parameter tuning approach is adopted to obtain the best possible parameters so that maximum classification accuracy can be achieved. Certain optimizers are utilized to evaluate errors for forwarding propagation and fine-tune features of the proposed CNN-DFLC model such as learning rate and feature weights. These optimizers are utilized to reduce computational training loss. The optimizers can be of different types such as *RMSProp*, Adam, and *AMSGrad*. Here, the *RMSProp* optimizer is used for evaluating the dynamic learning rate whereas the Adam optimizer is employed which supports the properties of *RMSProp* optimizer and regulates the dynamic components like mean or learning rate with respect to dynamic mean squared gradients. The Adam optimizer is evaluated by (10) and (11):

$$L_{jwz}(u+1) = L_{jwz}(u) - \Psi \cdot v_u \quad (10)$$

where,

$$v_u = \Gamma v_{u-1} + (1 - \Gamma) \left[ \frac{\Delta(M(p, q))}{\Delta(L_{jwz}(u))} \right] \quad (11)$$

where aggregation of gradients at time  $u$  is given by  $v_u$  and aggregation of gradients at time  $u - 1$  is given by  $v_{u-1}$ , weights at time  $u$  and  $u + 1$  are represented by  $L_{jwz}(u)$  and  $L_{jwz}(u + 1)$ , respectively. Here,  $\Psi$  represents the learning rate and  $\Delta(M(p, q))$  shows loss function derivative and derivative of weights at time  $u$  are given by  $\Delta(L_{jwz}(u))$  and  $\Gamma$  is a moving average coefficient. Furthermore, *AMSGrad* optimizer is one of the variants of the Adam optimizer which is used to optimize the learning rate. In this way, a proposed CNN-DFLC model is designed to perform efficient classification and identify plant species accurately. Figure 2 demonstrates the design of the proposed CNN-DFLC model.

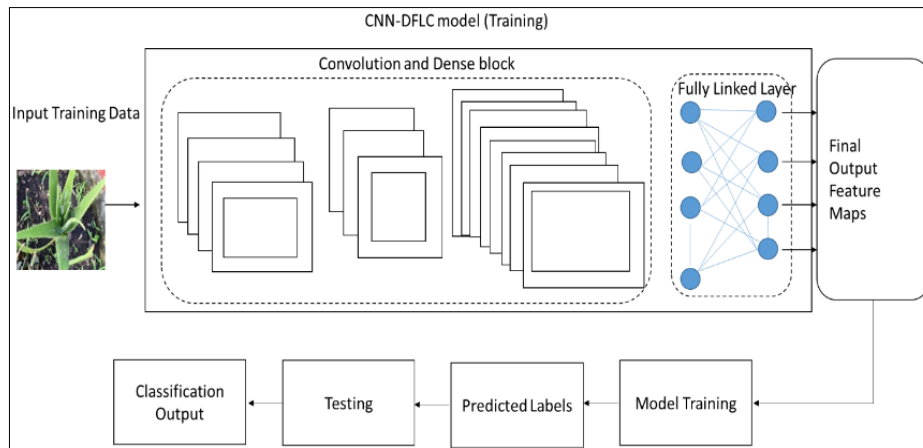


Figure 2. Design of proposed CNN-DFLC model

#### 4. RESULT AND DISCUSSION

This section demonstrates the performance results obtained from the deep analysis performed on the proposed CNN-DFLC model and compared against varied plant classification methods in terms of classification accuracy. Data misclassification, untrainable hyper-parameters, and inadequate training models can provide different challenges in performing accurate classification of the plant species. Therefore, a CNN-DFLC model is adopted in this research work to handle these mentioned problems and design an adequate and efficient training model to perform effective plant classification by analysing a large plant dataset. Although designing an efficient plant classification model is an extremely complicated task, especially for Vietnam plant dataset (VPN-200) [18] due to the presence of multiple leaves, flowers, stems together. Thus, the detection of the exact boundaries of leaves and distinguish between plant leaves and flowers is a complicated task. Therefore, an effective classification model based on CNN architecture is employed to perform adequate classification. The main base of the proposed classification process is an efficient architecture design that consists of multiple layers and blocks. Hence, CNN architecture can be segregated into two different blocks convolutional blocks and dense blocks. Inside these blocks, multiple players are present and each layer consists of different filters. These filters consist of multiple functions and packages and all these filters are assigned some specific tasks related to the plant classification. Those layers are the convolutional layer, pooling layer, ReLU activation functional layer, soft-max layer, and flatten layer and fully linked layers. From the generated feature maps in the training of the proposed CNN-DFLC model, the classification performance is observed by comparing predicted labels against ground truth labels. The testing results majorly depend upon the overall training performance to provide high classification accuracy and accurately predict which image belongs to which class among available numerous classes. However, multi-class classification is a complicated process and mainly depends upon predicted labels. Testing the given test dataset is an important step in a plant classification process. Testing is measured by different performance metrics and testing results are simulated using the trained model. Generate feature maps are a combination of feature weights obtained from each image. For every image, a ground truth label is assigned which is compared with their respective predicted label to get classification results.

##### 4.1. Dataset details

The training and testing performance of the proposed CNN-DFLC model is evaluated using VNP-200 dataset and compared against varied plant classification models in terms of classification accuracy. The VNP-200 dataset consists of a total number of 20,000 varied plant-related images. Moreover, the training, validation and testing ratio considered is 60:40 to measure classification performance i.e., total number of training images is 12,000 and testing images is nearly 8,000. The number of plant species present in this dataset is 200. These plant images are captured by an organization named as National Institute of Medicinal Materials and the plants are located in different nurseries in Vietnam City namely Ho Chi Minh City, Island Resort, Phu Tho City, and Ngoc Xanh. However, the conditions in which these plant images are captured can produce noise and illumination changes. As shown in Figure 3, some of the plant species are *Agave Americana*, *Alocasia macrorrhizos*, *Ampelopsis cantoniensis*, *Blackberry Lily*, *Bengal Arum*, *Breynia vitis*, *Citrus aurantifolia*, and *Curculigo gracilis*. All the images of 200 classes are selected for training and testing. The resolution of each image is 128×128 pixels. In each class, a different number of plant images are present. Fine-grained



discriminative features are obtained by analyzing plant images from these images to get classification performance. Due to the presence of different backgrounds, soil, tree bark, flowers, and several leaves together, noise can be present in the given images, which can be handled, in pre-processing stage using the proposed CNN-DFLC model.



Figure 3. An overview of the VPN-200 dataset

#### 4.2. Comparative analysis

A comparative analysis is performed in this section of multiple plant classification models against the proposed CNN-DFLC model in terms of classification accuracy. The proposed CNN-DFLC model is tested on

*Accurate plant species analysis for plant classification using convolutional neural ... (Savitha Patil)*



the VPN-200 dataset using different performance metrics like precision, recall, F1-score, and area under the curve (AUC). The proposed CNN-DFLC model focuses on achieving high classification accuracy with minimum computation cost and resources. Thus, fewer layers and blocks are used in the proposed CNN-DFLC model in comparison with the previous CNN classification models. Convolutional and pooling layers efficiently provide feature weights that can be utilized in the training of the model to generate feature maps and obtained feature maps are utilized for further testing of the model. The classification performance is evaluated by analysing confusion matrix results which are constructed using true positive, true negative, false positive, and false negative values. In other words, confusion matrix is a combination of two kinds of elements, which first discusses ground truth labels, and other shows predicted labels. Furthermore, a system with the configuration of an i7 processor, 16 GB RAM, 2 TB SSD+HDD, and GeForce RTX NITRO5 GPU memory is considered to perform all the plant classification experiments and simulation results. The performance of the proposed CNN-DFLC model is compared against varied CNN classification models such as VGG16 [19], InceptionV3 [20], MobileNet V2 [21], ResNet 50 [22], DenseNet 121 [23], and Xception [24]. Here, VGG16 is a deep neural network architecture that is designed using several convolutional and fully connected layers to analyze large datasets using small inception filters. Moreover, InceptionV3 is a combination of multiple local structures with varied sizes of convolutional operators. It is a multi-scale presentation and can be extended to generate pre-trained parameters. Here, MobileNet V2 artificial intelligence (AI) based is the built-in mobile device to compute high computation through mobile devices. Then, ResNet 50 is a mapping function used to optimize references to the multiple layers and restores the channel depth. Next, DenseNet 121 is a visual object detection model using dense block transition layers. Finally, Xception utilizes depth-wise separable convolutions to reduce inception module utilization. However, propose CNN-DFLC model is an efficient object classification model with minimum computational resource utilization. Table 1 represents simulation results for all 200 classes in terms of mean classification accuracy. The mean accuracy achieved using the proposed CNN-DFLC model is 96.42% considering all 200 classes. The highest previous accuracy achieved for the VPN-200 dataset considering all 200 classes is 88.26% and the model was Xception. So, the percentage increment of mean accuracy considering all 200 classes against VGG16 is 27%, InceptionV3 is 17%, MobileNet V2 is 10%, ResNet 50 is 10%, DenseNet 121 is 10%, and Xception is 9%. This shows the proposed CNN-DFLC model outperforms existing CNN plant classification modes and claims the highest performance than any other state-of-art classification model considering the VPN-200 dataset. The proposed CN [25], and F1-measure for all 200 classes.

Table 1. Classification performance results

Different CNN classification models	Classification accuracy
VGG16 [17]	76.00
InceptionV3 [18]	82.50
MobileNet V2 [19]	87.92
ResNet 50 [20]	88.00
DenseNet 121 [21]	88.00
Xception [22]	88.26
CNN-DFLC	96.42

Here, Figure 4 shows a graphical representation of performance metrics like validation accuracy and testing accuracy considering validation and testing data, respectively for varied CNN classification models such as InceptionResnet-2, InceptionV3, MobileNet V2, ResNet 50, GoogleNet, and Xception against the proposed CNN-DFLC model. Testing accuracy is denoted by green lines whereas validation accuracy is denoted by blue lines. Here, the number of epochs is considered 100 and the number of steps is 250. This shows each image is transformed or flipped with multiple orientations or angles and processed in training which means each image is processed multiple times so most of the essential pixels are trained. These graphs show that the testing results are slightly better than the validation metrics results. The previous best CNN classification model was exception net with 91.8% testing accuracy whereas the second-best CNN method was Inception ResNetV2 with 91.2% testing accuracy. The proposed CNN-DFLC model outperforms traditional CNN classification models with a testing accuracy of 96.42%. Here, Figure 5 shows a graphical representation of improvement in classification accuracy using proposed CNN-DFLC model against varied ensemble models such as mean ensemble, voting ensemble, weighted mean ensemble, and stacking ensemble. The percentage improvement in classification accuracy for mean ensemble is 4.1%, voting ensemble is 3.67%, the weighted mean ensemble is 4.24%, the stacking ensemble is 2.14% and the proposed CNN-DFLC model is 5%. These improvements are observed while keeping the individual best ensemble model as a reference with 91.80% classification accuracy. These graphs show that the classification improvement is slightly better than the varied

ensemble results. The previous best classification improvement is observed in the weighted mean ensemble model. The proposed CNN-DFLC model outperforms varied ensemble models in terms of classification accuracy improvement as well.

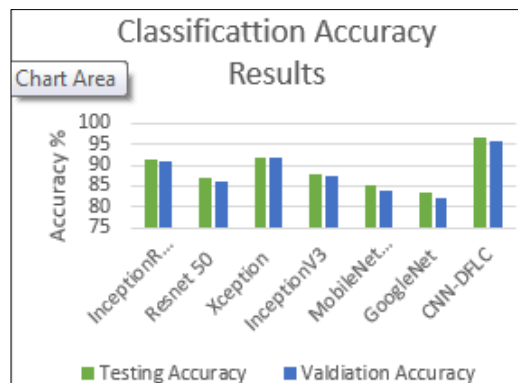


Figure 4. Classification accuracy for validation and test datasets for VP-200 dataset

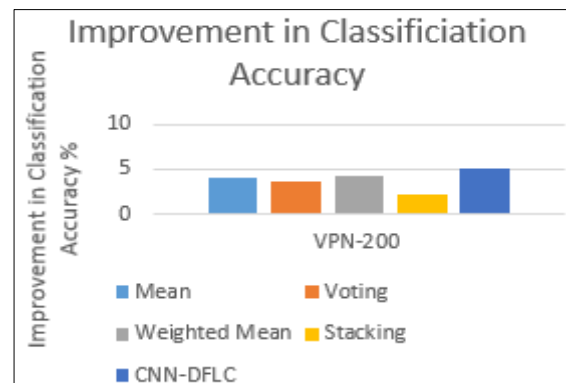


Figure 5. Percentage of improvement in classification accuracy proposed CNN-DFLC model against individual best classification ensemble model

## 5. CONCLUSION

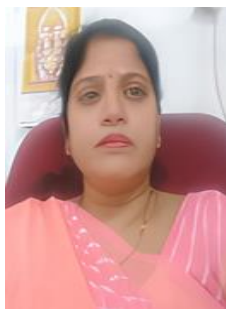
Plant classification is an interesting and challenging research area due to the presence of numerous plant species across the world, the same green color of the leaves in a maximum number of plants, the presence of flowers, and the presence of multiple leaves together. Thus, a CNN-DFLC model is proposed to analyze plant classification and detect plant species accurately by countering these challenges. The main objective of this work is the plant species identification of which plant image belongs to which class. The proposed CNN-DFLC model is constructed using several layers and blocks like convolutional layer, pooling layer, ReLU activation functional layer, soft-max layer, flatten layer, and fully linked layers. The proposed CNN-DFLC model is performed in different stages such as the data selection stage, data pre-processing stage, feature generation stage, training stage, and testing stage. Moreover, a comprehensive analysis is performed to understand specific parameters to enhance training and testing efficiency and capture fine-grained feature weights. Then, those obtained feature weights are utilized in the proposed CNN-DFLC model to get the maximum yield. A deep mathematical analysis of CNN architecture is also presented. The performance of the proposed CNN-DFLC model is tested on the Vietnam plant (VPN-200) dataset, which contains 200 plant species images. Performance is measured using the proposed CNN-DFLC model in terms of classification accuracy, precision, recall, and F1 score. The proposed CNN-DFLC model is compared against varied traditional CNN plant classification models in terms of classification accuracy. Mean classification accuracy is 96.42%, mean precision is 95.56%, mean sensitivity is 93.58%, mean specificity is 99.98%, and mean F1 measure is 94.23%. The model accurately detects which particular image belongs to which species. Thus, the proposed CNN-DFLC model shows decent performance against different traditional classification models.




## REFERENCES

- [1] J. Bruinsma, "The resource outlook to 2050: by how much do land, water and crop yields need to increase by 2050?," *FAO Expert Meeting on How to Feed the World in 2050*, 2009, [Online]. Available: <https://ftp.fao.org/docrep/fao/012/ak971e/ak971e00.pdf>
- [2] J. Ma, K. Du, F. Zheng, L. Zhang, Z. Gong, and Z. Sun, "A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network," *Computers and Electronics in Agriculture*, vol. 154, pp. 18–24, Nov. 2018, doi: 10.1016/j.compag.2018.08.048.
- [3] F. O. Faithpraise, P. Birch, R. C. D. Young, J. Obu, B. Faithpraise, and C. R. Chatwin, "Automatic plant pest detection and recognition using k-means clustering algorithm and corespondence filters," *International Journal of Advanced Biotechnology and Research*, vol. 4, no. 2, pp. 1052–1062, 2013.
- [4] X.-S. Wei *et al.*, "Fine-grained image analysis with deep learning: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 8927–8948, Dec. 2022, doi: 10.1109/TPAMI.2021.3126648.
- [5] J. Yin, A. Wu, and W.-S. Zheng, "Fine-grained person re-identification," *International Journal of Computer Vision*, vol. 128, no. 6, pp. 1654–1672, Jun. 2020, doi: 10.1007/s11263-019-01259-0.
- [6] S. D. Khan and H. Ullah, "A survey of advances in vision-based vehicle re-identification," *Computer Vision and Image Understanding*, vol. 182, pp. 50–63, May 2019, doi: 10.1016/j.cviu.2019.03.001.
- [7] X.-S. Wei, Q. Cui, L. Yang, P. Wang, L. Liu, and J. Yang, "RPC: a large-scale and fine-grained retail product checkout dataset," *Science China Information Sciences*, vol. 65, no. 9, Sep. 2022, doi: 10.1007/s11432-022-3513-y.




- [8] L. Li, S. Zhang, and B. Wang, "Plant disease detection and classification by deep learning-a review," *IEEE Access*, vol. 9, pp. 56683–56698, 2021, doi: 10.1109/ACCESS.2021.3069646.
- [9] A. Bakhshipour, "Cascading feature filtering and boosting algorithm for plant type classification based on image features," *IEEE Access*, vol. 9, pp. 82021–82030, 2021, doi: 10.1109/ACCESS.2021.3086269.
- [10] W. Albattah, M. Nawaz, A. Javed, M. Masood, and S. Albahli, "A novel deep learning method for detection and classification of plant diseases," *Complex & Intelligent Systems*, vol. 8, no. 1, pp. 507–524, Feb. 2022, doi: 10.1007/s40747-021-00536-1.
- [11] S. Mathulapragasan and K. Lanthong, "Cassava leaf disease recognition using convolutional neural networks," in *2021 9th International Conference on Orange Technology (ICOT)*, Dec. 2021, pp. 1–5. doi: 10.1109/ICOT54518.2021.9680655.
- [12] C. Zhou, S. Zhou, J. Xing, and J. Song, "Tomato leaf disease identification by restructured deep residual dense network," *IEEE Access*, vol. 9, pp. 28822–28831, 2021, doi: 10.1109/ACCESS.2021.3058947.
- [13] W. Haider, A.-U. Rehman, N. M. Durrani, and S. U. Rehman, "A generic approach for wheat disease classification and verification using expert opinion for knowledge-based decisions," *IEEE Access*, vol. 9, pp. 31104–31129, 2021, doi: 10.1109/ACCESS.2021.3058582.
- [14] X. Jin, J. Che, and Y. Chen, "Weed identification using deep learning and image processing in vegetable plantation," *IEEE Access*, vol. 9, pp. 10940–10950, 2021, doi: 10.1109/ACCESS.2021.3050296.
- [15] S. S. Chouhan, A. Kaul, U. P. Singh, and S. Jain, "Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology," *IEEE Access*, vol. 6, pp. 8852–8863, 2018, doi: 10.1109/ACCESS.2018.2800685.
- [16] L. Liu, W. Min, S. Mei, L. Wang, and S. Jiang, "Plant disease recognition: a large-scale benchmark dataset and a visual region and loss reweighting approach," *IEEE Transactions on Image Processing*, vol. 30, pp. 2003–2015, 2021, doi: 10.1109/TIP.2021.3049334.
- [17] C. Zhou, Z. Zhang, S. Zhou, J. Xing, Q. Wu, and J. Song, "Grape leaf spot identification under limited samples by fine grained-GAN," *IEEE Access*, vol. 9, pp. 100480–100489, 2021, doi: 10.1109/ACCESS.2021.3097050.
- [18] T. N. Quoc and V. T. Hoang, "VNPlant-200-a public and large-scale of Vietnamese medicinal plant images dataset," in *ICIS 2020: Integrated Science in Digital Age 2020*, 2021, pp. 406–411. doi: 10.1007/978-3-030-49264-9\_37.
- [19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Prepr. arXiv.1409.1556*, Sep. 2014, [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, Feb. 2017, doi: 10.1609/aaai.v31i1.11231.
- [21] A. G. Howard *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," *Prepr. arXiv.1704.04861*, Apr. 2017, [Online]. Available: <http://arxiv.org/abs/1704.04861>
- [22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [23] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.
- [24] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.
- [25] O. A. Malik, M. Faisal, and B. R. Hussein, "Ensemble deep learning models for fine-grained plant species identification," in *2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, Dec. 2021, pp. 1–6. doi: 10.1109/CSDE53843.2021.9718387.

## BIOGRAPHIES OF AUTHORS



**Savitha Patil**    completed B.E. in information science and engineering from Visvesvaraya Technological University in the year 2008 and obtained M.Tech. in computer science and engineering in the year 2011. She worked as a lecturer in the Appa Institute of Engineering and Technology from 2008 to 2010. Presently, she working as an assistant professor at Sharnbasva University in the Department of Computer Science and Engineering since 2012 to till date. She has teaching experience of about 12 years. She can be contacted at email: SavithaPatil123456789@gmail.com.



**Mungamuri Sasikala**    completed B.E. in electrical engineering from Osmania University Hyderabad in the year 1985 with first class distinction. She obtained M.E. in the year 1987, in power systems from Osmania University and was awarded a gold medal for standing first in the university. She completed her Ph.D. from JNTU Hyderabad in electrical engineering in the year 2008. She working as a professor/principal in various colleges since 2008 and currently working as the principal of Godutai Engineering College for Women, Sharnbasva University, Kalaburagi since 2014. She has 3 decades (30 years) of teaching experience. She can be contacted at email: sasi\_mum@rediffmail.com.