Ecological Informatics 67 (2022) 101492

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Performance analysis of real-time plant species recognition using bilateral network combined with machine learning classifier

S. Anubha Pearline [a,*], V. Sathiesh Kumar [b]

[a] Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani Campus, Chennai 600026, India
[b] Department of Electronics Engineering, Madras Institute of Technology Campus, Anna University, Chennai 600044, India

ARTICLE INFO

Keywords:
Plant species recognition
Bilateral architectures
Convolutional neural network
Dual architectures
Transfer learning

ABSTRACT

A real-time plant species recognition under an unconstrained environment is a challenging and time-consuming process. The recognition model should cope up with the computer vision challenges such as scale variations, illumination changes, camera viewpoint or object orientation changes, cluttered backgrounds and structure of leaf (simple or compound). In this paper, a bilateral convolutional neural network (CNN) with machine learning classifiers are investigated in relation to the real-time implementation of plant species recognition. The CNN models considered are MobileNet, Xception and DenseNet-121. In the bilateral CNNs (Homogeneous/Heterogeneous type), the models are connected using the cascade early fusion strategy. The Bilateral CNN is used in the process of feature extraction. Then, the extracted features are classified using different machine learning classifiers such as Linear Discriminant Analysis (LDA), multinomial Logistic Regression (MLR), Naïve Bayes (NB), k-Nearest Neighbor (k−NN), Classification and Regression Tree (CART), Random Forest Classifier (RF), Bagging Classifier (BC), Multi-Layer Perceptron (MLP) and Support Vector Machine (SVM). From the experimental investigation, it is observed that the multinomial Logistic Regression classifier performed better compared to other classifiers, irrespective of the bilateral CNN models (Homogeneous - MoMoNet, XXNet, DeDeNet; Heterogeneous - MoXNet, XDeNet, MoDeNet). It is also observed that the MoDeNet + MLR model attained the state-of-the-art results (Flavia: 98.71%, Folio: 96.38%, Swedish Leaf: 99.41%, custom created Leaf-12: 99.39%), irrespective of the dataset. The number of misprediction/class is highly reduced by utilizing the MoDeNet + MLR model for real-time plant species recognition.

1. Introduction

The primary aim of the plant species recognition system is to assist the non-botanists and non-experts in recognizing the plant species (Kaya et al., 2019a). Manual methods for identifying the plant species are complex and laborious. Furthermore, the plant species which are endangered needs to be recognized to preserve them from extinction. Hence, it becomes necessary to develop a computerized system for recognizing the plant species. The common mode of computerized plant species recognition is by examining the features of leaves (Geometry, Color, Texture, Vein pattern, leaf arrangement on the branches etc.). But, these characteristics may differ due to large diversity (inter-class or intra-class) in plant species, seasonal changes and aging. Also, the computerized plant species recognition system has to address the issues related to variation in ambient lighting conditions, cluttered background, viewpoint or orientation variations and scale changes.

With the development and progress of technology, the conventional machine learning domain is preferred in the classification and recognition tasks. Machine learning methods like k-Nearest Neighbor (Mostajer Kheirkhah and Asghari, 2019), Support Vector Machine (Ahmed and Hussein, 2020; Wang et al., 2020), Multi-Layer Perceptron (Kumar et al., 2019), Random Forest (Anubha Pearline et al., 2019) and other classifiers are employed in the process of plant species recognition. Deep learning-based computer vision applications mainly focused on Convolutional Neural Network (CNNs) (Guo et al., 2016) for classification and object recognition tasks. CNNs are capable of extracting detailed information from the input data compared to the conventional machine learning techniques. Numerous research activities have been reported on plant species recognition using CNN models. Some of the utilized CNN architectures are AlexNet (Kaya et al., 2019a), VGG-16 (Kaya et al., 2019a), Inception-V3 (Anubha Pearline et al., 2019), Inception ResNet-V2 (Anubha Pearline et al., 2019), LeafNet (Barré et al., 2017) and

* Corresponding author.
E-mail addresses: anubhaps@srmist.edu.in (S. Anubha Pearline), sathieshkumar@annauniv.edu (V. Sathiesh Kumar).

https://doi.org/10.1016/j.ecoinf.2021.101492
Received 7 August 2021; Received in revised form 15 November 2021; Accepted 20 November 2021
Available online 26 November 2021
1574-9541/

Table 1
Existing bilateral architectures for plant species recognition.

Method	Deep Learning CNN used	Early Fusion/Late Fusion/Hierarchical	Dataset	Transfer Learning (✓/✘)	Data Augmentation (✓/✘)	Author & Year
3-EnsCNNs and 5-EnsCNNs	Xception, MobileNet-V1, MobileNet-V2, DenseNet-121, NASNetMobile	Late Fusion	Mulberry leaf, tomato leaf and corn leaf datasets	✓	✓	Chompookham and Surinta (2021)
Dual-path CNN	Custom CNN	Early Fusion	Flavia	✘	✓	Rizk (2019)
MMCNN	EfficientNet and MobileNet	Early Fusion	Flavia, Folio and LeafSnap	✓	✓	Dat et al. (2021)
Siamese-Inception	Inception	Early Fusion	LeafSnap, Flavia and Swedish leaf	–	✘	Wang and Wang (2019)
TA-CNN	Xception CNN	Hierarchical	ICL, Flowers-102, CFH and Malayakew	✓	✓	Zhu et al. (2019)
HGO-CNN and Plant-StructNet Ensemble		Late Fusion	PlantCLEF2015	✓	✓	Lee et al. (2018)
Bi-channel Framework	VGG-16 and SqueezeNet	Late Fusion	Orchidaceae plant family	✓	✘	He et al. (2018)
MSF-CNN	Custom CNN	Hierarchical and early fusion	Malayakew and LeafSnap	✘	✓	Hu et al. (2018)
Hierarchical Classification (GoogleNet)	GoogleNet	Hierarchical and late fusion	ImageCLEF 2015 plant dataset	✓	✓	Araújo et al. (2018)
Hybrid global-local feature extraction (AlexNet)	AlexNet	Early Fusion (cascade, conv-sum) and Late Fusion	Malayakew	✓	✓	Lee et al. (2017)
GoogleNet + VGGNet	Combination of GoogleNet and VGGNet	Late Fusion	LifeCLEF2015	✓	✓	Ghazi et al. (2017)
Dual-path deep CNN	Custom CNNs	Early Fusion	Flavia and ImageCLEF2015	✘	✓	Shah et al. (2017)

ResNet (Sun et al., 2017).

2. Review of literature

In recent years, the Bilateral CNNs are widely used in different applications such as aerial scene classification (Dede et al., 2019), geospatial land classification (Minetto et al., 2019), smoke detection (Pundir and Raman, 2019), medical image segmentation (Wang et al., 2021) and sentimental analysis (Rao et al., 2021). The two CNNs are connected either in parallel or in a hierarchical manner to form the bilateral CNN. Few literatures utilizing the bilateral CNN in plant species recognition are described below.

Chompookham and Surinta (Chompookham and Surinta, 2021) proposed 3-EnsCNNs and 5-EnsCNNs ensemble methods of deep Convolutional Neural Network for recognizing the plant species. 3-EnsCNNs involves Xception, MobileNet-V2 and DenseNet-121. While, 5-EnsCNNs include MobileNet-V1, MobileNet-V2, NASNetMobile, DenseNet-121 and Xception. The individual CNNs of 3-EnsCNNs and 5-EnsCNNs are combined after passing each of the CNN's softmax layer using unweighted majority vote, unweighted average, and weighted average methods. The authors obtained the optimal results for weighted average method. The authors reported the results of ensemble methods on three datasets, namely, mulberry leaf (94.75%-for 5-EnsCNNs), tomato leaf (99.93%- for 3-EnsCNNs) and corn leaf (99.47%- for 3-EnsCNNs) datasets.

Rizk (Rizk, 2019) implemented a dual Convolutional Neural Network (CNN) for the classification of plant species. In the first path, the shape features are extracted from the thresholded binary images. Venation features are extracted in the second path using the leaf image patches generated by employing the Sobel operator. Dual-path CNN includes a simple CNN with four sets of convolution layer, Rectified Linear Unit (ReLU) and max-pooling layer. The output feature vector from the individual CNN path is concatenated to form the final output feature vector. Then, the extracted features are classified using the fully connected layer. The reported method is evaluated on the augmented Flavia dataset.

Dat et al. (Dat et al., 2021) proposed a multi-model Convolutional Neural Network (MMCNN) for plant species recognition. It combines the EfficientNet and MobileNet models. The authors performed leaf segmentation using U-Net architecture. MMCNN utilizes the joint learning multi-loss model for each CNN. The authors utilized four datasets namely, custom Vietnamese herbal leaf species dataset, Flavia, Folio and LeafSnap. Also, the authors reported that the MMCNN model is limited only to work under illumination variations.

Wang et al. (Wang and Wang, 2019) presented a Siamese-Inception (S-Inception) network for plant species recognition. It is based on a few-shot learning strategy. An Inception network is consists of two-way parallel CNN for feature extraction. Spatial Structure Optimizer (SSO) metric is proposed to improve the classification accuracy. k-Nearest Neighbor classifier is employed to recognize the plant species. S-Inception methodology is tested on LeafSnap, Flavia and Swedish leaf datasets. The authors reported accuracies of 91.75% (LeafSnap), 95.32% (Flavia) and 91.37% (Swedish leaf).

Zhu et al. (Zhu et al., 2019) introduced a Two-way Attention model (TA-CNN) focusing on the recognition of plant species. It utilizes the late fusion method. The individual CNN branches are used in the prediction of plant species and its family. The second CNN branch utilizes a max-sum attention network to generate a heat map. The Xception network is used as a backbone architecture in TA-CNN. The efficiency of the TA-CNN is assessed with the help of ICL, Malayakew, Flowers-102 and CFH augmented datasets. The authors reported accuracies of 99.8% (Malayakew), 99.9% (ICL), 97.2% (Flowers-102) and 79.5% (CFH).

Hybrid-Generic Convolutional Neural Network (HGO-CNN) (Lee et al., 2018) and Plant-StructNet architectures are ensembled and utilized in the process of plant species recognition. From the single plant image, the extracted organ and generic knowledge are fused with each other in the HGO-CNN model. Plant-StructNet model is made up of Recurrent Neural Network (RNN). The authors tested the methodology on PlantCLEF 2015 dataset and reported an accuracy of 68.5%.

He et al. (He et al., 2018) introduced a bi-channel deep learning system towards the classification of plant species from the Orchidaceae plant family. The reported Bi-channel system comprises of two pre-trained fine-tuned CNNs, namely, VGG-16 and SqueezeNet architectures. A stacking layer is used to integrate the outputs obtained from the two pre-trained CNNs. For the Orchidaceae plant dataset, this unified system resulted in an enhanced performance of 96.81%.

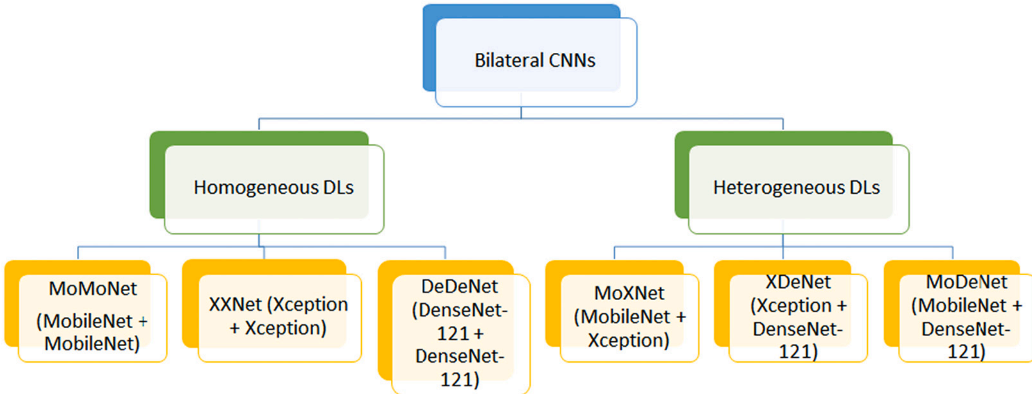

Fig. 1. Different combinations of bilateral CNNs.

Hu et al. *(*Hu et al., *2018)* designed a Multi-Scale Fusion Convolutional Neural Network (MSF-CNN) to perform plant species recognition. MSF-CNN architecture utilizes a downsampled image at multiple scales ([256 × 256], [128 × 128], [64 × 64] and [32 × 32]). The bilinear interpolation technique is used in the process of downsampling. The MSF-CNN model consists of two sets of Convolution layer and Max-pooling layer along with an average pooling layer. The model extracts multiple features by providing the input with multiple scales. Later, the outputs from the model are concatenated and utilized for the prediction of plant species. The authors reported accuracies of 97.35% on MalayaKew leaf D3 and 85.28% on LeafSnap datasets.

Araujo et al. (Araujo et al., 2018) aggregated the CNNs in a two-level hierarchical approach (GoogleNet CNNs) to implement a plant species recognition system. The two-level approach enabled the coarse-to-fine grained classification process. The complete image and image patches are given as inputs to the model towards the prediction of plant Genus. The author reported an accuracy of 86.44% using the ImageCLEF 2015 plant dataset.

Lee et al. (Lee et al., 2017) proposed a hybrid global-local feature extraction method in the context of plant species recognition. Features are extracted by utilizing two CNNs models. The leaf venation and leaf patches are provided as inputs to each of the CNN architectures. Three approaches for combining the architectures are reported. They are Cascade, Conv-sum and Late fusion approaches. The author reported that the Conv-sum approach (Accuracy = 96.3%) resulted in a better performance metric compared to Cascade (Accuracy = 95.5%) and Late fusion (Accuracy = 94.5%) methodologies.

Ghazi et al. (Mehdipour Ghazi et al., 2017) merged VGG and GoogleNet architectures to perform plant species recognition. The authors reported the performance metric obtained by utilizing individual pre-trained CNNs such as VGGNet (Accuracy = 71.24%), GoogleNet (Accuracy = 69.44%) and AlexNet (Accuracy = 49.63%). The fused CNN architecture resulted in an accuracy of 80% with the LifeCLEF2015 plant dataset.

Shah et al. (Shah et al., 2018) recommended a dual-path CNN towards the recognition of plant species. The individual paths are provided with a complete leaf image and leaf patches. The shape and texture features are extracted by the model. The output feature vector from the individual CNN paths is combined to form a final feature vector. Later, it is classified using the fully connected layer. The efficacy of the method is studied using three datasets namely, LeafSnap (Accuracy = 95.61%), Flavia (Accuracy = 99.28%) and ImageCLEF2015 (Accuracy = 96.42%).

From Table 1, it is noticed that the ensemble of deep learning models utilized for plant species recognition are Inception, MobileNet, GoogleNet, VGG-16, NASNetMobile, AlexNet and SqueezeNet. Each of the utilized architecture has its own characteristics and importance in relation to plant species recognition. Most of the reported architectures suffer from the class imbalance problem when it is exposed to real time data's. Also, the performance metrics obtained by utilizing the bilateral network could be further enhanced with the aid of combining advanced deep learning architectures such as Xception, DenseNet-121 and MobileNet.

The datasets (LeafSnap, Malayakew, Flowers-102, LifeCLEF2015, Flavia, Swedish leaf etc.) utilized in the reported literatures are heavily pre-processed and did not incorporate all the challenges (scale variation, viewpoint or orientation changes, different background, lighting changes, simple or compound leaf structures) in it. Hence, there is a requirement of creating a new dataset incorporating all the above specified challenges. Also, most of the reported literatures utilize the data augmentation method to increase the number of sample images in the considered dataset. As a result, there is an improvement in the performance metrics. The necessity of data augmentation is mitigated by incorporating the above specified challenges in the dataset itself.

In most of the reported bilateral networks, the complete network is utilized in the process of feature extraction and classification task. It results in larger computational time. The computation time is significantly reduced by incorporating the machine learning classifiers (classification task) with pre-trained CNN models (Feature extraction).

The major contributions of the paper are as follows:

1. The advanced CNN architectures such as Xception, DenseNet-121 and MobileNet are combined to form bilateral networks. These networks are investigated in the context of real-time plant species recognition. The class imbalance problem during the classification of real-time data is mitigated by utilizing the above specified architectures in bilateral mode.
2. A dataset named, Leaf-12 is custom developed incorporating the challenges (camera viewpoint or orientation changes, scale modifications, simple and complex structure of leaf, illumination variations, cluttered backgrounds etc.) encountered by the computer vision models. The requirement of data augmentation by the computer vision model is nullified by incorporating the above specified challenges during the process of image acquisition.
3. The inference time on real-time data is significantly reduced by utilizing the bilateral network (feature extraction process) with machine learning classifiers (classification task).
4. The optimal machine learning classifier is identified to be used in conjunction with the bilateral network. This is done to improve the performance metrics with lower computation time.

In this paper, a Bilateral Convolutional Neural Network (CNN) is formulated for plant species recognition. This is done to reduce the number of mispredictions on real-time plant images. The aggregation of two CNNs is referred to as Bilateral CNNs. MobileNet, Xception and DenseNet-121 CNN models are adopted for building the bilateral CNN

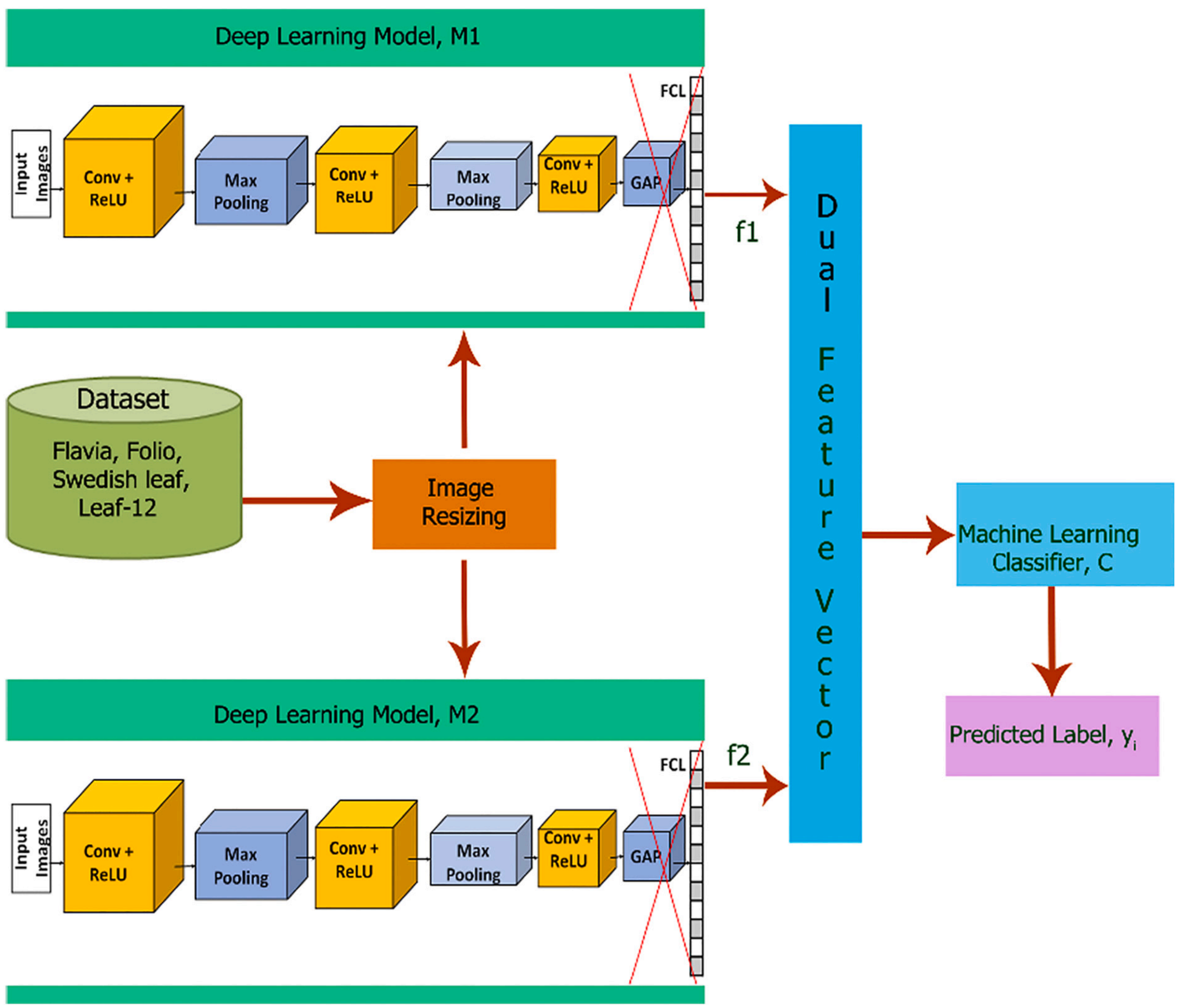

Fig. 2. Proposed bilateral convolutional neural network for the recognition of plant species.

architectures. The Bilateral CNN architecture is used as a feature extractor. Later, the extracted features are classified using Multinomial Logistic Regression (MLR), Linear Discriminant Analysis (LDA), Naive Bayes (NB), Classification and Regression Tree (CART), k-Nearest Neighbor (k-NN), Bagging Classifier (BC), Random Forest (RF), Multi-layer Perceptron (MLP) and Support Vector Machine (SVM). Six different combinations of architectures are considered in this study. The Bilateral CNNs are classified into Homogenous DLs and Heterogenous DLs. Different combinations of Bilateral CNNs are represented in Fig. 1.

3. CNN architectures

In recent years, the pre-trained CNN models are widely used in different applications (skin lesion (Yildirim-Yayilgan et al., 2021), sign language classification (Barbhuiya et al., 2021), Covid-19 detection (Hira et al., 2021), real-time anomaly detection (Ullah et al., 2021), etc.) and produces state-of-the-art results. Among different pre-trained CNN models, MobileNet, Xception and DenseNet-121 architectures are selected to form a bilateral CNN. The architecture selection is based on the significant properties of the model such as lightweight, residual connection and high recognition rates. The significant properties of the above-specified CNN architectures are discussed in the following subsections.

3.1. MobileNet

MobileNet (Howard et al., 2017) is one of the CNN models developed to facilitate the rapid execution of vision-based applications on mobile and embedded devices. MobileNet is a lightweight and portable CNN model. It uses a Depthwise Separable Convolutions (DSC) block to build a lightweight (less number of parameters to be trained) deep neural network model. The DSC block consists of Depthwise Convolution (DC) followed by a Pointwise Convolution (PW). DC produces three convoluted outputs for each color channel of the image. PW is a 1 × 1 convolution method, used to reduce the number of convolution operations. DSC minimizes the overfitting of the network and decreases the model complexity.

3.2. Xception

Xception (Chollet, 2017) is an updated version of the Inception CNN model. The model employs a modified depthwise separable convolution process (pointwise convolution followed by a depthwise convolution). It also makes use of residual connections which results in minimizing the vanishing gradient problem for very large networks. Also, the involvement of residual connections encourages the multiple feature gathering from the previous layers. The number of parameters to be trained by the model is larger compared to the MobileNet model.

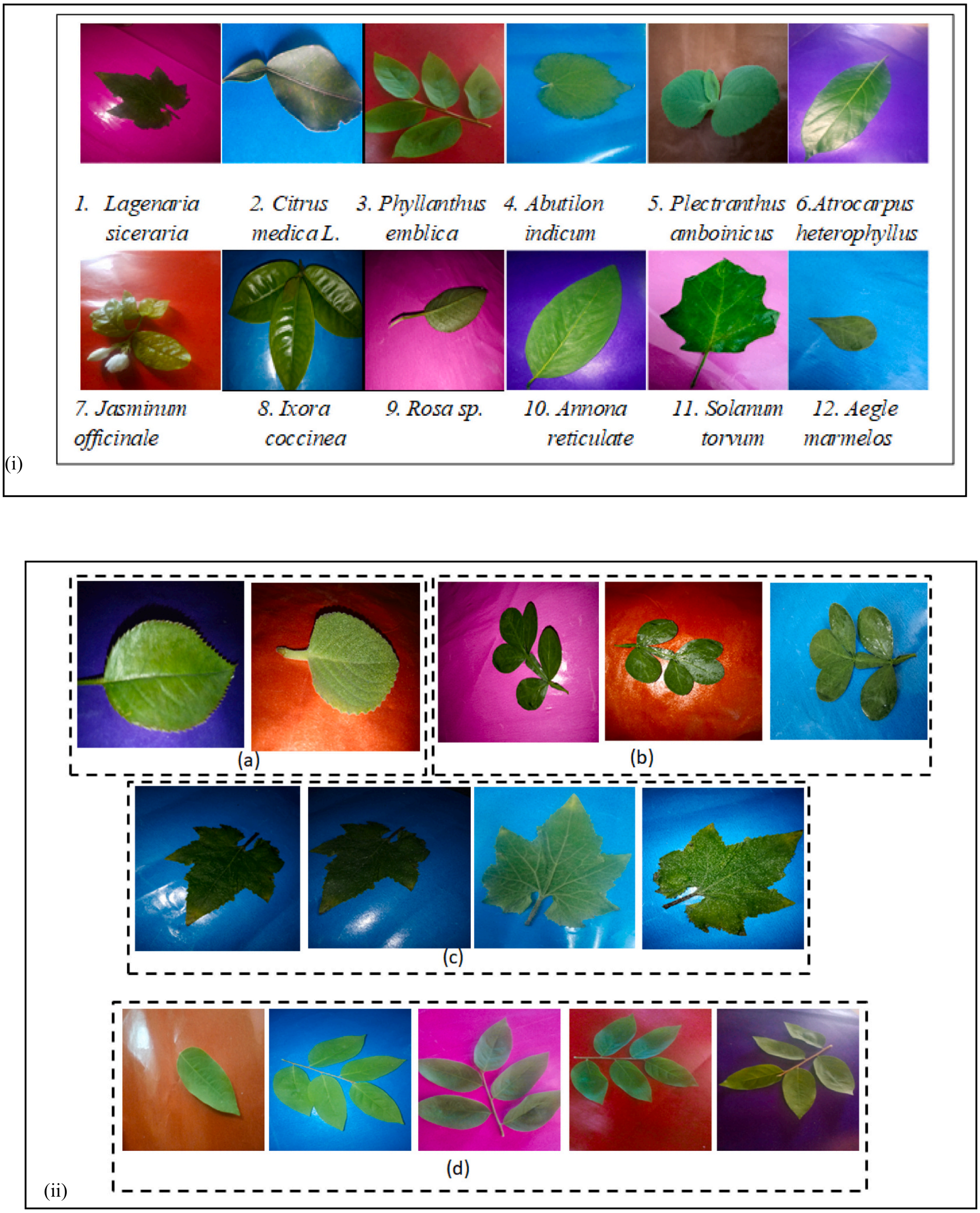

Fig. 3. (i) Sample leaf images of Custom created Leaf-12. 3(ii) Variation ((a) Inter-similar leaves, (b) Orientation, (c) Illumination, (d) Color backgrounds, Simple leaf and compound leaves) of leaf images in Leaf-12 dataset.

3.3. DenseNet-121

DenseNet-121 (Huang et al., 2017) consists of 121 trainable weight layers. Deep networks encounter vanishing input information (feedforward phase) issue and vanishing gradient problem (Backpropagation phase). The issues specified above are minimized in DenseNet models by ensuring the maximum information (feedforward phase) and gradient (Backpropagation) flow. It is accomplished by connecting every layer directly with each other. It results in minimizing the required number of trainable parameters by the network when compared to Residual networks. The model comprises dense blocks and transition blocks. The transition blocks are sandwiched between the dense blocks. In dense blocks, the dimension of feature maps remains constant within a block although the number of filters is varied. The transition blocks are used to downsample the input to that layer by applying the batch normalization, a 1×1 convolution and a 2×2 average pooling layers. As the feature maps from each layer are concatenated with each other, the channel dimension gets increased as the network deepens. Hence, the model performance depends on the acquired collective information from all the layers of the network.

4. Bilateral CNN for plant species recognition

The proposed Bilateral Convolutional Neural Network to perform the recognition of plant species is shown in Fig. 2. It consists of two deep learning models (M1 and M2) connected using the cascade early fusion strategy (Lee et al., 2017). M1 and M2 can be any of the three models (MobileNet, Xception, DenseNet-121). The final Global Average Pooling

Table 2
Details about the leaf datasets.

Dataset	Number of classes	Number of images/ class	Total number of images
Flavia	32	50	1600
Folio	32	18	576
Swedish leaf	15	75	1125
Leaf-12	12	320	3840

(GAP) and fully connected layers are removed from the individual architectures. It is indicated by the cross (X) mark in Fig. 2. Both the architectures utilize a transfer learning approach for initializing the weight parameters (ImageNet weights). The dataset (Flavia, Folio, Swedish Leaf and custom-developed Leaf-12) images are resized to 100 × 100 pixels and fed into the bilateral CNN's for the process of feature extraction. The obtained feature vectors from the individual architectures are concatenated to form a final feature vector. Then, it is passed on to the machine learning classifier (C) to get the final prediction results. The complete plant species recognition process is described in Algorithm 1.

Algorithm 1. Plant species recognition using a bilateral CNN
Input: Input Image, I; CNN Models, M1, M2; Classifier, C;
Output: Feature Vector, f_1, f_2, and f_{bi}; Predicted label, y_i
Step 1: Read the N number of input images (I)
Step 2: Resize the images (I) to 100 × 100 × 3
Step 3: Process the individual CNN models (M1 and M2). Take into account the number of images (N) in the dataset and the number of layers (L) in each CNN model.
Step 4: Feature Extraction
for $i = 1$ to N

i) Obtain the feature vector, f_1 from the CNN model, M1
ii) Obtain the feature vector, f_2 from the CNN model, M2

$$f_{bi} = cat(f_1, f_2)$$

Here, '*cat*' represents concatenation operation on f_1 and f_2.

iii) Save 'f_{bi}' features to a file.

End for
Step 5: Apply classifier, C to the features, f_{bi} to predict the label, y_i
Step 6: Generate the classification report

4.1. Datasets

The four datasets used in the experimental studies are Flavia (Wu et al., 2007), Folio (Munisami et al., 2015), Swedish leaf (Oskar and Söderkvist, 2001) and custom created Leaf-12. Sample leaf images with class labels in custom created Leaf-12 dataset are shown in Fig. 3(i). The images of Leaf-12 dataset are captured under varying color backgrounds (orange, pink, blue, red), illumination settings (dim, bright), orientation variations, inter-similar leaves, simple leaves, compound leaves and cluttered background. Sample leaf images with variations (Inter-similar leaves, Orientation, Illumination, Simple or Compound leaves, Color backgrounds) are shown in Fig. 3(ii). Details related to the datasets are provided in Table 2. The images in the datasets are of variable sizes. To maintain uniformity, the images are resized to 300 × 300 by adding extra white pixels to the images of smaller size (< 300 × 300 pixels). Further, the images are resized to 100 × 100 pixels by using the nearest interpolation method. The images in the dataset are split in the ratio of 70:30 (Training data: Test data).

Table 3
Extracted feature vector from homogeneous DL.

Model	Feature vector		Final feature vector
	M1	M2	
MoMoNet	9216-d	9216-d	18,432-d
XXNet	18,432-d	18,432-d	36,864-d
DeDeNet	9216-d	9216-d	18,432-d

Table 4
Extracted feature vector from heterogeneous DL.

Model	Feature vector		Final feature vector
	M1	M2	
MoXNet	9216-d	18,432-d	27,648-d
XDeNet	18,432-d	9216-d	27,648-d
MoDeNet	9216-d	9216-d	18,432-d

4.2. Feature extraction

The individual architectures (M1 and M2) in the bilateral CNN are used in the process of feature extraction. The models considered for M1 and M2 are MobileNet, Xception and DenseNet-121. The bilateral CNN is studied under two categories namely, Homogenous DL (same architecture is used as M1 and M2, named as MoMoNet, XXNet, DeDeNet) and Heterogenous DL (different architecture is used as M1 and M2, named as MoXNet, XDeNet, MoDeNet).

4.2.1. Homogeneous DL

In Homogeneous DL, the same architecture is used in the M1 and M2 branch of bilateral CNN. This is intended to improvise the plant species recognition rate. By concatenating the feature vectors obtained from M1 and M2 models, the dimension of the resultant feature vector is increased. The obtained feature vector dimension of Homogeneous DL are listed in Table 3. It results in obtaining higher performance metrics when the final feature vector is classified using the machine learning classifier.

4.2.2. Heterogeneous DL

Different architectures are used in the M1 and M2 branch of bilateral CNN. This is designed to improvise the plant species recognition rate. By concatenating the feature vectors obtained from M1 and M2 models, the dimension of the resultant feature vector is increased. The obtained feature vector dimension of Heterogeneous DL are listed in Table 4. In addition to the higher dimension of the feature vector, the properties of the individual architectures also play a vital role in obtaining higher performance metrics.

4.3. Classification

The supervised machine learning algorithms are used to construct a classifier (for categorizing C classes) (Krishnapuram et al., 2005) for the N number of training examples. Each training example has 'd' number of features i.e. input vector, $x = [x_1, x_2, \ldots x_d]^T \in \mathbb{R}^d$. The class labels are represented as an encoding vector, $y = [y^{(1)}, y^{(2)}, \ldots y^{(m)}]^T$. The feature vector with class labels is provided as an input to the classifier during the training phase of the model. The classifiers utilized in this study are Linear Discriminant Analysis (LDA), multinomial Logistic Regression (MLR), Naïve Bayes (NB), k-Nearest Neighbor (k-NN), Classification and Regression Tree (CART), Random Forest Classifier (RF), Bagging Classifier (BC), Multi-Layer Perceptron (MLP) and Support Vector Classifier (SVM). After the training phase, the model is evaluated using the test data to obtain the performance metrics.

Table 5
Plant species recognition accuracy obtained using the bilateral CNN with different classifiers on the Flavia dataset.

Feature extractor	Classifiers								
	LDA	MLR	NB	KNN	CART	RF	BC	MLP	SVM
MoMoNet	60.00	96.67	93.12	95.00	69.58	95.42	88.54	94.79	96.25
XXNet	81.25	92.50	64.38	82.92	57.71	86.04	73.96	90.83	91.25
DeDenet	87.08	97.92	83.54	93.96	73.54	96.25	88.33	96.88	96.67
MoXNet	83.12	96.88	93.12	94.58	70.00	94.38	87.92	95.42	96.04
XDeNet	94.79	98.12	85.42	94.79	70.42	96.04	89.79	96.37	97.29
MoDeNet	89.38	98.71	92.50	95.83	74.79	97.08	87.92	96.67	96.45

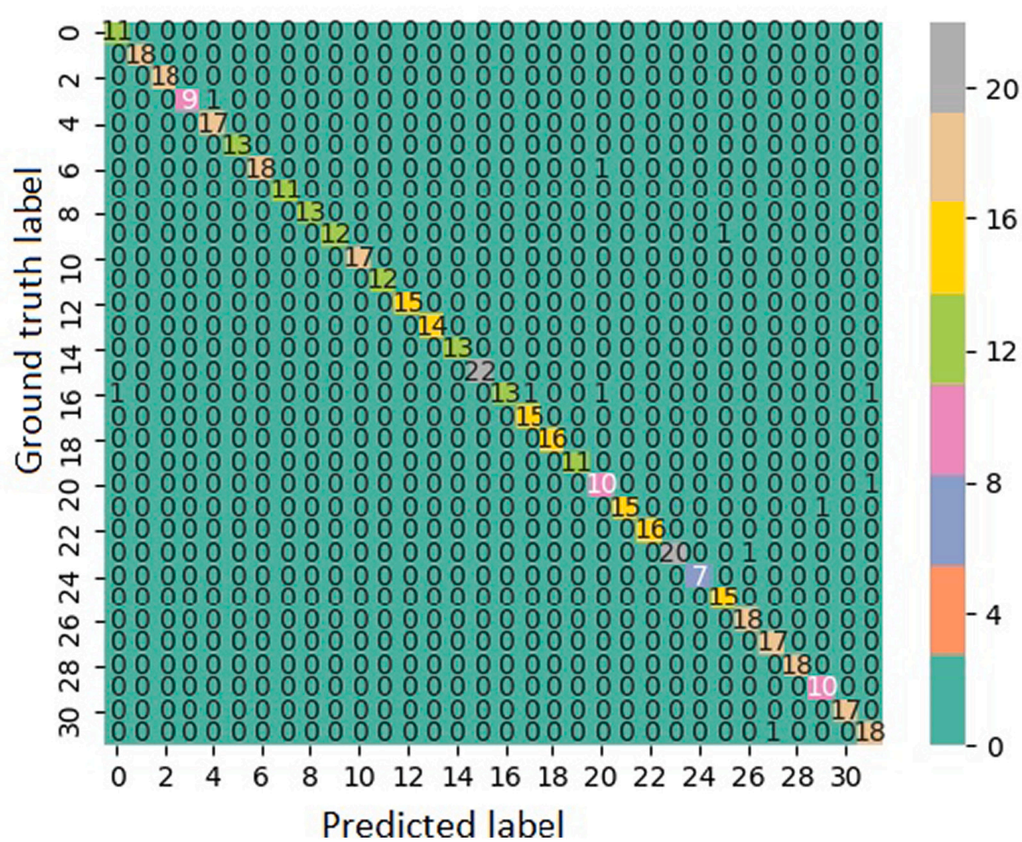

Fig. 4. Confusion matrix obtained using the proposed MoDeNet+MLR on Flavia dataset.

5. Results and discussion

The bilateral CNN model (Homogeneous DL and Heterogeneous DL) is experimented using an i7 processor with 20 GB RAM coupled with NVIDIA Titan X GPU. For the software implementation, Python framework and other assisting packages such as Numpy, OpenCV, OS, Scikit-Learn (Pedregosa et al., 2011), Matplotlib, Seaborn, Keras (Chollet, 2018) with Tensorflow Backend, and H5py are used. The CNN architectures in the bilateral CNN are used for feature extraction. Whereas, the machine learning classifiers are used for classification of the plant species. Few of the machine learning classifiers such as multinomial Logistic Regression (MLR), k-Nearest Neighbor (k-NN), Support Vector Classifier (SVM), Random Forest Classifier (RF) and Multi-Layer Perceptron (MLP) requires the setting of hyperparameters. For MLR, solver is set as 'newton-cg', L2- penalty term with value C = 1. In k-NN, the optimum hyperparameter k value is chosen to be 1, after evaluating the performance by varying the k value between 1 and 50. The number of trees is chosen to be 200 for RF Classifier. For MLP, optimizer is set as 'Adam', initial learning rate as 0.001, batch size as 32 and number of neurons in the hidden layer is chosen as 512. Support vector machine (SVM) is set with a linear kernal and C = 1.

5.1. Flavia dataset

The performance metric (accuracy in %) obtained using the bilateral CNN with different classifiers on the Flavia dataset is listed in Table 5. It is observed that the usage of multinomial Logistic Regression as a classifier resulted in higher accuracy when compared to other classifiers. It is also observed that the heterogeneous DL's in bilateral CNN produced better accuracy. It indicates that the architectures with different features/properties when combined to form bilateral CNN yields higher performance metrics. Out of all the models considered, the MoDeNet (MobileNet + DenseNet-121) combined with multinomial Logistic Regression classifier produced the state-of-the-art result (accuracy of 98.71%) on the Flavia dataset. The obtained result is better when compared to S-Inception (Wang and Wang, 2019) and Dual-path CNN (Rizk, 2019) by 3.39% and 1.91%, respectively. The confusion matrix obtained using the MoDeNet + MLR on the Flavia dataset is shown in Fig. 4. From the matrix data, it is observed that the number of mispredictions/class is on a smaller scale yielding higher accuracy.

Table 6
Plant species recognition accuracy obtained using the bilateral CNN with different classifiers on the Folio dataset.

Feature extractor	Classifiers								
	LDA	MLR	NB	KNN	CART	RF	BC	MLP	SVM
MoMoNet	86.71	93.06	88.44	89.60	54.91	89.60	75.14	88.44	91.33
XXNet	83.24	87.86	70.52	77.46	48.55	78.03	63.58	83.82	84.97
DeDenet	91.91	95.95	85.55	89.60	60.69	91.33	80.35	94.22	93.06
MoXNet	80.92	93.64	89.60	91.33	60.12	87.86	71.68	88.44	93.54
XDeNet	90.17	95.38	87.86	90.75	64.74	91.33	81.45	94.02	94.22
MoDeNet	90.75	96.38	90.75	93.06	58.96	92.49	81.50	91.91	94.80

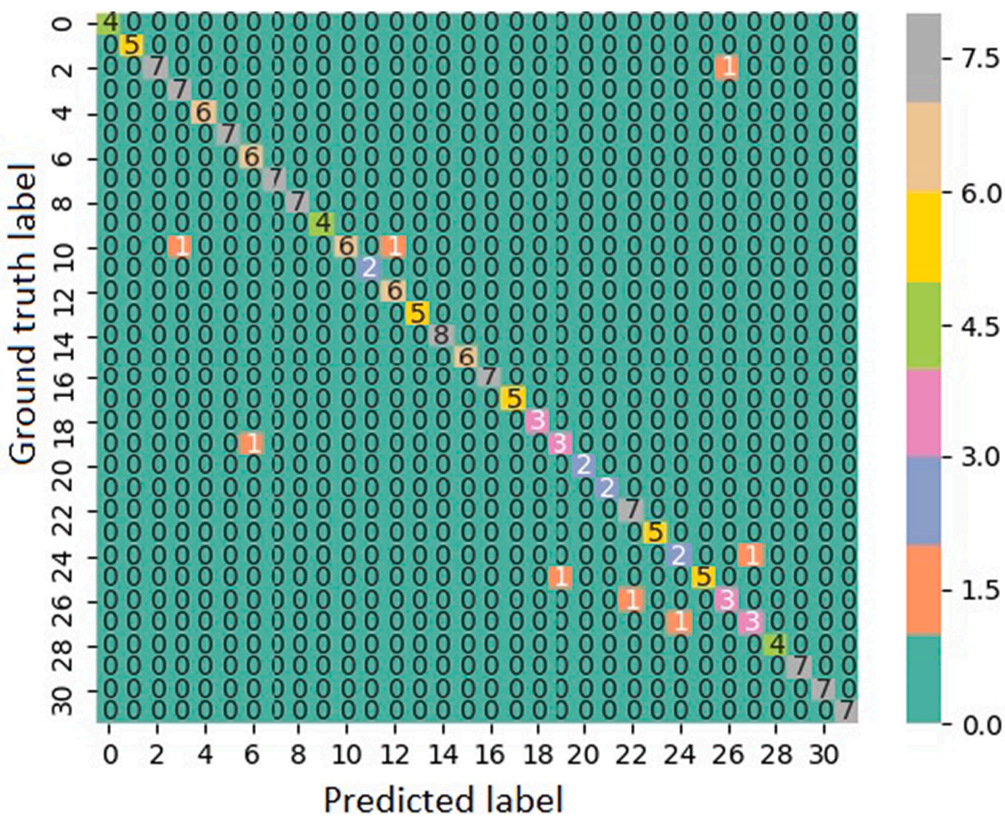

Fig. 5. Confusion matrix obtained using proposed MoDeNet+MLR on Folio Dataset.

Table 7
Plant species recognition accuracy obtained using the bilateral CNN with different classifiers on the Swedish leaf dataset.

Feature Extractor	Classifiers								
	LDA	MLR	NB	KNN	CART	RF	BC	MLP	SVM
MoMoNet	94.67	99.13	91.72	95.86	80.18	95.86	93.49	98.22	98.41
XXNet	90.53	96.75	79.88	89.35	74.56	92.01	86.98	94.38	96.15
DeDenet	93.20	99.23	85.21	94.97	79.29	97.93	93.79	97.63	98.22
MoXNet	96.15	99.00	92.60	96.45	79.29	96.75	94.38	96.75	98.11
XDeNet	98.52	98.82	89.35	96.15	82.25	97.63	94.67	97.34	97.62
MoDeNet	97.34	99.41	95.27	97.04	82.25	99.11	95.86	98.22	98.43

5.2. Folio dataset

The performance metric (accuracy in %) obtained using the bilateral CNN with different classifiers on the Folio dataset is listed in Table 6. It is observed that the usage of multinomial logistic regression as a classifier achieved higher accuracy when compared to other classifiers. It is also observed that the heterogeneous DL's in bilateral CNN produced comparatively better accuracy. It indicates that the architectures with different features/properties when combined to form bilateral CNN yields higher performance metrics. Out of all the models considered, the MoDeNet (MobileNet + DenseNet-121) combined with multinomial Logistic Regression classifier produced the state-of-the-art result (accuracy of 96.38%) on the Folio dataset. The confusion matrix obtained using the MoDeNet + MLR on the Folio dataset is shown in Fig. 5. From the matrix data, it is observed that the number of mispredictions/class is on a smaller scale yielding higher accuracy.

5.3. Swedish leaf dataset

The performance metric (accuracy in %) obtained using the bilateral

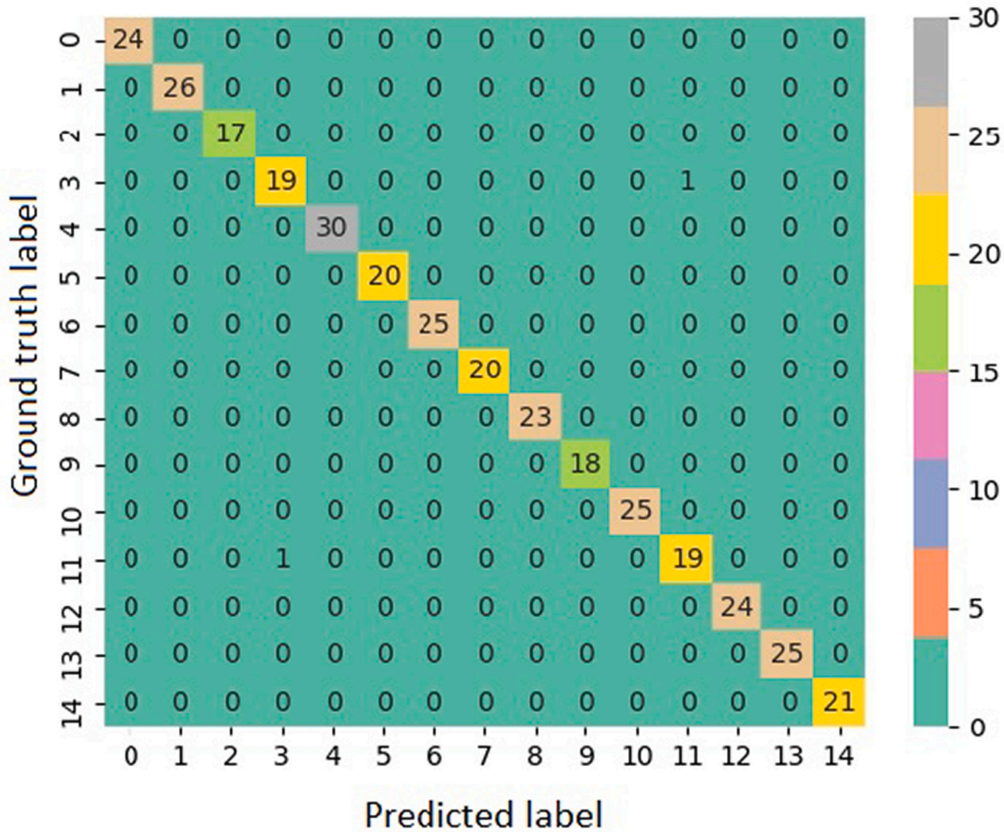

Fig. 6. Confusion matrix obtained using proposed MoDeNet + MLR on Swedish leaf dataset.

Table 8
Plant species recognition accuracy obtained using the bilateral CNN with different classifiers on the leaf-12 dataset.

Feature Extractor	Classifiers								
	LDA	MLR	NB	KNN	CART	RF	BC	MLP	SVM
MoMoNet	96.09	97.48	59.11	94.01	61.20	93.23	80.30	96.70	97.22
XXNet	92.45	94.53	66.93	87.76	56.08	85.94	72.14	93.84	94.36
DeDenet	98.00	98.09	88.11	95.23	68.32	95.49	84.55	97.74	97.83
MoXNet	96.96	97.83	59.20	94.62	61.20	92.36	79.25	97.14	97.74
XDeNet	97.57	98.18	81.08	95.57	67.62	95.05	86.20	97.92	97.82
MoDeNet	98.05	99.39	57.99	96.27	69.18	96.53	85.85	98.78	98.31

CNN with different classifiers on the Swedish Leaf dataset is listed in Table 7. It is observed that the usage of multinomial Logistic Regression as a classifier contributed to higher accuracy when compared to other classifiers. It is also observed that the heterogeneous DL's in bilateral CNN produced comparatively better accuracy. It indicates that the architectures with different features when combined to form bilateral CNN yields higher performance metrics. Out of all the models considered, the MoDeNet (MobileNet + DenseNet-121) combined with multinomial Logistic Regression classifier produced the state-of-the-art result (accuracy of 99.41%) on the Swedish Leaf dataset. The obtained result is 7.74% higher compared to the S-Inception method (Wang and Wang, 2019). The confusion matrix obtained by using the MoDeNet + MLR on the Swedish Leaf dataset is shown in Fig. 6. From the matrix data, it is observed that the number of mispredictions/class is on a smaller scale resulting in enhanced accuracy.

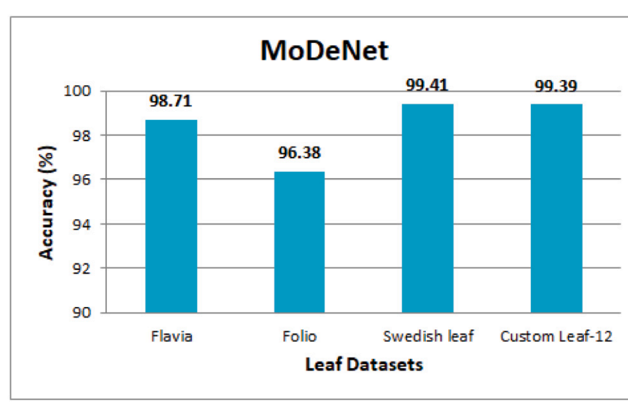

Fig. 7. Confusion matrix obtained using proposed MoDeNet+MLR on Leaf-12 dataset.

5.4. Leaf-12 dataset

The performance metric (accuracy in %) obtained using the bilateral CNN with different classifiers on custom created Leaf-12 dataset is listed

Table 9
Performance metrics obtained for optimal homogeneous and heterogeneous bilateral CNN.

Dataset	Bilateral CNNs	Accuracy		Performance Metrics			Computational Time (s)	
		Top-1	Top-5	Pr	Re	F1	Extract features and Train	Test
Flavia	DeDeNet + MLR	97.92	99.58	0.98	0.98	0.98	284.69	0.02
	Proposed MoDeNet + MLR	98.71	99.79	0.98	0.98	0.98	239.27	0.02
Folio	DeDeNet + MLR	95.95	99.42	0.96	0.96	0.96	80.2	0.02
	Proposed MoDeNet + MLR	96.38	100	0.96	0.95	0.95	60.82	0.01
Swedish leaf	DeDeNet + MLR	99.23	100	0.99	0.99	0.99	157.3	0.02
	Proposed MoDeNet + MLR	99.41	100	0.99	0.99	0.99	80.62	0.02
Leaf-12	DeDeNet + MLR	98.09	100	0.98	0.98	0.98	706.27	0.05
	Proposed MoDeNet + MLR	99.39	100	0.99	0.99	0.99	541.45	0.04

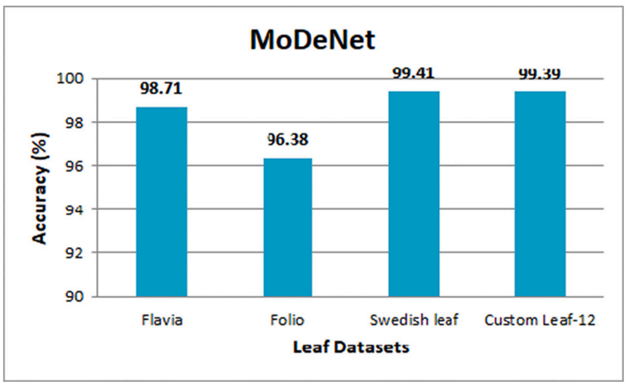

Fig. 8. Accuracy obtained by using bilateral MoDeNet CNN+ MLR on different datasets.

in Table 8. It is observed that the usage of multinomial Logistic Regression as a classifier produced higher accuracy when compared to other classifiers. It is also observed that the heterogeneous DL's in bilateral CNN produced comparatively better accuracy. It indicates that the architectures with different properties when combined to form bilateral CNN yields higher performance metrics. Out of all the models considered, the MoDeNet (MobileNet + DenseNet-121) combined with multinomial Logistic Regression classifier produced the state-of-the-art result (accuracy of 99.39%) on the Leaf-12 dataset. The confusion matrix obtained by using the MoDeNet+MLR on the Leaf-12 dataset is shown in Fig. 7. From the matrix data, it is observed that the number of mispredictions/class is on a smaller scale resulting in yielding high accuracy values. The results achieved using a custom created Leaf-12 dataset is in correlation with the results of the benchmark datasets (Flavia, Folio, Swedish Leaf).

5.5. Discussion of results

The optimal results obtained by using the bilateral CNN (Homogeneous and Heterogeneous) are listed in Table 9. The performance metrics considered are Top-1 accuracy, Top-5 accuracy, Precision (Pr), Recall (Re), F1-score (F1) and computational time. Prediction accuracy turns out to be good performance metric if we have a balanced dataset (equal number of images per class). As shown in Table. 2, the considered dataset for experimental investigation contains equal number of images per class (Flavia = 50 images/class, Folio = 18 images/class, Swedish Leaf = 75 images/class, Leaf-12 = 320 images/class). Hence, the prediction accuracy with computation time is considered as an important parameter to evaluate the model's performance. Fig. 8 shows the plant species recognition accuracy obtained by using MoDeNet CNN + MLR on different datasets such as Flavia (98.71%), Folio (96.38%), Swedish Leaf (99.41%) and Leaf-12 (99.39%).

In Table 9, the performance metrics (accuracy - Top 1, Top 5; Precision; Recall; F1-Score) and computation time is compared between the

Table 10
Comparison between the prediction accuracy in % obtained by the bilateral MoDeNet CNN with other reported methods.

Machine Learning (ML)/Deep Learning (DL)/ Ensemble of CNN Architecture (ECA)	Method	Accuracy obtained for different Dataset (%)			
		Flavia	Folio	Swedish leaf	Custom Leaf-12
ML	Improved Linear Regression (Goyal et al., 2020)	84.86%	81.32%	90.8%	–
ML	RF (Anubha Pearline and Sathiesh Kumar, 2020)	85.62%	78.03%	87.87%	82.38%
ML	RF with Whale Optimization Algorithm (WOA) (Pankaja and Suma, 2020)	97.58%	–	97.58%	–
DL	LeafNet (Barré et al., 2017)	97.9%	–	–	–
DL	D-Leaf (Tan et al., 2020)	94.63%	–	98.09%	–
DL	Custom CNN (4 Convolutions +2 Dense) (Bisen, 2021)	–	–	97%	–
DL	CNN-Recurrent Neural Network (RNN) (Kaya et al., 2019b)	92.65%	–	99.11%	–
DL	VGG-16 + Logistic Regression (Anubha Pearline and Sathiesh Kumar, 2020)	95%	93.64%	98.52%	97.14%
ECA	Dual-path CNN (Rizk, 2019)	96.8%	–	–	–
ECA	Multi-model Convolutional Neural Network (U-Net + simple CNN) (Dat et al., 2021)	95.42%	–	–	–
ECA	S-Inception (Wang and Wang, 2019)	95.32%	–	91.37%	–
ECA	B-CNN (Anubha Pearline and Sathiesh Kumar, 2020)	97.71%	–	98.67%	97.70%
ECA	Bilateral MoDeNet CNN + MLR model	98.71%	96.38%	99.41%	99.39%

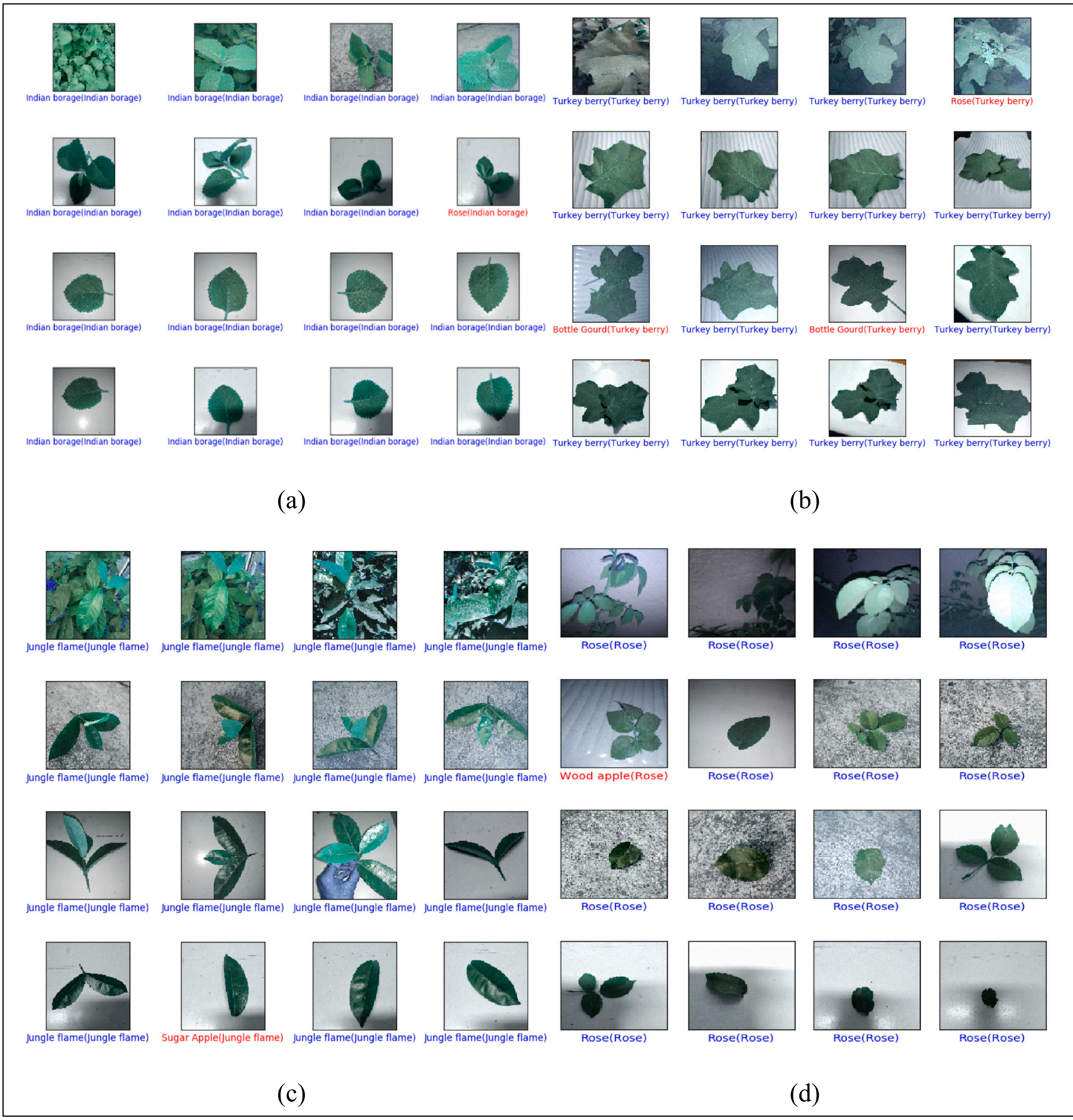

Fig. 9. Real-time leaf prediction using the proposed bilateral MoDeNet + MLR model (a) Indian Borage (b) Turkey berry (c) Jungle Flame (d) Rose. Note: In each image, labels are mentioned in the format: predicted label (ground-truth label). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

best Homogenous DL (DeDeNet + MLR) and Heterogenous DL (MoDeNet + MLR) models. It is observed that the MoDeNet + Multinomial Logistic Regression model performed better (higher performance metrics with lower training time) when compared to DeDeNet + Multinomial Logistic Regression model. The trend is true irrespective of the dataset considered in the experimental investigation. This emphasis on the proposed model (MoDeNet +MLR) to be highly accurate in predicting the plant species and also at a faster rate when compared with other bilateral network models (MoMoNet + MLR, XXNet + MLR, DeDeNet + MLR, MoXNet +MLR, XDeNet + MLR). Hence, the real-time leaf prediction or classification is carried out by utilizing the MoDeNet + MLR model.

Table 10 lists the comparison between the prediction accuracy in % obtained by the proposed method (Bilateral MoDeNet CNN) with other reported methods (Machine Learning - ML, Deep Learning - DL, Ensemble of CNN Architectures - ECA). The reported accuracy is segregated depending on the dataset (Flavia, Folio, Swedish Leaf and Leaf-12). From Table 10 data, it is observed that the proposed Bilateral MoDeNet CNN (MobileNet + DenseNet-121 + Multinomial Logistic Regression) yielded better performance metric- accuracy (Flavia =

98.71%, Folio = 96.38%, Swedish Leaf = 99.41% and custom created Leaf-12 = 99.39%) when compared with other reported methods. The obtained results are consistent irrespective of the dataset considered.

5.6. Real-time leaf prediction

The bilateral MoDeNet with MLR classifier is introduced to the acquired real-time images of few plant species. The leaf images are captured using a Vivo V9 pro mobile phone with a 13 MP rear camera, Android 8.1 OS, and Qualcomm Snapdragon 660 processor. The leaf images are acquired by varying the distance between the camera and leaf (10 cm to 30 cm). Also, the real-time conditions such as orientation changes, illumination variation, cluttered background, simple leaves, and compound leaves are considered during the process of collecting the real-time images. Fig. 9 (a-d) depicts the images obtained from plant species such as Indian Borage, Turkey Berry, Jungle Flame and Rose with their predicted outputs (MoDeNet +MLR model). Based on the results obtained, it is concluded that the proposed bilateral MoDeNet+MLR is highly suitable for recognizing the plant species in a real-time environment. As a future scope, more plant species will be recognized by increasing the number of classes in the Leaf-12 dataset.

6. Conclusion

Investigation of bilateral CNN with machine learning classifier is carried out in the context of real-time plant species recognition. The Bilateral CNNs (Homogeneous/Heterogeneous type) are used as feature extractors. The CNN models considered are MobileNet, Xception and DenseNet-121. The extracted features from the individual models are concatenated to form the final feature vector. Then, the final feature vector is classified using the machine learning classifiers. From the experimental results, it is observed that the MoDeNet + MLR (MobileNet + DenseNet-121 + multinomial Logistic Regression classifier) model produced the state-of-the-art results, irrespective of the dataset considered (Flavia: 98.71%, Folio: 96.38%, Swedish Leaf: 99.41% and custom created Leaf-12: 99.39%). The number of mispredictions/class is significantly reduced. The unseen real-time images collected are tested using the MoDeNet Bilateral CNN + MLR. The proposed methodology resulted in high recognition rate of plant species. Hence, the MoDeNet +MLR model is highly suitable for predicting the plant species in a real-time environment (scale variations, illumination changes, cluttered backgrounds, viewpoint or orientation changes, simple or compound structure of leaf). Future research will shed light by designing lightweight CNNs on plant species recognition. It is vital to establish a dedicated mobile application for recognizing multi-continental plant species.

Declaration of Competing Interest

None.

Acknowledgment

The authors would like to thank the NVIDIA Corporation for providing NVIDIA Titan X GPU under the University Research Grant Programme.

References

Ahmed, A., Hussein, S.E., 2020. Leaf identification using radial basis function neural networks and SSA based support vector machine. PLoS One. https://doi.org/10.1371/journal.pone.0237645.

Anubha Pearline, S., Sathiesh Kumar, V., 2020. High performance ensembled convolutional neural network for plant species recognition. In: International Conference on Computer Vision and Image Processing. Springer, Singapore, pp. 526–538. December.

Anubha Pearline, S., Sathiesh Kumar, V., Harini, S., 2019. A study on plant recognition using conventional image processing and deep learning approaches. J. Intell. Fuzzy Syst. https://doi.org/10.3233/JIFS-169911.

Araujo, V.M., Britto, A.S., Brun, A.L., Koerich, A.L., Oliveira, L.E.S., 2018. Fine-grained hierarchical classification of plant leaf images using fusion of deep models. Proc. Int. Conf. Tools with Artif. Intell. ICTAI. https://doi.org/10.1109/ICTAI.2018.00011.

Barbhuiya, A.A., Karsh, R.K., Jain, R., 2021. CNN based feature extraction and classification for sign language, multimed. Tools Appl. 80, 3051–3069. https://doi.org/10.1007/s11042-020-09829-y.

Barré, P., Stöver, B.C., Müller, K.F., Steinhage, V., 2017. LeafNet: a computer vision system for automatic plant species identification. Ecol. Inform. https://doi.org/10.1016/j.ecoinf.2017.05.005.

Bisen, D., 2021. Deep convolutional neural network based plant species recognition through features of leaf. Multimed. Tools Appl. 80, 6443–6456. https://doi.org/10.1007/s11042-020-10038-w.

Chollet, F., 2017. Xception: deep learning with depthwise separable convolutions. In: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR. https://doi.org/10.1109/CVPR.2017.195.

Chollet, F., 2018. Keras: The Python Deep Learning Library. https://keras.io/.

Chompookham, T., Surinta, O., 2021. Ensemble methods with deep convolutional neural networks for plant leaf recognition. ICIC Express Lett. 15, 553–565. https://doi.org/10.24507/icicel.15.06.553.

Dat, T.T., Le Thien Vu, P.C., Truong, N.N., Anh Dang, L.T., Thanh Sang, V.N., Bao, P.T., 2021. Leaf recognition based on joint learning multiloss of multimodel convolutional neural networks: a testing for Vietnamese herb. Comput. Intell. Neurosci. 2021, 1–19. https://doi.org/10.1155/2021/5032359.

Dede, M.A., Aptoula, E., Genc, Y., 2019. Deep network ensembles for aerial scene classification. IEEE Geosci. Remote Sens. Lett. 16, 732–735. https://doi.org/10.1109/LGRS.2018.2880136.

Goyal, N., Kumar, N., Kapil, 2020. On solving leaf classification using linear regression. Multimed. Tools Appl. 4533–4551. https://doi.org/10.1007/s11042-020-09899-y.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., Lew, M.S., 2016. Deep learning for visual understanding: a review. Neurocomputing. https://doi.org/10.1016/j.neucom.2015.09.116.

He, G., Xia, Z., Zhang, Q., Zhang, H., Fan, J., 2018. Plant species identification by bi-channel deep convolutional networks. J. Phys. Conf. Ser. https://doi.org/10.1088/1742-6596/1004/1/012015.

Hira, S., Bai, A., Hira, S., 2021. An automatic approach based on CNN architecture to detect Covid-19 disease from chest X-ray images. Appl. Intell. 51, 2864–2889. https://doi.org/10.1007/s10489-020-02010-w.

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications arXiv preprint arXiv:1704.04861.

Hu, J., Chen, Z., Yang, M., Zhang, R., Cui, Y., 2018. A multiscale fusion convolutional neural network for plant leaf recognition. IEEE Signal. Process. Lett. https://doi.org/10.1109/LSP.2018.2809688.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR, p. 2017. https://doi.org/10.1109/CVPR.2017.243.

Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H., Tekinerdogan, B., 2019a. Analysis of transfer learning for deep neural network based plant classification models. Comput. Electron. Agric. https://doi.org/10.1016/j.compag.2019.01.041.

Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H., Tekinerdogan, B., 2019b. Analysis of transfer learning for deep neural network based plant classification models. Comput. Electron. Agric. 158, 20–29. https://doi.org/10.1016/j.compag.2019.01.041.

Krishnapuram, B., Carin, L., Figueiredo, M.A.T., Hartemink, A.J., 2005. Sparse multinomial logistic regression: fast algorithms and generalization bounds. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2005.127.

Kumar, M., Gupta, S., Gao, X.Z., Singh, A., 2019. Plant species recognition using morphological features and adaptive boosting methodology. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2952176.

Lee, S.H., Chan, C.S., Mayo, S.J., Remagnino, P., 2017. How deep learning extracts and learns leaf features for plant classification. Pattern Recogn. https://doi.org/10.1016/j.patcog.2017.05.015.

Lee, S.H., Chan, C.S., Remagnino, P., 2018. Multi-organ plant classification based on convolutional and recurrent neural networks. IEEE Trans. Image Process. https://doi.org/10.1109/TIP.2018.2836321.

Mehdipour Ghazi, M., Yanikoglu, B., Aptoula, E., 2017. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing. https://doi.org/10.1016/j.neucom.2017.01.018.

Minetto, R., Pamplona Segundo, M., Sarkar, S., 2019. Hydra: an ensemble of convolutional neural networks for geospatial land classification. IEEE Trans. Geosci. Remote Sens. 57, 6530–6541. https://doi.org/10.1109/TGRS.2019.2906883.

Mostajer Kheirkhah, F., Asghari, H., 2019. Plant leaf classification using GIST texture features. IET Comput. Vis. https://doi.org/10.1049/iet-cvi.2018.5028.

Munisami, T., Ramsurn, M., Kishnah, S., Pudaruth, S., 2015. Plant leaf recognition using shape features and colour histogram with K-nearest neighbour classifiers. Proc. Comput. Sci. https://doi.org/10.1016/j.procs.2015.08.095.

Oskar, J., Söderkvist, O., 2001. Computer Vision Classifcation of Leaves from Swedish Trees. Linkoping University. Master's Thesis.

Pankaja, K., Suma, V., 2020. Plant leaf recognition and classification based on the whale optimization algorithm (WOA) and random forest (RF). J. Inst. Eng. Ser. B 101, 597–607. https://doi.org/10.1007/s40031-020-00470-9.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.

Pundir, A.S., Raman, B., 2019. Dual deep learning model for image based smoke detection. Fire. Technol 55, 2419–2442. https://doi.org/10.1007/s10694-019-00872-2.

Rao, D., Huang, S., Jiang, Z., Deverajan, G.G., Patan, R., 2021. A dual deep neural network with phrase structure and attention mechanism for sentiment analysis. Neural Comput. Applic. 6 https://doi.org/10.1007/s00521-020-05652-6.

Rizk, S., 2019. Plant Leaf Classification Using Dual Path Convolutional Neural Networks. Doctoral dissertation. Notre Dame University-Louaize.

Shah, M.P., Singha, S., Awate, S.P., 2018. Leaf classification using marginalized shape context and shape+texture dual-path deep convolutional neural network. Proc. Int. Conf. Image Process. ICIP. https://doi.org/10.1109/ICIP.2017.8296403.

Sun, Y., Liu, Y., Wang, G., Zhang, H., 2017. Deep learning for plant identification in natural environment. Comput. Intell. Neurosci. https://doi.org/10.1155/2017/7361042.

Tan, J.W., Chang, S.W., Abdul-Kareem, S., Yap, H.J., Yong, K.T., 2020. Deep learning for plant species classification using leaf vein morphometric. IEEE/ACM Trans. Comput. Biol. Bioinforma. 17, 82–90. https://doi.org/10.1109/TCBB.2018.2848653.

Ullah, W., Ullah, A., Haq, I.U., Muhammad, K., Sajjad, M., Baik, S.W., 2021. CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks. Multimed. Tools Appl. 80, 16979–16995. https://doi.org/10.1007/s11042-020-09406-3.

Wang, B., Wang, D., 2019. Plant leaves classification: A few-shot learning method based on siamese network. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2947510.

Wang, X., Du, W., Guo, F., Hu, S., 2020. Leaf recognition based on elliptical half Gabor and maximum gap local line direction pattern. IEEE Access. https://doi.org/10.1109/ACCESS.2020.2976117.

Wang, D., Li, M., Ben-Shlomo, N., Corrales, C.E., Cheng, Y., Zhang, T., Jayender, J., 2021. A novel dual-network architecture for mixed-supervised medical image segmentation. Comput. Med. Imaging Graph. 89 https://doi.org/10.1016/j.compmedimag.2020.101841.

Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.X., Chang, Y.F., Xiang, Q.L., 2007. A leaf recognition algorithm for plant classification using probabilistic neural network. In: ISSPIT 2007–2007 IEEE Int. Symp. Signal Process. Inf. Technol. https://doi.org/10.1109/ISSPIT.2007.4458016.

Yildirim-Yayilgan, S., Arifaj, B., Rahimpour, M., Hardeberg, J.Y., Ahmedi, L., 2021. Pre-trained CNN based deep features with hand-crafted features and patient data for skin lesion classification. Commun. Comput. Inf. Sci. 1382, 151–162. https://doi.org/10.1007/978-3-030-71711-7_13.

Zhu, Y., Sun, W., Cao, X., Wang, C., Wu, D., Yang, Y., Ye, N., 2019. TA-CNN: two-way attention models in deep convolutional neural network for plant recognition. Neurocomputing. https://doi.org/10.1016/j.neucom.2019.07.016.

Anubha Pearline S received her Bachelor's degree in Computer Science Engineering from Tagore Engineering College, Anna University. She pursued her Master's degree in Information Technology from Madras Institute of Technology Campus, Anna University. She completed her research at Madras Institute of Technology Campus, Anna University. Currently, she is an Assistant Professor at SRM Institute of Science and Technology, Vadapalani campus, Chennai. Her research interests include Deep Learning, Computer Vision, Pattern Recognition and Machine Learning.

Sathiesh Kumar V received his Doctorate of Philosophy from Indian Institute of Technology, Madras. Currently, he is serving as an Assistant Professor in the Department of Electronics Engineering, Madras Institute of Technology Campus, Anna University. Also, he is the Head of Opto-Mechatronics laboratory, Department of Electronics Engineering, Madras Institute of Technology Campus, Anna University. His research interests include Artificial Intelligence, Computer Vision, Machine Learning, Laser Spectroscopy and Embedded Systems. He has completed two funded projects supported by Science and Engineering Research Board (SERB) and Centre for Technology Development and Transfer (CTDT). He is one of the principal investigators of a project under University Grant Commission (UGC). Also, two other projects funded by TANII scheme is approved. He is a member of Indian Unit for Pattern Recognition and Artificial Intelligence (IUAPRI), IEEE, Indian Laser Association and SPIE.