International Journal of Academic Information Systems Research (IJAISR)
ISSN: 2643-9026
Vol. 6 Issue 5, May - 2022, Pages: 9-13

# Deep Learning Leaf Classification System based on Transfer Learning and Augmentation Strategy

**Yashas N Gowda**

Larsen & Toubro Infotech Limited (LTI)
Bangalore, India

yashasn1999@gmail.com

***Abstract****: Plant identification is an interdisciplinary focus of both botanical taxonomy and computer vision. Plant leaves have an assortment of features like design, shape, size, texture, color, and venation. These features are studied by botanists in order to identify the plant species. Today's botanists have the advantage of using computer vision based technology for automated leaf analysis and identification. Currently, computer vision researchers are increasingly inclined towards Convolutional Neural Networks and Deep Learning methods for feature extraction and classification problems given its efficiency and accuracy. This paper presents a Deep learning framework for leaf classification developed by transfer learning Resnet pretrained model. Mere transfer learning of Resnet was found inadequate for accurate classification of the heterogeneous dataset. Hence augmentation was performed to increase the samples in the dataset for a more powerful classification. The proposed methodology is evaluated on the Malayakew leaf dataset and was found to be efficient.*

**Keywords**— Leaf Classification, Transfer Learning, ResNet, Augmentation, Deep Learning, Malayakew leaf Dataset, Feature Extraction.

## 1. INTRODUCTION

Identification and classification of plant species is an important facet of researchers from different fields such as agronomists, botanists, foresters, gardeners, and environment protectors [1]. Identification of new and rare plant species is off utmost priority in order to maintain agricultural productivity, sustainability, and ecological balance. Botanical researchers study the characteristics of leaf like shape, size, color, texture, venation etc., in order to identify plants. Leaf analysis in specialized laboratory is a costly and elaborate task and requires expertise, due to the existence of enormous number of plant species (over 5000 species). But this is not a bottleneck today because computer vision provides intelligent tools which can automatically identify the species of plants by analyzing the leaf images. Nevertheless leaf identification is a challenging task because the leaf characteristics can be similar to many other leaves. This demands a powerful feature extraction phase which can extract the most descriptive and detective features from the leaf so that the performance of the classification phase can be better.

The recent Deep learning techniques eliminate the laborious task of handcrafted feature extraction and combine the feature extraction and classification into one unified model [2]. Deep learning models extract all possible features from the image from coarse to fine. Hence the recognition rate is quite high in these models.

## 2. RELATED WORKS

Recently many researchers have employed Deep learning and Convolutional Neural Networks to perform leaf classification on well known plant leaf datasets. [3] used CNN to extract leaf features and Deconvolutional network DNN to gain insight of the selected features. They suggested that hybrid feature extraction models can improve their discriminative power of the classification models. [4] developed SWP Leaf net model by transfer learning MobileNet V2 in order to model botanist’s behavior for efficient leaf identification. They reported more than 99% accuracy. [5] employed deep learning models for leaf classification from untrained images. A hybrid optimization algorithm called SS-WoH was introduced to achieve higher classification accuracy. [6] employed pre trained ImageNet CNN model for plant identification of Image CLF 2013 dataset. [7] proposed a 50 layer deep residual learning framework for plant identification. Their model achieved 93.9% recognition rate on LeafSNAP dataset. [8] Designed a multi scale fusion CNN for plant leaf recognition at multiple scales. The experimentation was done on Malayakew dataset and leaf SNAP dataset. [9] collected the first plant image data set called dataset- BJFU100 in natural scene through mobile phone. They proposed 26 layer resnet deep learning model for plant classification in a natural environment. Recognition rate was 91.78%. [10] proposed a deep learning neural network for recognition of Malaysian herbs with an accuracy of 93%. [11] constructed a 10 layer CNN for classification of Flavia Leep database. Their accuracy was 87.92%. [12] developed a CNN algorithm based on VGGH with discriminative loss functions to classify leaf images of different common bean cultivars. [13] used VGGH CNN model to extract topological feature and train attention module for leaf image recognition. This model was tested on various datasets like Flavia, Swedish,Folio and Cherry dataset. Over 99% accuracy was achieved.

[14] developed two deep Feature fusion methods - Distance fusion and Classifier fusion using deep learning technique to achieve a classification rate of 83.5% on

SoyCultivar200 leaf dataset. [15] transfer learnt Google Net CNN model and used it to classify Arabidposis and tobacco Plant images CVPPP database. The accuracy of their model was 98%. [16] performed transfer learning on 3 deep learning architectures- Google Net, Alex net and VGG Net and tested them on plant tusks dataset of LifeCLEF 2015 and obtained 80% accuracy. [17] developed a wide and deep learning framework which combines Linear model and a deep learning model using a logistic function. With this they could consider discrete features simultaneously with continuous image content to achieve better classification accuracy and plant image data set. [18] developed a customized CNN for classifying tomato plant images extracted from Kaggle dataset. [19] developed a VGGNET DL Architecture for plant seedling classification on a data set of Aarhus University signal processing group. They achieved 99.48% accuracy. [20] evaluated the effectiveness of four different transfer learning models for plant classification on four public datasets. They suggested that deep learning can improve low performance plant classification models. [21] employed VGGNET to extract deep learning features and use them to train Light GGM Classification model. They achieved a recognition rate of 93.6% on herbal plant images collected from natural environment.

In this work an alternate for botanical laboratory method of leaf identification is proposed using Deep Learning model. The pretrained ResNet50 Deep Learning Architecture is transfer learnt to customize it for leaf recognition. In order to improve the training accuracy data augmentation was performed to increase the number of leaf samples. The experimentation was conducted on the Malayakew Dataset which consists of leaf images belonging to 44 different species. The next section of this paper discusses the materials and methods used in this study. This is followed by results and discussion. The subsequent conclusion section will conclude the paper.

## 3. MATERIALS AND METHODS

The proposed deep learning methodology for leaf classification is depicted in the figure 1.There are 3 stages in the proposed model namely- Image Augmentation, Transfer learning the Resnet50 pretrained model and Leaf Classification. In order to increase the performance of the deep learning model, the Malayakew leaf image dataset is first augmented using three affine transformations – Rotation, Flipping and Brightness modification. The Resnet50 pretrained deep learning model is transfer learnt and customized to suit leaf image classification. The augmented dataset is divided into training and testing set in 80:20 ratio. In the classification stage, the training set is passed to the transfer learnt Resnet50 model for feature extraction and training. Next the trained model is validated using the test set for classifying the leaf images as appropriate species.

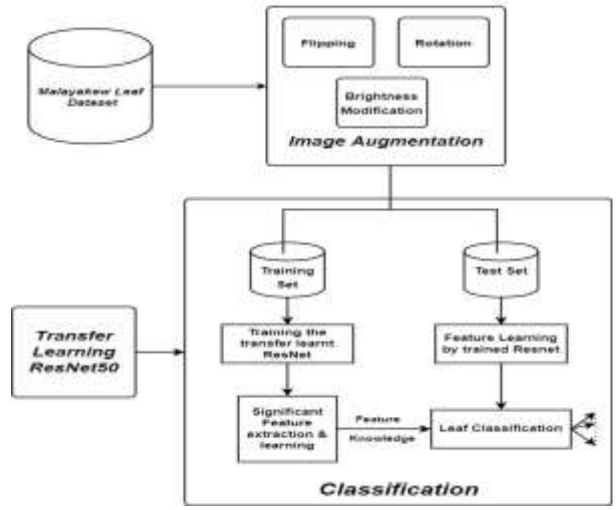


**Fig.1.** *Block Diagram of the Proposed Leaf Classification Methodology*

### 3.1 Malayakew dataset

Malayakew dataset is a collection of leaf images which are camera captured from the Royal Botanic Gardens, Kew, England. It consists of leaves belonging to 44 different plant species. Most of the species have similar appearance hence it is challenging task to classify leaves from this dataset. There are 2288 training images and 528 testing images of size 256*256 each in this collection. Figure 2 below shows some sample Malayakew leaf images.

4.

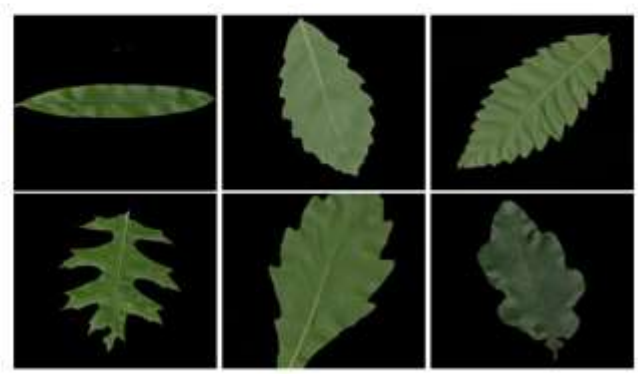

**Fig.2.** *Sample images from Malayakew Dataset*

### 3.2 Image Augmentation

Image Augmentation is a process on simulation the existing images in order to increase the number of image samples in the dataset. Recently this process has proven to optimize the recognition rate of deep learning models. In the proposed method three types of augmentations namely Rotation, Flipping and Brightness modification are performed on the dataset. These augmentations are demonstrated using a sample leaf image in the figure 3 below. For every image in the dataset clockwise(+30$^0$) and anti clockwise rotation(-30$^0$), vertical and horizontal flipping, brightness increase and decrease by 50% is performed, thus increasing the dataset size. The more the number and versatility of training images the better the recognition accuracy of the deep learning model. Hence Image augmentation is an important phase in the proposed model.

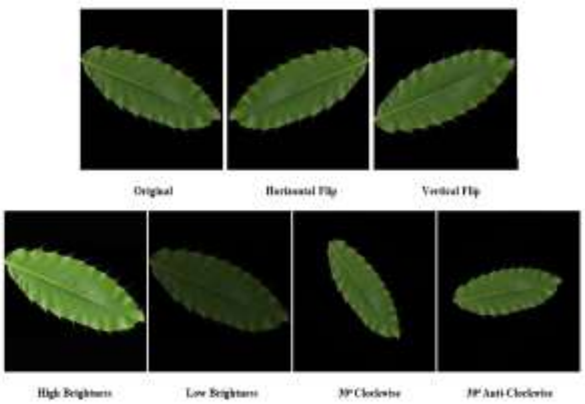


**Fig. 3**. *A sample Augmented leaf image*

### 3.3 Transfer Learning of Resnet50Image Augmentation

Resnet50 is a deep residual network which is a 50 layer deep Convolutional Neural network. This architecture stacks residual blocks on top of each other in order to form a residual neural network model as shown in figure 4 below. ResNet50 is a variant of Resnet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. For the proposed model all the layers of the pretrained Resnet50 were freezed and weights were tuned to customize it for leaf recognition.

**Fig. 4**.*Transfer learnt Resnet50 Deep Learning Architecture*

**Leaf Classification**

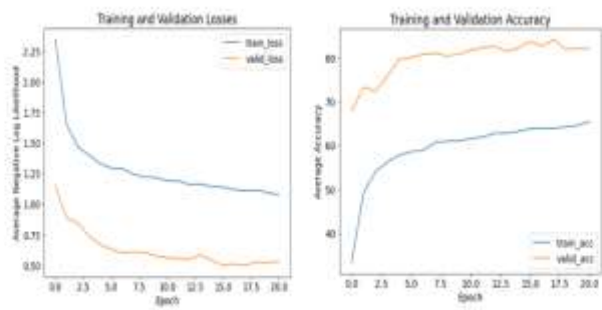


**Fig. 5**.*Training and Validation Performance of the proposed model*

The transfer learnt Resnet50 was then trained on the augmented Malayakew dataset using SGDM(Stochastic Gradient Descent Momentum) algorithm. A batch size of 10 leaf images was used for training. The weight for each batch was updated based on the error of previous batch as per the back propagation strategy of SGDM algorithm. The learning rate α was set to 0.0001 and the training was performed for 20 epochs. The training and validation losses and accuracies as per figure 5 were monitored. The proposed model was evaluated using the test set for leaf classification and labeling.

## 4 RESULTS AND DISCUSSION

The figure 6 shows some sample results of the leaf classification model. The recognition accuracy of the model was calculated using the formula:

International Journal of Academic Information Systems Research (IJAISR)
ISSN: 2643-9026
Vol. 6 Issue 5, May - 2022, Pages: 9-13

$$Accuracy = 100 * \frac{number\ of\ labels\ predicted\ correctly}{total\ number\ of\ labels}$$

The model was able to classify the leaf images into their belonging class labels at an accuracy of 93% .

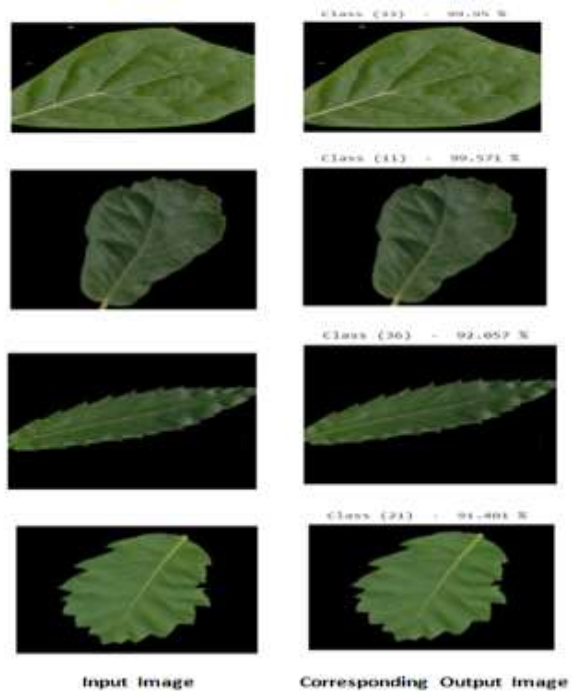


**Fig. 6.** *Sample labeled outputs of the proposed model*

The Confusion matrix is a graphical summarization of a classification model. The figure 7 below shows the confusion matrix of the proposed leaf classification model. The matrix shows a diagonal which indicates that the actual and derived output leaf classes are same. This proves that the model is highly accurate.

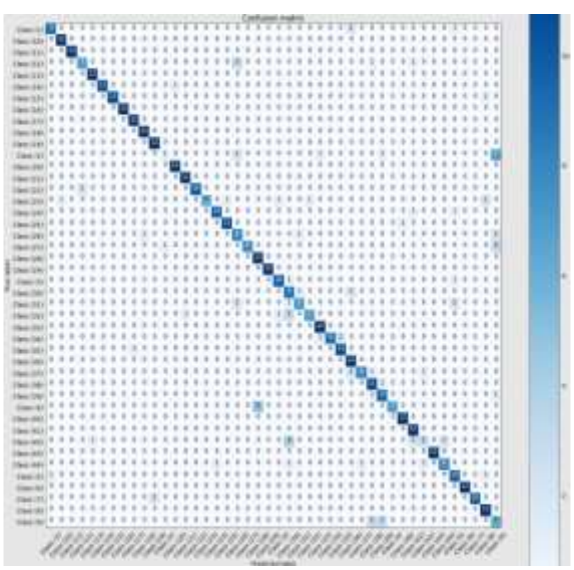

**Fig.7**.*Confusion Matrix of the proposed Leaf Classification Model*

## 5 Conclusions

The paper presents an automatic and intelligent deep learning based leaf image classification system. This system employs image augmentation and transfer learning of ResNet50 deep CNN model. It was found that augmentation categorically increased the number of leaf images in the dataset. This improved the training accuracy and eventually classification accuracy of the model. The model was tested on Malayakew Leaf dataset and could classify the leaf images very efficiently and accurately. The future scope of this work is to implement this model into a mobile application so that any leaf input image from the camera can be classified into a class and the output class can be shown to the user.

## 6 References

**[1]** Hall D, McCool C, Dayoub F, Sunderhauf N, Upcroft B. . Evaluation of features for leaf classification in challenging conditions. In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2015,797-804.

**[2]** Chandrakala H T, Thippeswamy G. Deep Convolutional Neural Networks for recognition of Historical Handwritten Kannada Characters . In: Suresh Chandra Satapathy, Vikrant Bhateja, Bao Le Nguyen, Nhu Gia Nguyen, Dac-Nhuong Le(Eds). Advances in Intelligent Systems and Computing . Springer, 2018, ISBN: 978-9811399190.

**[3]** Sue Han Lee , Chee Seng Chana , Simon Joseph Mayo , Paolo Remagninoc. How deep learning extracts and learns leaf features for plant classification. Pattern Recognition journal, Elsevier, 2017; 0031-3203.

**[4]** Beikmohammadi A, Faez K, Motallebi A. SWP-LeafNET: a novel multistage approach for plant leaf identification based on deep CNN. Expert Systems with Applications ,2022;117470.

**[5]** Bhanuprakash Dudi , V. Rajesh. Optimized threshold-based convolutional neural network for plant leaf classification: a challenge towards untrained data. Journal of Combinatorial Optimization ,Springer ,2021.

**[6]** Hedjazi M A, Kourbane I, Genc Y. On identifying leaves: A comparison of CNN with classical ML methods. In: 25th Signal Processing and Communications Applications Conference (SIU), IEEE,2017, 1-4.

**[7]** Bodhwani V, Acharjya D P, Bodhwani U. Deep residual networks for plant identification. Procedia Computer Science, 2019;152, 186-194.

**[8]** Hu J, Chen Z, Yang M, Zhang R, Cui Y. A multiscale fusion convolutional neural network for plant leaf recognition. IEEE Signal Processing Letters, 2018;25(6), 853-857.

**[9]** Sun Y, Liu Y, Wang G, Zhang H. Deep learning for plant identification in natural environment. Computational intelligence and neuroscience, 2017.

**[10]** Fati Mohamed, Amgad Muneer. Efficient and automated herbs classification approach based on shape and texture features using deep learning. IEEE Access,2020; 8:196747–196764.

**[11]** Liu J, Yang S, Cheng Y, Song Z . Plant leaf classification based on deep learning. In: Proceedings of the Chinese automation congress (CAC),2018, pp 3165–3169.

**[12]** Tavakoli H, Alirezazadeh P, Hedayatipour A, BanijamaliNasib AH, Landwehr N . Leaf image based classification of some common bean cultivars using discriminative convolutional neural networks. Computers and electronics in agriculture,2021; vol 18.

**[13]** Y. Zhang. Improved Leaf Image Classification Using Topological Features and CNN with Attention Module. In: Proceedings of the IEEE 6th International Conference on Signal and Image Processing (ICSIP), 2021, pp. 311-315.

**[14]** BinWang, HaoLi, JiaweiYou, XinChen, XiaohuiYuan, XianzhongFeng. Fusing deep learning features of triplet leaf image patterns to boost soybean cultivar identification . Computers and Electronics in Agriculture,2022; Volume 197, 106914.

**[15]** A Tapas. Transfer learning for image classification and plant phenotyping. Int. J. Adv. Res. Comput. Eng. Technol. 2016; vol. 5, no. 11, pp. 2664–2669.

**[16]** M. M. Ghazi, B. Yanikoglu, E. Aptoula. Plant identification using deep neural networks via optimization of transfer learning parameters. Neurocomputing,2017; vol. 235, pp. 228–235.

**[17]** J. W. Lee, Y. C. Yoon. Fine-grained plant identification using wide and deep learning model. In: Proceedings of the Int. Conf. Platform Technol. Service (PlatCon), 2019, pp. 1–5.

**[18]** M. A. Alajrami ,S. S. Abu-Naser. Type of tomato classification using deep learning. Int. J. Academic Pedagogical Res,2019; vol. 3, no. 12, pp. 21–25.

**[19]** A Ashqar, B. S. Abu-Nasser, S. S. Abu-Naser. Plant seedlings classification using deep learning. Int. J. Academic Inf. Syst. Res,2019; vol. 3, no. 1, pp. 7–14.

**[20]** A Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin, B. Tekinerdogan. Analysis of transfer learning for deep neural network based plant classification models. Comput. Electron. Agricult.2019; vol. 158, pp. 20–29.

**[21]** H. Vo, H. T. Dang, B. T. Nguyen, V.-H. Pham. Vietnamese herbal plant recognition using deep convolutional features. Int. J. Mach. Learn. Comput, 2019; vol. 9, no. 3, pp. 363–367.

**[22]** Amuthalingeswaran, M. Sivakumar, P. Renuga, S. Alexpandi, J. Elamathi, S. S. Hari. Identification of medicinal plants and their usage by using deep learning. In: Proceedings of the 3rd Int. Conf. Trends Electron. Informat. (ICOEI), 2019, pp. 886–890.

**[23]** A Patil K. S. Bhagat. Plant identification by leaf shape recognition—A review. Int. J. Eng. Trends Technol.2016; vol. 8, no. 35, pp. 359–361.

**[24]** Rawat W, Wang Z . Deep convolutional neural networks for image classification: a comprehensive review. Neural Computing, 2017; 29:2352–2449.

**[25]** Horaisova K, Kukal J . Leaf classification from binary image via artificial intelligence. Biosystems Engineering, 2016; 142:83–100.

**[26]** Keivani M, Mazloum J, Sedaghatfar E . Automated analysis of leaf shape, texture, and color features for plant classification. Int Inf Eng Technol Assoc, 2020; 37(1):17–2.

**[27]** Kamilaris A, PrenafetaBoldú F X. Deep learning in agriculture: A survey. Computers and electronics in agriculture, 2018; 147, 70-90.

**[28]** Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017; 1251-1258.

**[29]** Shreyan Ganguly, Pratik Bhowal, Diego Olivace, Ram Sarkard. BLeafNet: A Bonferroni mean operator based fusion of CNN models for plant identification using leaf image classification. Ecological Informatics journal, Elsevier,2022; Volume 69, 101585.

**[30]** Prabhjot Kaur, Vinay Gautam. Research patterns and trends in classification of biotic and abiotic stress in plant leaf. In: Proceedings of the Materials Today , 2021,Volume 45, Part 6, Pages 4377-4382.