Plant Identification in a Combined Imbalanced Leaf Dataset

[1]Dr.P. Pandi Selvi (MCA. M.Phil. PhD.), [2]Ms. A. Parameshwari

[1]Assistant Professor, [2]PG Student
Department of Computer Science
Mangayarkarasi College of Arts and Science for Women
Paravai, Madurai.

Abstract- **Plant identification has applications in ethnopharmacology and agriculture. Since leaves are one of the distinguishable features of a plant, they are routinely used for identification. Recent developments in deep learning have made it possible to accurately identify the majority of samples in five publicly available leaf datasets. However, each dataset captures the images in a highly controlled environment. This paper evaluates the performance of Efficient Net and several other convolutional neural network (CNN) architectures when applied to a combination of the Leaf Snap, Middle European Woody Plants 2014, Flavia, Swedish, and Folio datasets. To normalize the impact of imbalance resulting from combining the original datasets, the authors used oversampling, under sampling, and transfer learning techniques to construct an end-to-end CNN classifier. Emphasis is placed greater on metrics appropriate for a diverse-imbalanced dataset rather than stressing high performance on any one of the original datasets. A model from Efficient Net's family of CNN models achieved a highly accurate F-score of 0.9861 on the combined dataset.**

INTRODUCTION:

The one and only area that serves the food desires of the intact human race is the Agriculture zone. It has played a key responsibility in the development of human civilization. Plants exist all over the place, we live as well as places without us. Plant disease is one of the essential causes that reduces quantity and degrades quality of the agricultural merchandises. Images form important data and information in biological sciences. Until recently photography was the only method to reproduce and report such data.

It is difficult to quantify or treat the photographic data mathematically. Digital image processing and image analysis technology based on the advances in microelectronics and computers circumvent these problems associated with traditional photography. Digital image processing is the use of computer algorithms to perform image processing on digital images. As a subcategory or field of digital signal processing, digital image processing has many advantages over analogy image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and signal distortion during processing. Since images are defined over two dimensions (perhaps more) digital image processing may be modelled in the form of multidimensional systems.

Using this new tool helps to improve the images from microscopic to telescopic range and also offers a scope for their analysis. It, therefore, has many applications in biology. However, as is the case with any new technology, imaging technology also has to be optimized for each application, since what each user is looking for in an image is quite unique. Images of the leaves, captured by a camera or a scanner for Colour image analysis for estimation of normal leaf, infected leaf and chlorophyll. Many times, a viral or a fungal attack on plants results in degradation of chlorophyll pigments in leaves. Such infected leaves have patches of green and yellow colour. In plant breeding, it is important to quantify the leaf infection. In this paper, the authors develop a software for the automatic identification & classification of plants based on leaves. Here the end-user is the farmer. It, classifies the plant leaves and stems at hand. The developing software provides a fast and accurate method in which the leaves detected and classified using k-means based segmentation and neural networks-based classification.

LITERATURE REVIEW:

Plant species identification based on leaf images has been an active area of research [9], [10]–[1]. Kumar et al. [4] used features related to a leaf's curvature to identify species in the proposed Leaf Snap dataset and achieved a top-5 accuracy of 96.8%. Fourier descriptors for leaf contours combined with the nearest-neighbour classifier were used to classify leaf images in the proposed MEW 2012 dataset with an accuracy of 88.91% [6]. Wu et al. [11] proposed the Flavia dataset with 32 different species of plants and 12 leaf features in combination with a probabilistic neural network (PNN) to achieve a classification accuracy of 90%. Features based on the geometry, eigen leaves, and grey level co-occurrence matrix (GLCM) were used to train a support vector machine (SVM) classifier for leaf identification by [2].

Munisami et al. [5] used shape features and colour histogram with k-nearest neighbours to classify plant leaves in the Folio dataset with an accuracy of 87.3%.

Kumar et al. [3] used morphological features extracted using a multilayer perceptron with adaboost to train a classifier and attained an accuracy rate of 95.42% on the Flavia dataset.

The DDLA method [8] achieved the highest accuracy of 98.71%, 96.38%, and 99.41% on the Flavia, Folio, and Swedish datasets, respectively.

METHODOLOGY:
Various steps involved in the proposed methodology is as follows,
1. Image
2. Segmentation
3. Splitting Dataset into Train and Test Data
4. Classification
5. Performance analysis
6. Result and discussion

Modules Description
 Data Collection:
The data selection is the process of selecting the data for Plant Leaf detection from the Image dataset.

PRE-PROCESSING:
Image Resize:
 In computer graphics and digital imaging, scaling refers to the re-sizing of a digital image. In video technology, the magnification of digital material is known as up-scaling or resolution enhancement. When scaling a vector graphic image, the graphic primitives which make up the image can be scaled using geometric transformations, without any loss of image quality. When scaling a raster graphics image, a new image with a higher or lower number of pixels must be generated.   In the case of decreasing the pixel number (scaling down) this usually results in a visible quality loss. From the standpoint of digital signal processing, the scaling of raster graphics is a two-dimensional example of sample rate conversion, the conversion of a discrete signal from a sampling rate (In this case the local sampling rate) to another.

SPLITTING DATASET INTO TRAIN AND TEST DATA:
Data splitting is the act of partitioning available data into two portions, usually for cross-validation purpose.
One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance. Separating image data into training and testing sets is an important part of evaluating image processing models.
Typically, when one separates a data set into a training set and testing set, most of the image data is used for training, and a smaller portion of the data is used for testing.

CLASSIFICATION:
In Deep learning, a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyse visual image. It uses a special technique called Convolution.
In mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.

PERFORMANCE ANALYSIS:
• Estimations
• True positive (TP) = the number of cases correctly identified as patient.
• False positive (FP) = the number of cases incorrectly identified as patient.
• True negative (TN) = the number of cases correctly identified as healthy.
• False negative (FN) = the number of cases incorrectly identified as healthy.
• Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, one should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as: Accuracy = (TP+TN)/ (TP+TN+FP+FN);
 • Sensitivity: The sensitivity of a test is its ability to determine the patient cases correctly. To estimate it, we should calculate the proportion of true positive in-patient cases. Mathematically, this can be stated as: Sensitivity = (TP) / (TP + FN)
• Specificity: The specificity of a test is its ability to determine the healthy cases correctly. To estimate it, we should calculate the proportion of true negative in healthy cases. Mathematically, this can be stated as: Specificity = (TN) / (TN + FP)

RESULT AND DISCUSSION:
     The results obtained while testing the trained networks before and after performing oversampling and under sampling is as given below,
1)   EFFECT OF OVERSAMPLING AND UNDER SAMPLING:
     Comparisons between all trained models show that balancing the dataset improves both Top-1 accuracy and F-score; however, the improvement is much more significant in the Top-1 accuracy metric than the F-score.

Network	Before Balancing		After Balancing	
	Top-1 Accuracy	F-score	Top-1 Accuracy	F-score
MobileNetV2 [33]	0.9190	0.9400	0.9327	0.9416

DenseNet121 [31]	0.9455 0.9586	0.9618 0.9663
Xception [32]	0.9475 0.9657	0.9731 0.9753
Inception V3 [30]	0.9599 0.9642	0.9696 0.9719
Resnet50V2[13]	0.9377 0.9537	0.9591 0.9624
EfficientNet B0 [15]	0.9335 0.9601	0.9590 0.9639
EfficientNet B1 [15]	0.9433 0.9618	0.9646 0.9697
EfficientNet B2 [15]	0.9462 0.9651	0.9683 0.9727
EfficientNet B3[15]	0.9480 0.9681	0.9751 0.9778
EfficientNet B4 [15]	0.9506 0.9698	0.9771 0.9787
EfficientNet B5 [15]	0.9540 0.9728	0.9796 0.9811
EfficientNet B6 [15]	0.9593 0.9759	0.9841 0.9861

TABLE 1. Comparison of experimental results before and after performing oversampling and under sampling on the dataset.

2) PERFORMANCE ANALYSIS

The models used in this paper were trained efficiently. If computational complexity is a concern, then models such as Mobile Net and Efficient Net B0 can be trained to use considerably fewer parameters and still provide high accuracy and F-scores. Fig. 1 summarizes the number of parameters versus the F-score efficiency.
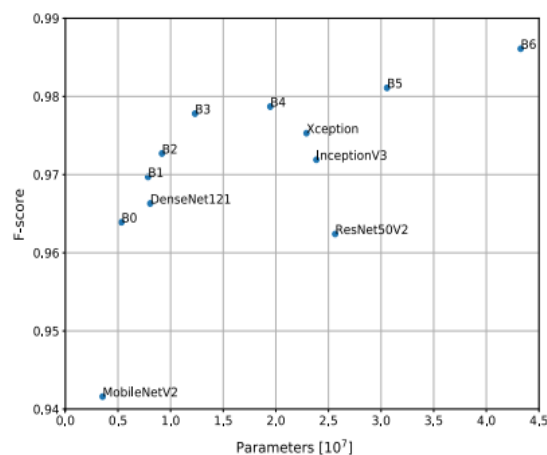

FIGURE 1. Number of model parameters vs. mean F-score. The plot shows that Efficient Net family of models perform better than other models for leaf identification in F2LSM dataset.

3) GRAD-CAM VISUALIZATION

In order to obtain gradient-weighted class activation mapping (Grad-CAM) visualization for any given class of image, the image is forward propagated through the CNN part of the model to obtain a raw score for the class. The gradient of the desired class is set to 1, and the remaining classes are set to 0. This signal is then backpropagated through the CNN to compute a heat map, which shows where the model had to look before making a decision. This heat-map then can be superimposed on the original image to create a Grad-CAM visualization [7]. Fig.2 shows Grad-CAM visualizations for a leaf using the following four models from left to the right: Mobile Net V2, Efficient Net B0, Exception, and Efficient Net B6. It can be observed from the Grad-CAM visualizations that as either the number of parameters or the input image size increases, models look at more details in an image before making a decision. The MobileNetV2 model primarily looks at the vein structure close to the midrib in the middle portion of the leaf, whereas Efficient Net B6 looks at significantly more detailed features such as detailed vein structure and the edge of the leaf before making a decision. This detailed feature extraction by the Efficient Net B6 model somewhat explains its superior performance compared to other models used in this paper.
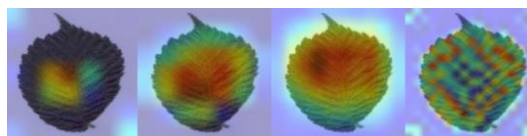
FIGURE 2. Grad-Cam visualizations of a leaf image for different networks. From left to right, MobileNetV2, Efficient Net B0, Exception, and Efficient Net B6, networks examine more detailed features as the number of parameters increase.

4) COMPARISONS AND BENCHMARKS

The performance of the CNN architectures on the F2LSM dataset was bench-marked by comparing its performance with the accuracy obtained on the individual datasets used to create the new dataset, as shown in Table 2. It can be observed from the table that the Top-1 accuracy generally drops as the number of classes increases. Even though F2LSM is a combined dataset with 374 classes (more than twice of Leaf Snap), Efficient Net B6 achieves comparable accuracy compared to the models trained on the individual datasets.

Dataset	Number of classes	Model	Top-1 accuracy (%)
Swedish [23]	15	Dual deep learning architecture [19]	99.41 [19]
Folio [24]	32	Google Net [11]	99.42 [46]
Flavia [22]	32	Extreme learning machines [47]	99.10[47]
MEW 2012 [21]	153	SVM [48]	96.54[49]
Leaf Snap [20]	185	Siamese CNN [50]	96.00[50]
F2LSM	374	Efficient Net B6 [15]	98.41

TABLE 2. State-of-the art accuracies on different datasets.

CONCLUSION:

This paper combines five publicly available leaf datasets into one F2LSM dataset. The combined dataset is highly imbalanced due to some of the individual datasets imbalanced nature, some classes having very few image samples, and certain classes overlapping across different datasets while combining. The authors used oversampling and under sampling to mitigate the imbalance in the dataset. TL is then used to train several CNN architectures for plant species identification using leaf images in the F2LSM dataset and tested their performance using metrics such as precision, recall, and F-score, considering the imbalanced nature of the combined dataset. Efficient Net B6 achieved comparable accuracy on the F2LSM dataset when compared to the state-of-the-art accuracy on the individual datasets.

Future work may include further expansion of the dataset by including more plant species and using leaf images from other applications such as weed identification, plant phenotyping, and identification of leaf diseases and pests. Also, plant identification for occluded leaves in a field would be an interesting problem.

REFERENCES:

[1] O. H. Babatunde, L. Armstrong, J. Leng, and D. Diepeveen, ''A survey of computer-based vision systems for automatic identification of plant species,'' J. Agricult. Informat., vol. 6, no. 1, pp. 61–71, Jan. 2015.

[2] V. Gajjar, Z. H. Lai, and K. Kosbar, ''Fast classification of leaf images for agriculturalremotesensingapplications,''inProc.Int.TelemeteringConf., Glendale, AZ, USA, vol. 54, Nov. 2018, pp. 569–578

[3] M. Kumar, S. Gupta, X.-Z. Gao, and A. Singh, ''Plant species recognition using morphological features and adaptive boosting methodology,'' IEEE Access, vol. 7, pp. 163912–163918, 2019.

[4] N. Kumar, P. N. Belhumeur, A. Biswas, D. W. Jacobs, W. J. Kress, I. C. Lopez, and J. V. B. Soares, ''Leafsnap: A computer vision system for automatic plant species identification,'' in Proc. Eur. Conf. Comput. Vis. Florence, Italy: Springer, 2012, pp. 502–516.

[5]T.Munisami,M.Ramsurn,S.Kishnah,and   S.Pudaruth,''PlantleafrecognitionusingshapefeaturesandcolourhistogramwithK-nearestneighbour classifiers,'' Proc. Comput. Sci., vol. 58, pp. 740–747, Jan. 2015.

[6] P. Novotný and T. Suk, ''Leaf recognition of woody species in central Europe,'' Biosyst. Eng., vol. 115, no. 4, pp. 444–452, 2013.

[7] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, ''Grad-CAM: Visual explanations from deep networks via gradient-based localization,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 618–626.

[8] A. P. Sundara Sobitha Raj and S. K. Vajravelu, ''DDLA: Dual deep learning architecture for classification of plant species,'' IET Image Process., vol. 13, no. 12, pp. 2176–2182, Oct. 2019.

[9] J. Wäldchen and P. Mäder, ''Plant species identification using computer visiontechniques: Asystematicliteraturereview,''Arch.Comput.Methods Eng., vol. 25, no. 2, pp. 507–543, 2018.

[10] J. Wldchen, M. Rzanny, M. Seeland, and P. Mäder, ''Automated plant speciesidentification— Trendsandfuturedirections,''PLoSComput.Biol., vol. 14, no. 4, 2018, Art. no. e1005993.

[11] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, ''A leaf recognition algorithm for plant classification using probabilistic neural network,'' in Proc. IEEE Int. Symp. Signal Process. Inf. Technol., Dec. 2007, pp. 11–16.