IJCSE International Journal of Computer Sciences and Engineering Open Access
Technical Paper Vol.8, Issue.4, Apr 2020 E-ISSN: 2347-2693

An Ensemble Deep Learning Technique for Plant Identification

P. Siva Prasad[1*], A. Senthilrajan[2]

[1] Dept. of Computer Science, Alagappa University,Karikudi- 630 003, Tamil Nadu, India
[2]Dept. of Computational Logistics, Alagappa University, Karikudi -630 003, Tamil Nadu, India

DOI: https://doi.org/10.26438/ijcse/v8i4.133135 | Available online at: www.ijcseonline.org

Received: 24/Feb/2020, Accepted: 19/Mar/2020, Published: 30/Apr/2020

***Abstract*:** Plant identification system is helped to find unidentified plants. Plant identification is most difficult task with the existing classification algorithms. Many existing classifiers are present to identify the plant species with the help of leafs. With the various drawbacks, the system will not reach that much. In recent years, many applications belong to various domains and technologies are using the Deep Learning (DL) for rapid and better results. In this paper, the Novel Approach (NA) is introduced with the combination of CNN adopted with ensemble methods such as bagging and boosting. This paper addresses that the Convolutional Neural Network (CNN) with ensemble methods is better than Machine Learning methods to identify the plant by leaf. The ensemble methods are to improve the accuracy and sensitivity of plant identification model. The parameters such as sensitivity and accuracy are the two metrics to show the performance.

***Keywords*:** *CNN, Bagging, Boosting, Novel Approach.*

I. INTRODUCTION

Plant identification system developed by researchers has helped to find unidentified plants more rapidly. Plant identification is important to agricultural development. Plants play an important in environmental protection. There are about three million types of plants in world. Only some of the plants are identified. Many of the plants are still unidentified. It is difficult to identify all the plants manually. Hence, there is an essential to develop an automated or computerized. Plant identification System is to recognise and classify plants. The method of plant identification can be achieved using the plants leaves. Each plant has a unique leaf. The leaves of the plants bring a lot of data about the plants. This data is used to develop plant identification model. Deep learning is a dedicated form of machine learning. Deep learning methods are able to extract more detailed data as compared to the machine learning methods. The machine learning approaches are used to find and categorize plants with manually extracted features from leaf images. With the evolution of Information Technology, We can simply transfer the leaf image to a computer and then the computer can extract necessary features automatically using Deep learning approaches. So, with deep learning workflow, applicable features are automatically extracted from leaf images. In deep learning, one of the most popular methods is convolutional neural network (CNN). This method eliminates the need of manual feature extraction. CNN is used to extract the leaf features automatically from the leaf images.

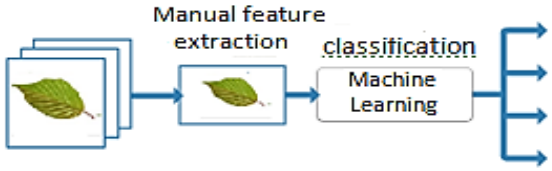

Fig.1:Work Flow of Machine Learning

Fig.2: Work Flow of Deep Learning

II. RELATED WORK

This section is deeply discussed on various DL techniques that are used in identification of plat species. Cope et al. [1] started his research on the classification of a vein based on traditional algorithms such as genetic algorithms (GA) and Ant Colony algorithms (AOC) to get the structure of vein. The developed vein is most compatible with getting a similar and sum of first and second vein patterns with only a very low noise.

Features such as color and texture combination based on plant identification system had been proposed by Anami et al. [2]. Another operator is used to extract the color and edge histogram as color and texture features respectively from 1000 images of different types of herbs, shrubs, and trees which is called Sobel. These are trained by the radial basis exact fit neural network (RBENN) and a SVM.

III. DRAWBACKS OF EXISTING SYSTEM

- For the existing systems processing time will take more.
- Only simple datasets are compatible for existing algorithms.

IV. NOVEL APPROACH WITH CNN

In this paper, the novel approach (NA) is introduced with the combination of CNN adopted with ensemble methods such as bagging and boosting.

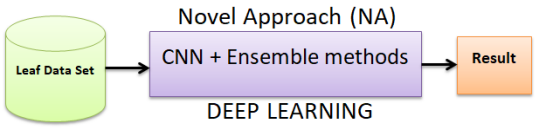

Fig.3: Proposed Architecture

In my research, Ensemble methods are used for handling of huge complicated datasets. Novel Approach is the key stage in proposed system. This stage is used to improve the accuracy and sensitivity

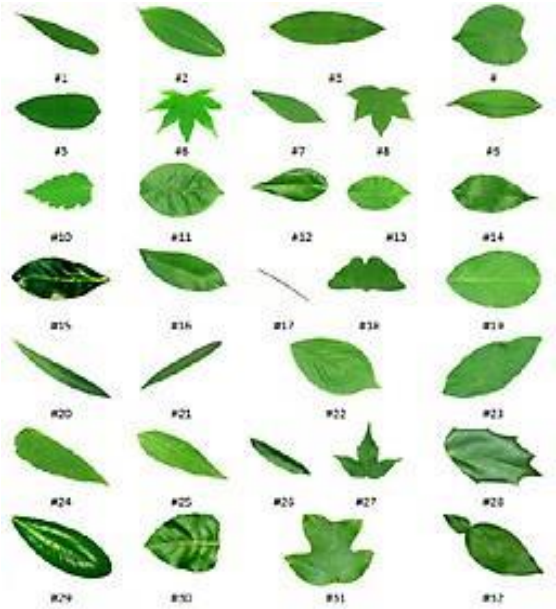

Fig.4

Ensemble methodology

Ensemble is Deep learning concept that various models are educated by the identical learning algorithm. Ensemble methodologies combine some decision trees classifiers to create enhanced predictive performance than a single decision tree classifier. The source behind the ensemble is that an assembly of weak learners come together to form a robust learner, hence increasing the accuracy of the model. This is typically used to improve the accuracy of the outcomes. Therefore, Ensemble methods are learning methods that achieve accurate performance. Ensemble is a learning procedure for merging several weak models to create a robust model with the idea of Bagging and Boosting. Bagging is a way to descent the alteration in the prediction by making extra data for exercise from dataset using groupings with duplications to produce multi-sets of the novel data. Boosting is a repetitive procedure that changes the weight of an opinion based on the earlier classification. The idea of boosting came out of the idea of whether a weak learner can be improved to become better.

In this paper, Gradient Boosting is used to learn from the residual error, rather than update the weights of data points. Gradient boosting contains three features: a loss function to be improved, a weak learner to mark prediction, and a preservative model to enhance weak learners to reduce the loss function. So, The Gradient Boosting creates a novel prediction by just adding up the predictions.

V. PERFORMANCE EVOLUTION

Accuracy and Sensitivity mostly show the performance of the proposed system. Novel Approach (CNN adopted with ensemble methods) is used to improve the accuracy and sensitivity in higher levels compared with the traditional algorithms.

The performance growth is done with the numerous parameters such as FPR, FNR, Sensitivity and Accuracy. The traditional count values such as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are used through these processes.

False Positive Rate (FPR)

The fraction of cases where an image was separated into normal images, such as

$$FPR = \frac{FP}{FP + TN}$$

False Negative Rate (FNR)

The proportion of cases where an image was categorized to irregular images, with

$$FNR = \frac{FN}{FN + TN}$$

Sensitivity

The proportion of original positives which are correctly identified is the measure of the sensitivity. It relates to the ability of the test to identify positive results.

$$Sesitivity = \frac{No.of\ TP}{No.of\ TP + No.of\ TN}$$

Accuracy: This will calculate the overall accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The proposed system predicted 119 out of 140 observations, giving accuracy of 96.54% and sensitivity of 95.7%.There were 21 cases that system got wrong. These 21 cases were separated between false negatives and false positives as 7 and 14 respectively

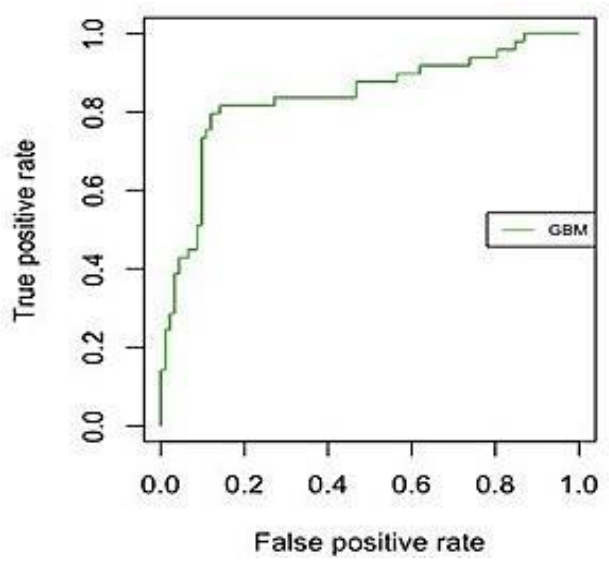

Fig.5

V1. ADVANTAGES OF PROPOSED SYSTEM

1. Accuracy and Sensitivity more compare with various existing systems.
2. More compatibility for the many datasets.

VII. EVALUATION RESULTS

The evolution of this implementation is finished with R-language and several packages are used to get the better results. For the research point of view R-language is most widely used for many applications based on research. The results have proved a higher degree of accuracy and sensitivity compared to the existing methods. The ensemble methods play a prominent role in the performance.

Table**1**: Performance Evaluation

Classifiers	Sensitivity	Accuracy
SVM	76%	82.75
ANN	87%	94.88
CNN	71%	79.03
NA	95.7%	96.54

VIII. CONCLUSION

In this paper, the proposed system performs well compare with the various existing algorithms. Various datasets are used to extract the features to shows the performance. The novel approach (NA) is used the ensemble methods to get the better accuracy compare with the existing methods. Metrics such as sensitivity and accuracy are shown in this paper. This work addresses and provides better solutions for plant classification using Deep Learning Methods. Experimental result shows that Gradient Boosting Algorithm is feasible with accuracy and sensitivity greater than 95% on 117 types of plants. The experimental results indicate that the proposed approach is a valuable approach which can significantly support an accurate detection of plant in a little computational effort. Compared with existing methods, proposed Algorithm is fast in performance, efficient in recognition and easy in implementation.

REFERENCES

[1] Cope, J. S., Remagnino, P., Barman, S., & Wilkin, P. (2010, December). The extraction of venation from leaf images by evolved vein classifiers and ant colony algorithms. In International Conference on Advanced Concepts for Intelligent Vision Systems (pp. 135-144). Springer Berlin Heidelberg.

[2] Anami, B. S., Suvarna, S. N., & Govardhan, A. (2010). A combined color, texture and edge features based approach for identification and classification of indian medicinal plants. International Journal of Computer Applications,6(12), 45-51.

[3] A. Aakif, M. F. Khan, "Automatic classification of plants based on their leaves", Biosyst. Eng., vol. 139, pp. 66-75, Nov. 2015.

[4] Go¨eau, H., Bonnet, P., Joly, A.: Plant identification in an open-world (lifeclef 2016). In: CLEF working notes 2016. (2016)

[5] J. Wäldchen, P. Mäder, "Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review" in Arch. Comput. Methods Eng., pp. 1-37, Jan. 2017.

Authors Profile

Mr.P.Siva Prasad is currently Research Scholar of Computer Science Department in Alagappa University. His Research interests are Artificial Intelligence, Data Mining and Image Processing and Deep Learning

Dr.A.Senthilrajan is Professor and Director of the Department of Computational Logistics in Alagappa University. His Research interests are Image Processing, Networks, Artificial Intelligence and Deep Learning.

