IJARSCT

ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 4.819

Volume 2, Issue 2, February 2021

Leaf Features Extraction for Plant Classification using CNN

Mr. P. Siva Prasad[1] and Dr. A. Senthilrajan[2]
Research Scholar, Department of Computer Science[1]
Professor and Director, Department of Computational Logistics[2]
Alagappa University, Karikudi, Tamil Nadu, India

Abstract: *Deep learning is now an active research area. Deep learning has done a success in computer vision and image recognition. It is a subset of the Machine Learning. In Deep learning, Convolutional Neural Network (CNN) is popular deep neural network approach. In this paper, we have addressed that how to extract useful leaf features automatically from the leaf dataset through Convolutional Neural Networks (CNN) using Deep Learning. In this paper, we have shown that the accuracy obtained by CNN approach is efficient when compared to accuracy obtained by the traditional neural network.*

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Feature Extraction, Convolutional Neural Network (CNN), Artificial Neural Network (ANN), leaf

I. Introduction

Each leaf has its individual features and carries significant information that can help people to recognise the plant. In medical perspective, images have been used by doctors to diagnose diseases. Using the same method as doctors, researchers try to simulate the same principle to recognise a plant using high quality leaf images. The persons who study plants can simply find the plants using the features of the leaf. Leaf of different plants has different features that can be used to identify the plants. The extraction of leaf features is helped to identify the plants. Many research works have been completed on the extraction of features from the leaf dataset. Most of the proposed approaches started the process with leaf image pre-processing. Many techniques have been implemented during pre-processing; for instance, to convert the RGB image to grey scale, transform the grey scale image to binary image and the use image enhancement. The purpose of implementing these processes is to minimize the noise in the image that can disturb the extraction and classification process. In this paper, we have focused only on the leaf feature extraction. We can simply transmission the leaf image to a system and then the system can extract essential features automatically and then can recognize the plant leaf using deep learning. Deep learning is itself a self-learning system used on huge amounts of data. Deep learning (DL) is subset of machine learning (ML) and also machine learning is a subset of artificial intelligence (AI).

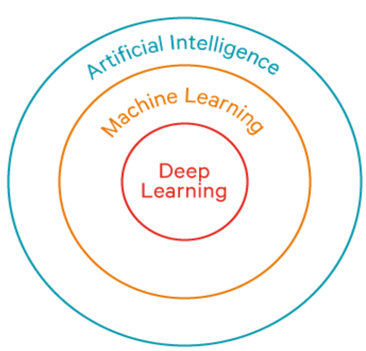


Figure 1: Relationship between AI, ML and DL

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 4.819
Volume 2, Issue 2, February 2021

In Machine learning, leaf features are mined manually and the classification algorithm categorises the leaf images separately. But, In Deep Learning, the network itself extracts the leaf features and also uses that leaf features to classify the leaf images.

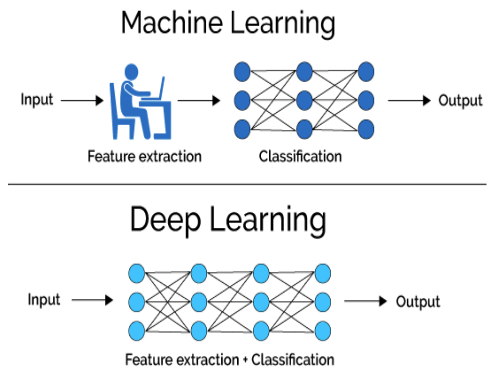


Figure 2: Workflow of ML and DL

1.1 Structure of a Leaf

A leaf has three main parts. These are petiole, leaf base and leaf blade. Each can perform specific task.

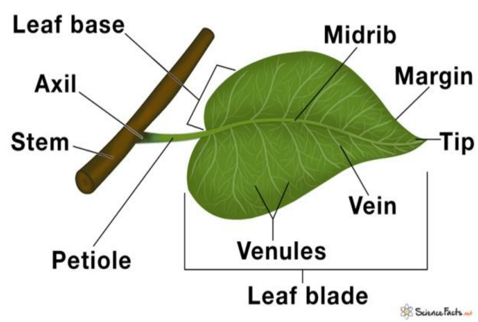


Figure 3: Leaf external structure

1.2 Leaf Taxonomy

Leaves are different from plant to plant. Below figure shows types of leaves:

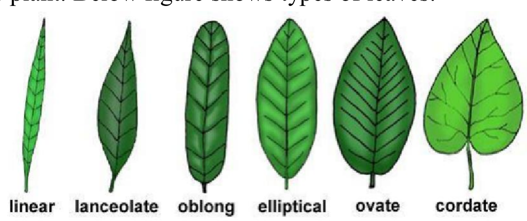


Figure 4: Common leaf types.

II. Deep Learning Approach

In deep learning, convolutional neural network (CNN) is an approach. This approach eliminates the manual extraction of leaf features. CNN can extract the leaf features automatically from the leaf images. A Convolutional Neural Network (CNN) is the foundation of most computer vision technologies. CNN is popular due to its automatic feature extraction for image classification involving large datasets.CNN uses two operations called 'convolution' and pooling' to reduce an image into its essential features, and uses that features to understand and classify the image.

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
IJARSCT
Impact Factor: 4.819
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2021

2.1 How to Store Leaf Images in the Machine?

We need to understand how a machine can read and store leaf images. Loading the leaf image, read them and then process them through the machine is difficult. Any computer can maintains and represent any images in the form of a matrix of numbers. The size of this matrix based on the number of pixels of the input leaf image. The pixel values for each of the pixels describe how bright that pixel is, and/or what colour it should be. So, in the simplest case of the binary images, the pixel value is a 1-bit number indicating either foreground or background. So pixels are the numbers, or the pixel values which denote the intensity or brightness of the pixel.

III. Proposed Work

3.1 Leaf Features Extraction through CNN

The objective of Feature Extraction is to reduce the number of features in a leaf dataset by creating new features from the existing ones. These new reduced set of features should then be able to summarize most of the information contained in the original set of features. The process of feature extraction is useful when you need to reduce the number of resources needed for processing without losing important or relevant information. Feature extraction can also reduce the amount of redundant data for a given analysis.CNN can extract the leaf features automatically through convolution layer and pooling layer. The Convolution layer can extract features from a leaf image. Convolution preserves the relationship between pixels by extracting leaf image features by small squares of input data. It is a mathematical procedure that takes the two inputs such as image matrix and a filter.

- Dimension of the leaf image matrix is **h×w×d**
- Dimension of the filter is **fh×fw×d.**
- Dimension of the output is **(h-fh+1)×(w-fw+1)×1**

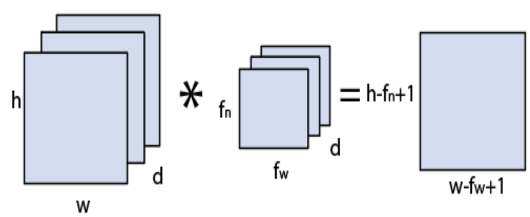


Image matrix multiplies kernl or filter matrix

Figure 5: Image Matrix Multiplies Filter Matrix

Let us start with consideration a 5*5 image whose pixel values are 0, 1, and filter matrix 3*3 as:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

5 × 5 − Image Matrix   3 × 3 − Filter Matrix

The convolution of 5*5 image matrix multiplies with 3*3 filter matrix is called "Features Map" and display as an output.

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 4 \\ 2 & 4 & 3 \\ 2 & 3 & 4 \end{bmatrix}$$

Convolved Feature

DOI: 10.48175/IJARSCT-807

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 4.819
Volume 2, Issue 2, February 2021

Convolution of an image through different filters can perform an operation such as blur, sharpen, and edge detection by applying filters.

Pooling layer reduces the amount of information in each feature obtained in the convolutional layer while maintaining the most important information Pooling layer shows an important role in pre-processing of a leaf image. Pooling layer shrinks the no. of parameters when the leaf images are too big. Pooling is "downscaling" of the leaf image found from the previous layers. It can be compared to shrinking a leaf image to reduce its pixel density. Spatial pooling is called down sampling, which reduces the dimensionality of each map but preserves the important information. Max pooling is a discretization procedure. Its objective is to downscale an input exemplification, sinking its dimensionality and permitting for the statement to be completed about features contained in the sub-region binned. Max pooling is done by applying a max filter to non-overlapping sub-regions of the initial representation.

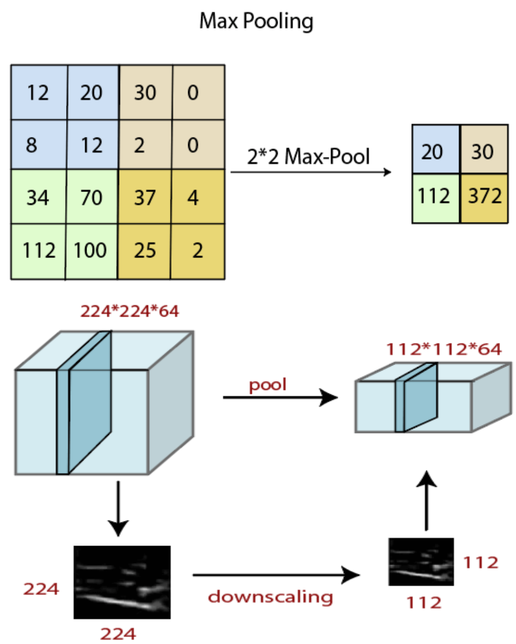


Figure 6: Max Pooling

So, the convolution layer and pooling layer extract the features automatically through CNN approach. We have extracted leaf features through convolutional layer and pooling layer and then these leaf features as an input are given to fully connected layer. The objective of a fully connected layer is to take the results of the convolution/pooling process and use them to classify the image into a label

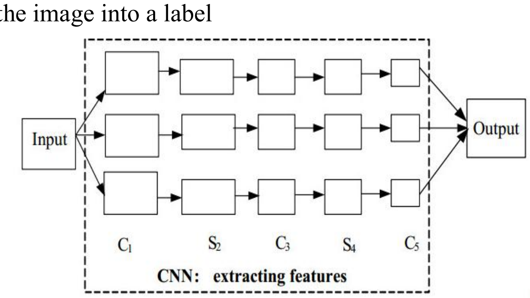
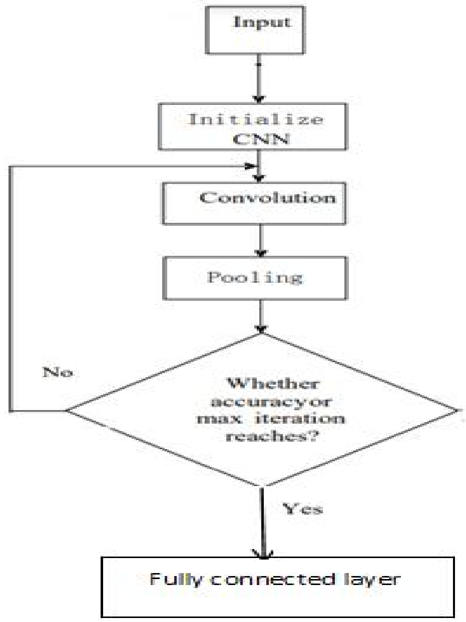


Figure 7: Workflow of Convolutional and Pooling

In CNN, input is the leaf image, C1 and C3 are the convolutional layers, S2 and S4 are the pooling layers and C5 is fully connected layer.

3.2 Flow Diagram of CNN for Test the Accuracy



3.3 Flavia Leaf Dataset for experimentation

The Flavia leaf dataset has selected as it contains huge number of samples for each leaf group. Deep learning needs huge datasets, so it is suitable to select Flavia dataset. Flavia leaf dataset has more than 1800 leaf images from more than 36 different kinds of plants. This amounts to an average of 50 images per category.

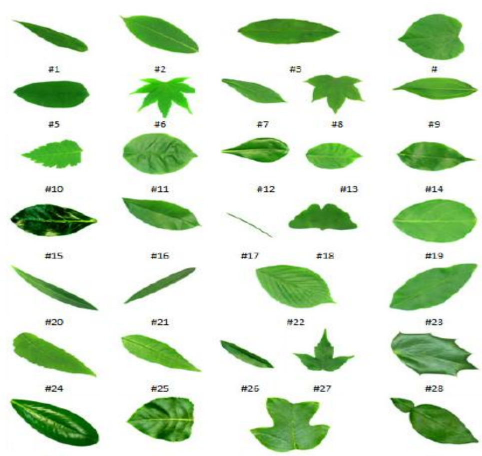


Figure 8: Leaves in Flavia dataset

Each leaf image was taken with a high-resolution camera on an identical white background.

3.4 Experimental Setup

The deep learning system has implemented through Python running on Anaconda Navigator implementing Jupyter Notebook. OpenCV image processing has also used to make modifications to the leaf images. The specification of the computer is of very limited use with deep learning models, as they are computationally intensive as to require Graphical Processing Units. Jupyter notebooks are a great way to run deep-learning experiments.

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 4.819
Volume 2, Issue 2, February 2021

IV. RESULT ANALYSIS

In this paper, we have extracted 5 features that can be used to recognise a leaf and those features are basic geometric features that are diameter, physiological length, physiological width, leaf area and leaf perimeter and also extracted 12 digital morphological features that have been defined based on the basic geometric features

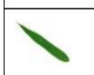
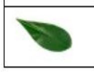
Leaf/Feature	Diameter	Physiological Length	Physiological Width	Leaf Area	Leaf Perimeter
	877.4058	766.5445	523.2447	7.0955e+004	1.7424e+003
	791.0209	529.3611	534.0206	2.2625e+005	1757
	751.8045	423.8629	531.7377	1.1498e+005	4523e+003
	842.0665	806.2175	385.5412	204990	1.9344e+003

Table 1: Results after extracting geometrical features

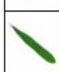
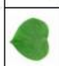
Leaf\Feature	Smooth Factor	Aspect Ratio	Form Factor	Rectangularity	Narrow Factor	Perimeter ratio of Diameter	Perimeter ratio of L_P and W_P
	0.9915	1.4650	0.2937	5.6527	1.6769	1.9858	1.3509
	0.9977	0.9913	0.9210	1.2494	1.4813	2.2212	1.6523
	0.9947	0.7971	0.1212	0.9601	1.4139	4.5920	3.6126
	0.997	2.0911	0.6884	1.5163	2.1841	2.2972	1.6231

Table 2: Result after extracting morphological features

The CNN approach of deep learning play a prominent role in the performance.The performance growth obtained by CNN approach of deep learning is better, which is efficient when compared with performance growth obtained by traditional machine learning approaches.

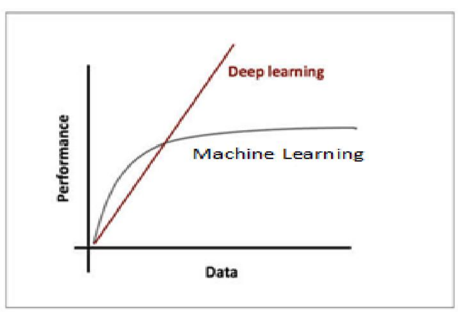


Figure 9: Performance growth of ML and DL

In this work, we have also observed a significant difference in test classification accuracy when the same set of features is classified through CNN and when classified through ANN. when the both methods have used soft max at the last layer(fully connected layer), we have observed a small difference in accuracies. With CNN, we have achieved the accuracy score is 98.3%.With ANN, we have achieved the accuracy score is 95.7% the evolution is finished by Python and different python packages are used to get the well results. For the research point of view, Python is used for many applications. In this paper, the accuracy obtained by CNN approach is 98.3%, which is efficient when compared to accuracy obtained by ANN.

Approach	Accuracy
ANN	95.7%
CNN	98.3%

Table 3: Approaches and its accuracy

V. CONCLUSION

In this paper, a deep learning approach (CNN) of deep learning for automatically extracting features of leaf is presented. The CNN performs well compare with the various existing approach. Flavia Leaf datasets is used to extract the features to shows the performance. This work addresses and provides better solutions for extracting the leaf features using Deep Learning approach. Experimental result shows that CNN is feasible. The experimental results indicate that the deep learning approach is a valuable approach which can support an accurate of leaf features extraction in a little computational effort. Compared with machine learning methods, CNN approach of deep learning is fast in performance, efficient in extraction and easy in implementation. In this paper, the accuracy obtained by CNN approach is efficient when compared to accuracy obtained by traditional neural network.

REFERENCES

[1]. Shabanzade, M.M.Zahedi and S.A.A ghvami, 2011. Combination of local descriptors and global features for leaf recognition signal and image processing. International Journal, vol. 2 : 23-21. DOI : 10 .5121/SIPIJ

[2]. Ji-Xiang Du, Xiao-FengWang,Guo-Jun Zhang. “Leaf Shape based plant species recognition”, Applied Mathematics abd Computation 185 (2007) 883-893

[3]. CholhongIm, Hirobumi Nishida, Tosiyasu L. Kunii .“ Recognizing Plant Species by Normalized Leaf Shapes”. Vision Interface ’99, Trois-Rivieres,Canada, 19-21 May

[4]. Najjar, A. and E.Zagrouba, 2012. Flower image segmentation based on color analysis and a supervised evaluation. Proceddings of the International Conference on Communications and Information Technology (ICCIT) , Jun. 26-28, IEEE Xplore Digital Library, pp : 397-401. DOI : 10.1109/ICCITechnol.2012.6285834

[5]. N. Valliammal and S.N. Geethalakshmi, (2011). Automatic Recognitin System Using Preferential Image Segmentation for Leaf and Flow.

[6]. Stefan Fiel and Robert Sablantnig .”Leaf classification using local features.”

[7]. Abdul Kadir,Lukito Edi Nugroho, AdhiSusanto and Paulus InsapSantosa , “ Foliage plant retrieval using polar fourier transform, color moments and vein features”, An international Journal(SIPIJ) Vol. 2, No.3, September 2011.

[8]. Du. J.X ., Huang, D.S.Wang,X.F., &Gu, X.(2006).: “Computer –aided Plant Species Identification (CAPSI) Based on Leaf Shape Matching Technique”. Transactions of the Institute of Measurement and Control, 28(3),pp.275-284.

[9]. Jing, L.Z.Shawan and L.Jiandu , 2009. “A method of Plant Leaf Recognition based on Locally Linear Embedding and Moving CenterHypersphere Classifier”. Procedding of Intelligent Computing Technology and Application Springer-verlag(ASV’09), Uslan, South Korea,PP : 645-651. DOI : 10.1007/978-3-642-04020-769.

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.