International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099

VOL 1 ISSUE 3 (2017) PAGES 120 - 127

Received: 24/03/2018. Published: 17/04/2018

Medicinal Plant Leaf Recognition and Show Medicinal Uses Using Convolutional Neural Network

[1]C. Sabarinathan, [2]Abhisekh Hota, [2]Ashish Raj, [2]Vivek Kumar Dubey, [1]V. Ethirajulu

[1]Assistant Professors, [2]UG Scholars
Department of Computer Science and Engineering
SRM Insittute of Science and Technology, Chennai, India
corresponding Author : abhisek@gmail.com

***Abstract*—**Plant identification by using leaf image is an essential issue which is seen in computer vision and biology. Methods used in the past for leaf image classification depend on features like hand-crafted techniques or texture information. In this paper, we proposed Convolutional Neural Network (CNN) for medicinal plant leaf recognition and medicinal uses of the plant is shown. We collected leaf images of 50 medicinal plants from google images. First, the leaf input image is down sampled into low resolution image and converted to a grayscale image. Then, edge detection is performed on the input image. Then, the input image is fed into the CNN architecture to learn features which can be used to distinguish different plants. In the last layer of CNN, all the discriminative data is combined to acquire the final feature. Then, the medicinal plant specie is predicted and its medicinal uses are shown. Our outcomes on three large public leaf datasets exhibit that our CNN shows better accuracy and consistency than the current methods used for leaf classification.

***Index terms*—** Medicinal plant, plant leaf classification, convolutional neural network, leaf shape, medicinal use

1. Introduction

Plant species can be recognized through plant parts like roots, stem, leaves, flowers and fruits. But in many cases these attributes don't show much differences to demonstrate mathematically and computationally. However, plant leaves are a rich source of data for species identification. Leaves remain connected to the plants for longer spans which is not true for fruits and flowers. Leaves have unique features such as shape and texture. In this paper, we have used leaf image for identifying medicinal plant species.

We view shape of an object as the geometrical informa- tion staying in the boundary. The size of the leaf changes with increase in leaf age. But the shape of a leaf remains generally unaltered over time and geographical location.

Most of the techniques used for leaf classification de- pend basically on shape of the leaf and handcrafted features. To give an example, the prominent plant species classifi- cation system Leafsnap [7] utilizes curvature-based shape features at different scales without utilizing any colour or texture qualities. Different techniques catch shape qualities by using "moment invariants" or "the multi-scale distance matrix". Different methodologies depend entirely on infor- mation from texture, overlooking information from shape. To be sure, shape information and texture information in leaf are very integral.

In this paper, we proposed a Convolutional Neural Net- work (CNN) to learn discriminative features to identify medicinal plant leaves, utilizing their shape qualities and show their medicinal uses. We analyze our approach against numerous prevalent handcrafted shape qualities. The out- comes on 3 huge public leaf datasets show that our CNN prompts higher precision and consistency than the best in class.

2. Related Work

Works on classification of plant species can be divided on the basis of features and classifiers used. To give an example, (i) most methods depend on information from shape [4], [6], [7], (ii) some methods depend on information from colour and texture [5] (overlooking shape information), and (iii) some methods depend on information from leaf shape and "vein geometry within the leaf" [9], [27] (ignoring information from colour and texture). [31] emphasis just on "vein structure extraction" without classification. Identification rates for [27] are very low for specific classes of leaves. The issue of vein detection [9], [27] depends on edge detection. It also does not perform good in situations where there is low contrast-to-noise ratio.

Most techniques utilize support vector machine (SVM) as the primary classifier. The latest leaf classification strate- gies utilize "neural networks" as the fundamental classi- fier. The technique in [14] utilizes a "probabilistic neural network" for leaf identification, showing 90% recognition accuracy on dataset like Flavia. Conversely, our CNN shows near perfect classification on dataset like Flavia by using information from leaf. Lee et al. [12] utilizes CNN for classification by using information from texture of the leaf ignoring information from shape of the leaf. [17], [18] utilize "hand crafted features" by taking information like leaf shape and leaf texture. Zhao et al. [13] utilizes CNN with "progressive sample learning". [12], [13] fail to classify huge leaf datasets.

3. Methodologies

3.1. Image Pre-processing

In image pre-processing, some of the leaf images are rotated manually. It is done to align the "leaf apex direction". Some images are also brightened. Then the input leaf image is cropped to 256×256 pixels. Cropping the image to low resolution helps in noise reduction. Then the leaf image is converted to grayscale image. Then in the feature extraction step edge detection is performed to get the shape information of the leaf.

3.2. Feature Extraction

After pre-processing, the shape of the plant leaf is de- tected. Shape is the unique feature for every leaf. Using shape as feature extraction, leaf classification can be performed. In feature extraction, edge detection on the leaf image is done by using OpenCV's edge detection tool. OpenCV's edge detection tool uses Canny algorithm. Canny algorithm reduces the amount of data to be processed.

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 1 ISSUE 3 (2017) PAGES 120 - 127
Received: 24/03/2018. Published: 17/04/2018

3.3. Classification

3.3.1. Convolutional Neural Network (CNN).

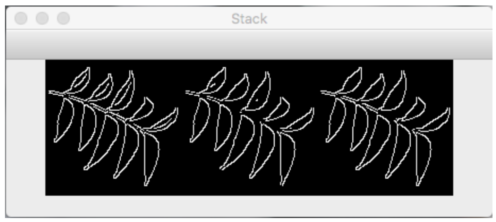

Figure 1: Leaf edge detection

CNN Architecture. Our CNN architecture learns shape information of the leaf to distinguish different plant species. The input leaf size is 256×256 pixels. The colour channel used for input to the CNN is grayscale. CNN consists of mainly three layers: Convolutional Layer, Pooling Layer and Fully-Connected Layer. By using all these layers, we get the final class scores which helps to distinguish different leaves. Feature extraction is done by convolutional and max- pooling layers. Batch Normalization (BN) layer follows each convolutional layer. Activation Function or Rectified Linear Unit (ReLU) is applied to all layers except the final layer. All the features learned are combined to form shape representation of the leaf. This shape representation is used as input for classification.

In CNN, when multiple input is passed it gives one output. The input consists of images and weights. The output is softmax. When the input leaf image is passed to the CNN it generates multiple input window. Then it is passed to the convolution layer. First Convolution layer consists of input, convolution and output. The input consists of 1 channel. The convolution consists of 2 filters. Activation function or Rectified Linear Unit (ReLU) is applied in this layer. It gives 2 tensors as output. Tensors are generated using sliding window. In the second convolution layer, 2 tensors from previous convolution layer is passed as input to this layer. In this layer there are 3 filters. This layer generates 3 tensors as output. In the pooling layer, 3 tensors from previous convolution layer are passed as input to this layer. Pooling layer consist of 1 filter. It generates 3 tensors as output. Pooling layer is used to reduce the height and width of input image. In the fully connected layer, the output from convolution layers and pooling layer is combined. 3 tensors from previous convolution layer is passed as input to this layer. This layer generates 4 neurons. In the softmax layer, 4 neurons from previous layer is passed as input. Softmax value for the input image is calculated in this layer. Softmax value is calculated using the input neurons and the weights. The highest value generated in softmax is used for classification.

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 1 ISSUE 3 (2017) PAGES 120 - 127
Received: 24/03/2018. Published: 17/04/2018

Objective Function.
Objective function is calculated by applying softmax. The formula used to calculate objective function is:

$$J(W) = \frac{-1}{N}\left[\sum_{n=1}^{N}\sum_{c=1}^{C}\ell\left(y_n == c\right) log \frac{e^{W_C^T X_i}}{\sum_{p=1}^{C} eW_p^T X_i}\right] \quad (1)$$

where (X1,y1),(X2,y2),...,(XN,yN) represents training set. N represents the number of training samples. C represents the number of training classes. ℓ(·) represents indicator function.

3.3.2. Transfer Learning

CNN requires to be trained with vast number of images to identify an image with great accuracy. Because of less availability of medicinal plant leaf images, we utilized transfer learning approach. In transfer learning, pre-trained network is used as a feature extractor. In transfer learning, all the pre-trained layers remain

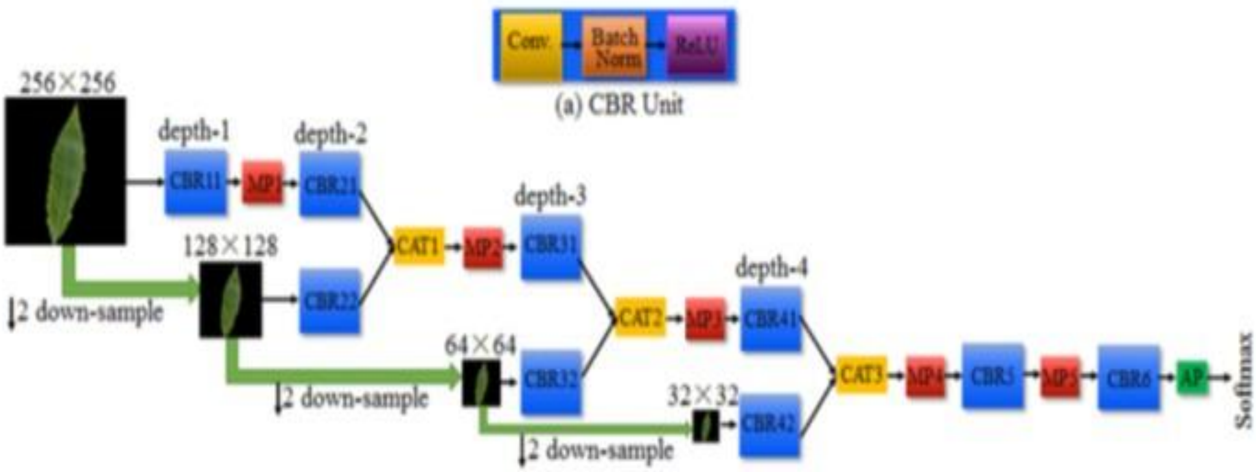

Figure 2: CNN Architecture

unchanged except the last fully connected layer. The most popular CNN architectures used for image classification are Inception-v3, OverFeat and Xception. We have considered Inception-v3 as our pre-trained network.

3.4. Medicinal Uses

Medicinal uses information of the medicinal plants dataset is stored in the MongoDB database. The database contains information about the plant's botanical name, fam- ily, about and medicinal uses. When the plant is classified from the classification step, the plant name is passed to the database using PyMongo connector. If the plant name exists in the MongoDB

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 1 ISSUE 3 (2017) PAGES 120 - 127
Received: 24/03/2018. Published: 17/04/2018

database, then all the information stored about the plant will be retrieved and shown in a html page using Flask.

4. Results

4.1. Dataset

Our medicinal plant leaf dataset consists of 50 medicinal plant species like aloe vera, amla, ashoka, ashwagandha, bael, brahmi, chandan, cinnamon, henna, lavender, marigold, neem, peppermint, tulsi, turmeric, etc. All the images are collected from google images. Each class consists of 30

Figure 3: Medicinal Plant Leaf Dataset

images. The resolution of the images used are not same for every image. Other datasets used are Leafsnap, Flavia and ImageClef.

4.2. Evaluation

By using our CNN, we evaluate leaf identification method on our medicinal plant leaf dataset. It showed 98% accuracy. We also evaluate leaf identification methods on 3 large publicly available datasets: (i) Flavia Dataset, (ii) Leafsnap Dataset, (iii) ImageClef Dataset. Flavia dataset consists of 1907 images and 32 classes. Leafsnap dataset consists of 7710 images and 150 classes. ImageClef dataset consists of 6630 images and 126 classes. We divide each dataset into two parts. One part consists of random subset of 70% leaves from each leaf class for training. The other part consists of random subset of 30% leaves from each leaf class for testing.

Our proposed CNN showed highest accuracy compared to other leaf classification

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 1 ISSUE 3 (2017) PAGES 120 - 127
Received: 24/03/2018. Published: 17/04/2018

methods on all datasets.

Method	Dataset	Top-1 Accuracy (%)	Top-3 Accuracy (%)
Uni-Path CNN	Flavia	99.28	99.97
	Leafsnap	95.61	99.67
	ImageClef	96.42	99.35
Texture-Patch CNN	Flavia	88.19	97.80
	Leafsnap	78.72	92.43
	ImageClef	70.35	87.83
Marginalized SC + SVM	Flavia	93.22	98.76
	Leafsnap	85.37	91.94
	ImageClef	81.63	88.72
Curvature SC + SVM	Flavia	82.84	94.62
	Leafsnap	71.13	78.43
	ImageClef	68.59	81.17
MDM + SVM	Flavia	82.55	95.12
	Leafsnap	75.76	85.96
	ImageClef	53.22	61.11

TABLE 1: Leaf Image Classification Performance

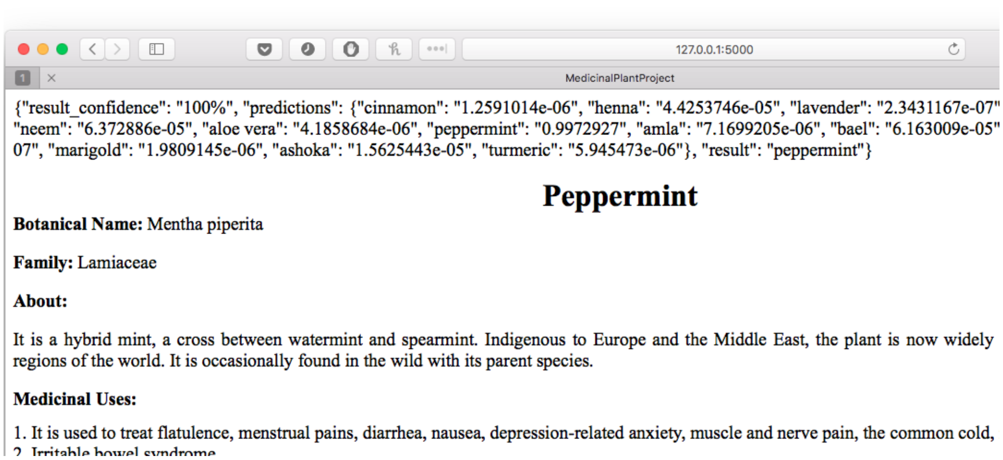

Figure 4: Screenshot of Medicinal uses of identified leaf

5. Conclusion

In this paper, we proposed Convolutional Neural Net- work (CNN) for medicinal plant leaf image classification and show its medicinal uses using MongoDB database. Shape information of leaf is used for leaf identification. Our CNN beats the best in class methods including SVM classifier and hand-crafted features on our medicinal plants dataset. The approach of this solution is non-specific. So it could be trained to classify any sort of leaf and not confined to medicinal plant leaves. This system could be utilized in many applicable areas where leaf classification is required.

VOL 1 ISSUE 3 (2017) PAGES 120 - 127

Received: 24/03/2018. Published: 17/04/2018

6. Future Work

There are thousands of medicinal plant species all over the world. Each medicinal plant has its own unique qualities which could be used to cure different diseases. So the future work would be to make a larger database with medicinal plant leaf images gathered from various sources all over the world.

References

[1] J. C. Liu, C. Y. Chiang and S. Chen, ”Image-Based Plant Recogni- tion by Fusion of Multimodal Information,” 2016 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), Fukuoka, 2016, pp. 5-11.

[2] M. P. Shah, S. Singha and S. P. Awate, ”Leaf classification using marginalized shape context and shape+texture dual-path deep convolu- tional neural network,” 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 860-864.

[3] J. Hu, Z. Chen, M. Yang, R. Zhang and Y. Cui, ”A Multi-Scale Fusion Convolutional Neural Network for Plant Leaf Recognition,” in IEEE Signal Processing Letters.

[4] P Belhumeur, D Chen, S Feiner, D Jacobs, W Kress, H Ling, I Lopez, R Ramamoorthi, S Sheorey, and S White, “Searching the world’s herbaria: A system for visual identification of plant species,” in Proc. Euro. Conf. Comp. Vis., 2008, pp. 116–129.

[5] J Cope, P Remagnino, S Barman, and P Wilkin, “Plant texture classifi- cation using Gabor co-occurrences,” in Proc. Adv. Vis. Comput., 2010, pp. 669–77.

[6] R. Hu, W. Jia, H. Ling and D. Huang, ”Multiscale Distance Matrix for Fast Plant Leaf Recognition,” in IEEE Transactions on Image Processing, vol. 21, no. 11, pp. 4667-4672, Nov. 2012.

[7] N Kumar, P Belhumeur, A Biswas, D Jacobs, W Kress, I Lopez, and J Soares, “Leafsnap: A computer vision system for automatic plant species identification,” in Proc. Euro. Conf. Comp. Vis., 2012, pp. 502–516.

[8] S Mouine, I Yahiaoui, and A Verroust-Blondet, “Advanced shape context for plant species identification using leaf image retrieval,” in Proc. Int. Conf. Multimedia Retrieval, 2012, p. 49.

[9] Y Nam, E Hwang, and D Kim, “A similarity-based leaf image re- trieval scheme: Joining shape and venation features,” Comp. Vis. Imag. Understanding, vol. 110, no. 2, pp. 245–59, 2008.

[10] Z Wang, Z Chi, and D Feng, “Shape based leaf image retrieval,” IEE Proc. Vis. Imag. Signal Proc., vol. 150, no. 1, pp. 34–43, 2003.

[11] S Mouine, I Yahiaoui, and A Verroust- Blondet, “A shape-based ap- proach for leaf classification using multi-scale triangular representation,” in Proc. Int. Conf. Multimedia Retrieval, 2013, pp. 127–34.

[12] S Lee, C Chan, P Wilkin, and P Remagnino, “Deep-plant: Plant identification with convolutional neural networks,” in IEEE Int. Conf. Imag. Proc., 2015, pp. 452–6.

[13] Z Zhao, B Xie, Y Cheung, and X Wu, "Plant leaf identification via a growing convolution neural network with progressive sample learning," in Proc. Asian Conf. Comp. Vis., 2014, pp. 348–61.
[14] S Wu, F Bao, E Xu, Y Wang, Y Chang, and Q Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in Proc. Int. Symp. Sig. Proc. Info. Tech., 2007, pp. 11–16.
[15] "Flavia dataset," flavia.sourceforge.net, Accessed: 2017-01-01.
[16] A Krizhevsky, I Sutskever, and G Hinton, "Imagenet classification with deep convolutional neural networks," in Adv. Neural Info. Proc. Systems, 2012, pp. 1–8.
[17] J Chaki, R Parekh, and S Bhattacharya, "Plant leaf recognition using texture and shape features with neural classifiers," Pattern Recog. Let., vol. 58, 61-68.
[18] T Kumar, V Reddy, and P Bora, "Leaf identification using shape and texture features," in Int. Conf. Comp. Vis. Imag. Proc., 2016, pp. 531–541.
[19] R Caruana, S Lawrence, and C Giles, "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping," in Proc. Neural Info. Proc. Systems, 2000, pp. 402–408.
[20] S Ioffe and C Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in Proc. Int. Conf. Mach. Learn., 2015, pp. 448–56.
[21] V Nair and G Hinton, "Rectified linear units improve restricted Boltzmann machines," in Proc. Int. Conf. Mach. Learn., 2010, pp. 807–814.
[22] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in Workshop on Performance Evaluation of Tracking and surveillance, 2007.
[23] J. Zhu, H. Zeng, S. Liao, Z. Lei, C. Cai, and L. Zheng, "Deep hybrid similarity learning for person re-identification," arXiv preprint arXiv:1702.04858, 2017.
[24] "Leafsnap dataset," leafsnap.com/dataset, Accessed: 2017-01-01.
[25] "ImageClef 2012 dataset," www.imageclef.org/2012/plant, Accessed: 2017-01-01.
[26] François Chollet, "Xception: Deep Learning with Depthwise Separa- ble Convolutions", arXiv:1610.02357 [cs.CV], 2016.
[27] J Park, E Hwang, and Y Nam, "Utilizing venation features for efficient leaf image retrieval," J. Systems Software, vol. 81, no. 1, pp. 71–82, 2008.
[28] S Belongie, J Malik, and J Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 4, pp. 509–522, 2002.
[29] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper with Convo- lutions", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
[30] I. Gogul and V. S. Kumar, "Flower species recognition system us- ing convolution neural networks and transfer learning," 2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN), Chennai, 2017, pp. 1-6.
[31] J Clarke, S Barman, P Remagnino, K Bailey, D Kirkup, S Mayo, and P Wilkin, "Venation pattern analysis of leaf images," in Int. Symp. Vis. Comput., 2006, pp. 427–36.