Plant Species Identification using Discriminant Bag of Words (DBoW)

Fiza Murtaza[1][a], Umber Saba [1][b], Muhammad Haroon Yousaf [1,2][c] and Serestina Viriri[3][d]

[1]*Computer Engineering Department, University of Engineering and Technology Taxila, Pakistan*

[2]*Swarm Robotics Lab-NCRA, University of Engineering and Technology Taxila, Pakistan*

[3]*University of KwaZulu-Natal, Durban, South Africa*

Keywords: Plant Species Identification, SURF, Kernel Descriptors, ImageClef, Flavia, Bag of Words.

Abstract: Plant species identification is necessary for protecting biodiversity which is declining rapidly throughout the world. This research work focuses on plant species identification in simple and complex background using Computer Vision techniques. Intra-class variability and inter-class similarity are the key challenges in a large plant species dataset. In this paper, multiple organs of plants such as leaf, flower, stem, fruit, etc. are classified using hand-crafted features for identification of plant species. We propose a novel encoding scheme named as Discriminant Bag of Words (DBoW) to identify multiple organs of plants. The proposed DBoW extracts the class specific codewords, and assigns the weights to codewords in order to signify discriminant power of the codewords. We evaluated our proposed method on two publicly available datasets: Flavia and ImageClef. The experimental results achieved classification accuracy rates of 98% and 94% on FLAVIA and ImageClef datasets, respectively.

1 INTRODUCTION

The identification of plants using manual method is time consuming, complex and frustrating for a layperson because of specialized botanical terms. The difficulties encountered in manual classification have prompted the necessity of computer vision to automate the process (Kala et al., 2016a). Nowadays, researchers have developed approaches to automate the process of plant species identification. The automatic classification of plant species is important for scientists and botanist who desire to obtain the series of features that describe plant structure. The availability of relevant technologies, remote access to databases, new techniques in image processing and pattern recognition let the idea of automated species identification become a reality. Automatic plant species identification using computer vision has applications in weeds identification, species discovery, plant taxonomy, natural reserve park management, etc.

The datasets can be categorized based on plant organs which are used for plant identification. The most common organ used for identification is leaf because it is the most abundant organ in nature and easiest to obtain in the field study throughout the year. Although leaf, flower, stem, fruit and entire plant can also be used for identification of plant species, they are not addressed adequately in literature due to their complexity and unavailability throughout the year. Some examples of different organs of plant are shown in Figure 1.

A combination of hand-crafted and deep learning features is proposed in (Nguyen et al., 2017) for the identification of plant species using leaf and flower organs. Convolution Neural Networks (CNN) and Kernel descriptors (KDES) are used in parallel for comparison purpose in (Nguyen et al., 2017). Experimental results show that hand-crafted features outperform deep features in case of simple images with plain background. However, deep learning performs better in case of natural and complex environment, but data augmentation is required to convert small dataset into a large dataset. Furthermore, the techniques like Borda Count (BC) (Goëau et al., 2012) and Inverse Rank Position (IRP) (Goëau et al., 2012) are used for combining results of different plant organs but they did not perform well for all types of organs.

A deep learning-based model LeafNet is proposed in (Barré et al., 2017) for plant identification using

[a] https://orcid.org/0000-0001-8259-7389
[b] https://orcid.org/0000-0000-0000-0000
[c] https://orcid.org/0000-0001-8255-1145
[d] https://orcid.org/0000-0002-2850-8645

Murtaza, F., Saba, U., Yousaf, M. and Viriri, S.
Plant Species Identification using Discriminant Bag of Words (DBoW).
DOI: 10.5220/0009161004990505
In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020) - Volume 5: VISAPP*, pages 499-505
ISBN: 978-989-758-402-2; ISSN: 2184-4321

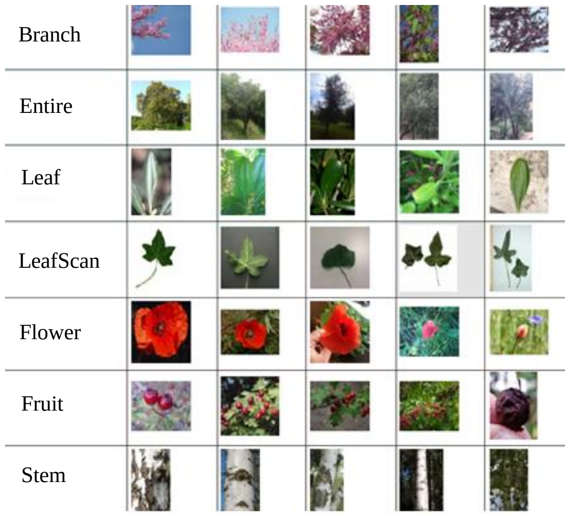

Figure 1: Different organs of the plant species.

Flavia dataset (Wu et al., 2007). Flavia is a small dataset with 32 plant species so data augmentation has been performed to generate 44242 images from 1526 images before training the network and achieved 97.9% classification accuracy. The use of deep learning approach on smaller datasets is dependent on data augmentation, which is time consuming and an additional overload. Kala et. al. (Kala et al., 2016b) proposed an approach for leaf classification using convexity moment of polygon. In (Kala et al., 2016b) series of experiments are performed on Flavia dataset and 92% accuracy is achieved using Radial Basis Function (RBF) classifier. In (Kala et al., 2016b) geometric features i.e. rectangularity, aspect ratio and circularity are combined with convexity moment of polygon for identification of plant species. The technique worked well for the small and plain background images like in Flavia but not suitable for complex datasets like ImageClef (Goëau et al., 2014a).

Ghazi et. al. (Mehdipour Ghazi et al., 2015) took the challenge to use complex and large dataset with natural background i.e. ImageClef for the validation of their approach and used deep learning based Principal Component Analysis Net (PCANet). PCANet is a simple but efficient deep learning network for image classification, that has been used with the combination of the scale-invariant feature transform (SIFT) based VLAD features. Classification accuracy of 21.2% has been obtained for LeafScan organ using PCANet. PlantNet (Goëau et al., 2014b) used hand-crafted features Random Maximum Margin Hashing (RMMH) for different types of organs of plant identification and classification accuracy of 51% has been achieved using adaptive KNN rule.

To conclude the literature, two streams can be observed for classification of plant species i.e. based on hand-crafted features or deep learning. Although, deep learning-based approaches obtained promising results but at the cost of dedicated computational resources and data augmentation for training. However, hand-crafted features give better results than deep learning-based methods when small and simple datasets are involved. Deep learning-based approaches also demand more time for training as compared to hand-crafted features.

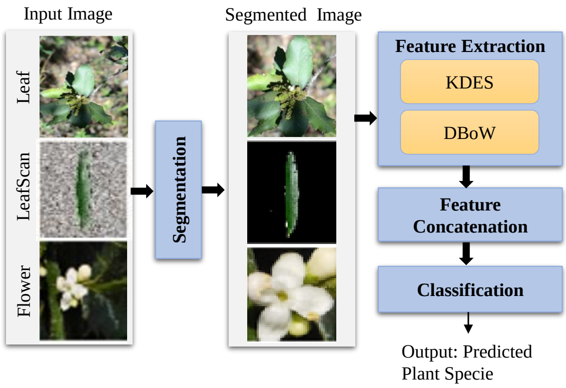

Figure 2: Flow diagram of proposed methodology.

A lot of research work has been done in this emerging field such as describe in (Kala et al., 2016a; Nguyen et al., 2017; Barré et al., 2017; Kala et al., 2016b) but there still exist problems in this domain which are still unaddressed by the researchers. We can observe lot of work on computer vision-based plant species identification in simpler scenarios but identification of species with cluttered background, non-uniform illumination and viewpoint variation is still a challenging task and there is ample room for researchers to propose new techniques in this domain.

In this paper, we have proposed a hand-crafted feature-based approach for plant species identification in simple as well as complex scenario. Our contributions are as follows: firstly, we extract multiple feature descriptors to identify plant species using multiple organs e.g. leaf and flowers of plants for better accuracy. Secondly, we proposed an encoding scheme named as Discriminant Bag of Words (DBoW) that can classify plant species with high accuracy on complex and large datasets. The detailed methodology and results are explained in Section 2 and Section 3, respectively.

2 PROPOSED METHODOLOGY

The proposed methodology for the identification of plant species is depicted in Figure 2. The steps re-

(a) (b) (c)

Figure 3: Segmentation results for (a) LeafScan, (b) Leaf and (c) Flower organ.

quired in the proposed methodology are preprocessing/segmentation, feature description and classification. Due to the discriminative nature of organs, different segmentation modules are introduced for different organ type. We can observe the simple background of LeafScan as compared to complex background of leaf in Figure 5. Similarly, flower has different color and shape feature as compared to leaf. Two different feature descriptors i.e. KDES and speeded up robust features (SURF) are merged in feature extraction module and then Support Vector Machine (SVM) is used for classification.

2.1 Preprocessing/Segmentation

Preprocessing enhances plant organs segmentation before their feature extraction. CIELAB color space is used for the segmentation of region of interest of each organ. The color space L represents the brightness level, in the range $[0,100]$ with the darkest black at $L = 0$ and the brightest white at $L = 100$. The color space A represents the green–red component, with green in the negative direction and red in the positive direction in the range: $[-128,127]$. Whereas the color space B represents the blue–yellow component, with blue in the negative direction and yellow in the positive direction in the range: [-128, 127]. CIELAB color space seems to be the most appropriate in segmentation of all organ types. For example, if we want green and yellow color in case of leaf organ, we can accordingly set the threshold for color in A and B space respectively. Moreover, we can also segment out dark and light part of image using L space of CIELAB color space. For three organs (LeafScan, flower and leaf), we set the threshold greater than 60 for L color space to ignore the dark objects.

For Leaf and LeafScan images, green color region of the leaf is extracted by setting threshold less than 50 for channel A to ignore the dark red objects while retrieving green objects. As the leaves may have yellow color, therefore, for channel B, we set threshold greater than 0 to ignore the blue component i.e. the sky region in the images while retrieving the yellow objects. A sample image of segmentation for LeafScan and Leaf organ is shown in Figure 3. Note that in Figure 3(a) the LeafScan image has plain background therefore the boundary of the leaf can be easily obtained by multiplying the binary image with the original RGB image. For Leaf organ in Figure 3(b) the background is not plain rather it has some of the background leaves and stems etc. therefore we cropped the region where the leaf region is detected by applying above defined threshold on the CIELAB color space.

As flower organ varies a lot in terms of color and are surrounded by leaves, most segmentation algorithms e.g. active contour, expectation maximization, watershed etc. are not able to segment the flowers accurately. However, flowers can be segmented using color thresholds on CIELAB color space. For A channel we set the threshold greater than -1 to ignore the green objects i.e. leaves are ignored. Similarly, for channel B, we set threshold greater than -50 to ignore the sky region the flower images. Results for the segmentation for different flower organs are shown in Figure 3(c).

2.2 Feature Extraction

After preprocessing, the next step is to extract discriminant features from each organ image. We proposed a novel feature encoding method named Discriminant Bag of Words (DBoW) features which includes hand-crafted feature descriptor i.e. SURF descriptor followed by discriminant encoding method as shown in Figure 4. KDES and DBoW features are extracted from each image in parallel and they are concatenated before passing to the SVM classifier for final classification. KDES and DBoW based feature extraction method is explained in the following subsections.

2.2.1 Kernel Descriptors (KDES) Features

Kernel Descriptors (Bo et al., 2010) have been proved to be robust for visual object and scene recognition. Different types of kernel are used in KDES i.e. Gradient kernel (G-KDES), RGB kernel (RGB-KDES) and Local Binary Pattern kernel (LBP-KDES). We applied different types of KDES for different organs

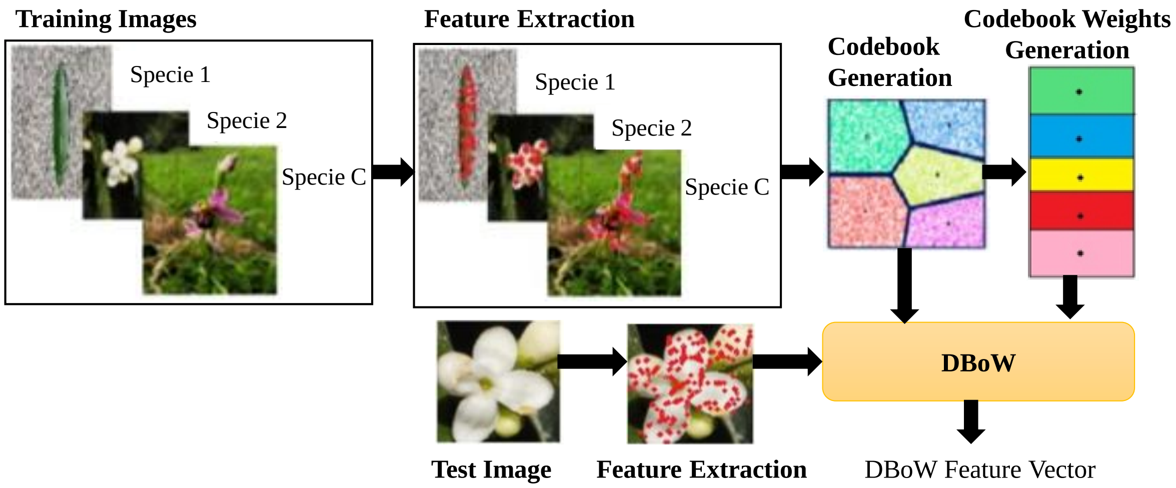

Figure 4: Feature extraction using DBoW approach.

type and conclude that RGB-KDES works better for flower organ and G-KDES works better for leaf organ.

In KDES, patch level and image level features are extracted. With the help of uniform grid, a reduced number of pixels are selected and patches of predefined area around each pixel is then extracted. We build a metric to find the similarity between two image patches. On each patch level feature, k-means clustering is applied and then visual word dictionary is built. The last step is to compute image level feature on a learned dictionary using spatial pyramid matching (Bo et al., 2010).

2.2.2 Discriminant Bag of Words (DBoW) Features

The proposed DBoW approach contains four steps. Firstly, the SURF features are extracted from each organ image. Secondly, the codebook is generated per each plant species using k-means clustering. Thirdly, the discriminant codewords are selected from each codebook by learning the weights of each codeword as proposed in (Murtaza et al., 2018). The idea behind discriminant codewords is to generate codewords that vote for the features of their own class. The major challenge faced in complex datasets is inter-class similarity and intra-class variability. Using the traditional codewords, features not only vote for the codewords of their own class but they also vote for the codewords of other similar classes. Therefore, we proposed to solve this issue by finding the codewords from each class which have higher similarity with their own classes. Finally, the discriminative codewords are used to encode the features from the training and testing images into a fixed-length feature vector per image.

We integrate discriminate codewords with the traditional Bag of Words (BoW) methods therefore we named our proposed method as DBoW. Main steps required to extract DBoW features are given in Algorithm 1.

When we multiply weight with codewords (Step 8 of Algorithm 1), we are actually enhancing the votes of features that is actually part of that class and suppress the votes of those features that is not the part of that class and hence increase in accuracy in resultant.Both feature vectors, DBoW and KDES are fused and passed to SVM classifier for training and testing purpose resulting in plant species identification.

3 EXPERIMENTAL RESULTS

3.1 Datasets

We have used two publicly available datasets i.e. Flavia (Wu et al., 2007) and ImageClef (Goëau et al., 2014a) to validate our proposed approach on simple as well as complex scenario. ImageClef dataset has more than sixty thousand images belonging to 500 plants species existing in West Europe. In this paper, leaf including Leaf and LeafScan and Flower organs as images are used. Table 1 gives details of the training and testing set.

We also conducted experiments on Flavia dataset (Wu et al., 2007) which only contains leaf organ composed of 1907 leaf images with 32 species. Each specie has about 50 to 60 images. As a standard practice, this dataset has been divided into 70% training and 30% test set. The main challenges of ImageClef are illustrated in Figure 5.

Algorithm 1: Steps required to extract DBoW features.

Input: Segmented images of plant organs.
Output: DBoW features for each plant organ.

1: Detect and extract SURF features for each organ image.
2: Extract K clusters from each class by apply k-means clustering on the training images of each class separately to find the class-specific codebooks. This process will result in $K \times C$ clusters where C represents total number of plant species present in the dataset.
3: Initiate two vectors A and $A^{'}$ of length $K \times C$ with zeros, where A represents within class assignment and $A^{'}$ represents out of class assignment.
4: Compare all feature vectors in the training set one by with the $K \times C$ codewords and find the best matching cluster i using Euclidian distance.
5: If the feature vector and its nearest codeword i belong to the same class, then assign vote to $A[i]$ otherwise assign vote to $A^{'}[i]$. This will result in the within class and out of class assignments for each codeword.
6: Then for each codeword find the discriminate weight using following equation:

$$w[i] = \frac{A[j]}{A[i] + q[i]'} \qquad \forall\, j \in [1, 2, 3 \cdots, K \times C] \tag{1}$$

If $w[i]$ has values closer to 1 then it means more features vectors are assigned to codeword of its own class
7: Extract BOW representation using these $K \times C$ clusters.
8: Then the DBoW features are calculated by dot multiplication of BOW representation with w resulting in a $K \times C$ dimensional feature vector.

Table 1: Train and test set provided by ImageClef dataset.

Plant Organ	Leaf	LeafScan	Flower
Training Set	7754	11335	13164
Test Set	2058	696	4559
Total	9812	12031	177223

Flavia is less complex dataset as compared to ImageClef and contain plain background of images. The leaf images were acquired by scanners or digital cameras on plain background therefore we did not perform segmentation on this dataset. The isolated leaf images contain blades only, without petioles as shown in Figure 6. For both datasets, accuracy is used as a standard evaluation metric to evaluate our approach.

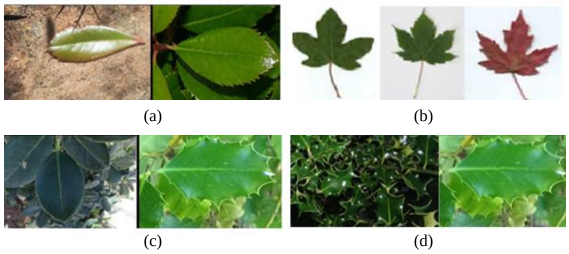

Figure 5: Some Challenging images of ImageClef are: (a) same class but images with different resolution (b) images of different plant species with similar shape (c) diversity of leaf Image in same class (d) images of Complex, uncontrolled and cluttered environment.

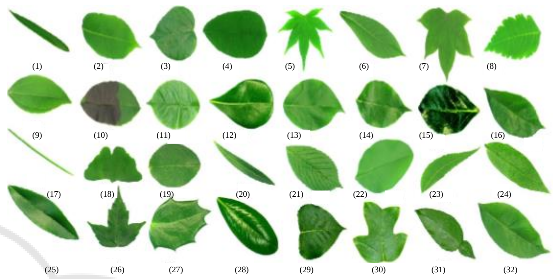

Figure 6: Sample images from 32 species of Flavia Dataset.

3.2 Implementation Details

For ImageClef dataset, it has been observed that dataset is imbalanced. There are some classes in test dataset having no image present in train data and vice versa so we discarded that classes at the time of evaluation. Furthermore, it is also observed that classes are imbalanced which affect the accuracy of the proposed method therefore we balanced dataset by finding the minimum number of images m from each class and randomly selecting m number of images from each class.

For KDES, in patch level feature extraction we used 200 eigenvectors. For building dictionary, we selected 1000 numbers of visual words. In image level feature extraction, spatial pyramid matching is applied on 3 layers. Hence, we have a feature vector of length 21000 for each image.

For DBoW, we used SURF as a feature descriptor and take only 256 valid points hence we have 256×64 length of features against each image. After that we performed class specific k-means clustering with $K = 128$ resulting in $K \times C$ dimensional feature vector for each image.

3.3 Results and Discussion

The classification results of proposed approach on ImageClef dataset using different combination of the

Table 2: Evaluation of the proposed method in terms of accuracy (%) on ImageClef dataset using KDES features and their combination with DBoW features.

Variants of the proposed method	Segmentation	LeafScan	Leaf	Flower
G-KDES	✗	85.45	36.91	34.50
G-KDES	✓	89.42	38.04	38.90
LBP-KDES	✓	85.50	40.11	35.02
G-KDES+LBP-KDES	✓	92.34	40.51	35.60
G-KDES+LBP-KDES+DBoW (Our)	✓	**94.03**	**42.11**	**54.75**

Table 3: Comparison results in terms of accuracy (%) on ImageClef dataset with the state-of-the-art methods.

Method	LeafScan	Flower	Leaf
Mica (Nguyen et al., 2017)	78.3	10.9	24.3
Mica Run 3 (Goëau et al., 2012)	43.7	20.7	34.2
Mica Run 1 (Goëau et al., 2012)	43.7	19.5	34.2
Sabanci Run 1 (Mehdipour Ghazi et al., 2015)	21.6	18.9	11.1
PlantNet Run 4 (Goëau et al., 2014b)	54.1	36.6	16.5
FINKI Run 3 (Wu et al., 2007)	44.9	25.5	16.0
IBM AU Run 2 (Goëau et al., 2014a)	61.2	55.5	30.0
Sabanci Run 3 (Chen et al., 2014)	90.5	34.0	33.8
Sabanci Run 2 (Mehdipour Ghazi et al., 2015)	90.5	34.0	33.8
G-KDES+LBP-KDES+DBoW (Our)	**94.0**	**54.7**	**42.1**

KDES and DBoW based features are shown in Table 2. There are two different parts of results as shown in the Table 2. First part shows the results with and without segmentation of the different organ images. Results show that segmentation step resulted in higher accuracy on all of three organs as compared to without segmentation. Using G-KDES with segmentation, there is an increase of about 4% for Flower and LeafScan and 1% for Leaf organ.

In the second part of the Table 2, we compute the results using the combination of the G-KDES and LBP-KDES on the segmented images and it can be seen that accuracy further improved using the combination of these features. Finally we combine G-KDES and LBP-KDES with the BoDW and the results show that the proposed DBoW features complement the state-of-art KDES features. The accuracy achieved using our approach clearly shows the efficiency of the proposed discriminant features that boast the accuracy when combined with traditional KDES because of class-wise clustering and generating discriminate codewords.

Table 3 presents the comparison results of the proposed method with the stat-of-the-art methods on ImageClef dataset. Our proposed method outperforms the state-of-the-art methods because it enhances discriminant codewords and suppress non-discriminant codewords through discriminate encoding scheme. The results show that this approach can help identification of multiple species in large and complex dataset.

The classification accuracy of proposed approach on Flavia dataset is shown in Table 4. The results show that the proposed method achieved a higher accuracy rate as compared to other methods. This shows that our proposed method with discriminant codewords resulted in discriminative representation of images using handcrafted features i.e. SURF features.

4 CONCLUSIONS

In this paper, discriminate codeword generation technique is proposed to identify plant species captured in challenging background. The experiments have been tested on two different types of datasets i.e. Flavia and ImageClef. The Flavia dataset has images with simple background of 32 species while ImageClef is very complex multi-organ images dataset captured in natural environment. The obtained results are promising on both datasets. On Flavia dataset, 98% accuracy rate has been achieved using 32 species while 94% accuracy rate of LeafScan organ of ImageClef dataset has been achieved using our proposed method.

For future work, it is envisioned that fusion techniques can be used to combine different feature extraction techniques for different plant organs. Although flower and leaf are the most commonly used

Table 4: Comparison result of the proposed method on Flavia dataset with the state-of-the-art other methods.

Method	Accuracy (%)
Thi-Lan Le,2014 (Nguyen et al., 2017)	97.5
Wu et al., 2007 (Wu et al., 2007)	90.3
Kadir, 2014 (Kadir, 2014)	97.2
Pierre Barr, 2017 (Barré et al., 2017)	97.9
Serestina Viriri, 2016 (Kala et al., 2016b)	95.0
G-KDES+LBP-KDES+DBoW (Our)	**98.0**

organs for plant species identification as they remain available throughout the year, other organs of plant such as stem and fruits can also be considered for different plant species identification. Furthermore, this work can also be extended to achieve favorable results by utilizing deep convolutional neural networks in order to evaluate their ability to identify plant species at a large-scale.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support by the Centre for Computer Vision Research (C^2VR) and Swarm Robotics Lab-NCRA, University of Engineering and Technology (UET) Taxila, Pakistan.

REFERENCES

Barré, P., Stöver, B. C., Müller, K. F., and Steinhage, V. (2017). Leafnet: A computer vision system for automatic plant species identification. *Ecological Informatics*, 40:50–56.

Bo, L., Ren, X., and Fox, D. (2010). Kernel descriptors for visual recognition. In *Advances in neural information processing systems*, pages 244–252.

Chen, Q., Abedini, M., Garnavi, R., and Liang, X. (2014). Ibm research australia at lifeclef2014: Plant identification task. In *CLEF (Working Notes)*, pages 693–704.

Goëau, H., Bonnet, P., Barbe, J., Bakic, V., Joly, A., Molino, J.-F., Barthelemy, D., and Boujemaa, N. (2012). Multi-organ plant identification. In *Proceedings of the 1st ACM international workshop on Multimedia analysis for ecological data*, pages 41–44. ACM.

Goëau, H., Joly, A., Bonnet, P., Selmi, S., Molino, J.-F., Barthélémy, D., and Boujemaa, N. (2014a). Lifeclef plant identification task 2014. In *CLEF2014 Working Notes. Working Notes for CLEF 2014 Conference, Sheffield, UK, September 15-18, 2014*, pages 598–615. CEUR-WS.

Goëau, H., Joly, A., Yahiaoui, I., Bakić, V., Verroust-Blondet, A., Bonnet, P., Barthélémy, D., Boujemaa, N., and Molino, J.-F. (2014b). Plantnet participation at lifeclef2014 plant identification task.

Kadir, A. (2014). A model of plant identification system using glcm, lacunarity and shen features. *arXiv preprint arXiv:1410.0969*.

Kala, J. R., Viriri, S., and Moodley, D. (2016a). Sinuosity coefficients for leaf shape characterisation. In *Advances in Nature and Biologically Inspired Computing*, pages 141–150. Springer.

Kala, J. R., Viriri, S., Moodley, D., and Tapamo, J. R. (2016b). Leaf classification using convexity measure of polygons. In *International Conference on Image and Signal Processing*, pages 51–60. Springer.

Mehdipour Ghazi, M., Yanıkoğlu, B., Aptoula, E., Muslu, Ö., and Özdemir, M. C. (2015). Sabanci-okan system in lifeclef 2015 plant identification competition. CLEF (Conference and Labs of the Evaluation Forum).

Murtaza, F., HaroonYousaf, M., and Velastin, S. A. (2018). Da-vlad: Discriminative action vector of locally aggregated descriptors for action recognition. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3993–3997. IEEE.

Nguyen, T. T.-N., Le, T.-L., Vu, H., Nguyen, H.-H., and Hoang, V.-S. (2017). A combination of deep learning and hand-designed feature for plant identification based on leaf and flower images. In *Asian Conference on Intelligent Information and Database Systems*, pages 223–233. Springer.

Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y.-X., Chang, Y.-F., and Xiang, Q.-L. (2007). A leaf recognition algorithm for plant classification using probabilistic neural network. In *2007 IEEE international symposium on signal processing and information technology*, pages 11–16. IEEE.