

Wlab of University of Szeged at LifeCLEF 2014 Plant Identification Task

Dénes Paczolay¹, András Bánhalmi¹, László G. Nyúl¹, Vilmos Bilicki¹, Árpád Sárosi²

University of Szeged, Szeged, Hungary

¹{pdenes, banhalmi, nyul, bilickiv}@inf.u-szeged.hu

²sarosi.arpad@gmail.com

Abstract. Our aim is to implement a plant identification application that can run on smartphones, and this shared task includes it. After the plant identification task of 2013 we concluded that the most frequent trees (e. g. in Hungary) can be identified well by a leaf, when there is a white paper background behind it at the time of photographing. This is why we want to classify more accurately this kind of pictures. Our other important goal is to develop a system which can be trained online while smartphone users take more and more photos. In the previous year there was a separate shared task for the 'Sheet as Background' photos, however this year the task is very complex, and supposes solutions for hard image processing problems, not just feature extraction and classification solutions.

Keywords: Plant identification, sheet as background, leaf classification, contour and metric feature combination, contour histograms, gradient histograms in color channels

1 Introduction

The ImageCLEF Plant identification task [3] is becoming increasingly more complex year after year, while more and more photos are being taken of more and more plants. This year 500 different types of plant species need to be identified. Although there are currently 47815 photos, this number is still not sufficiently high, because for many specific classes there are only a few examples available, which means that the usual classification methods may not be trained well, and other solutions are required. Because the number of examples for the scan-like leaf images is the highest (11335), and because the preprocessing of these kinds of pictures seems to be the least complex, we focused on this task, and all the others were treated to get a better solution than a random or a most-is-best classifier. In addition, we will concentrate on scan-like images because our aim is to improve an initial training by online learning methods, which refines a basic model continuously, when the users of smartphones take new photos of leaves with a white background. Our idea is that when users take new photos, there

will be a voting mechanism using human votes, which decides if a new photo belongs to one specific species or not. There is also a confidence value assigned to all the users, and this value is also updated when the smartphone application is used continuously for a long time.

2 Experiments

The training dataset [2] was downloaded from the PlantCLEF webpage, and it contained 47815 images (1987 of "Branch", 6356 photographs of "Entire", 13164 of "Flower", 3753 "Fruit", 7754 of "Leaf", 3466 "Stem" and 11335 scans and scan-like pictures of leaf) with complete xml files associated to them.

Table 1. The number of species for which there is a photo in a given category, the number of species for which there is fewer than 10 examples, and the average example number. *The scan-like leaf category has the best (image) example/species ratio, but with a lot of visually redundant near duplicate images, and with fewer informative images from the same observation. When observations are considered, LeafScan and Flower categories have mostly the same observation example/species ratio.*

category	branch	entire	flower	fruit	leaf	stem	scan-like	leaf
# of species	356	490	483	374	470	399	212	
# of species with < 10 examples	293	191	64	236	237	290	83	
average examples/species	5.6	13.0	27.3	10.0	16.5	8.9	53.5	

Table 1 lists some statistics concerning the photos for each category. As can be seen, in the case of "Branch", "Fruit", "Leaf" and "Stem" the majority of species are strongly underrepresented; very little data is available. In the case of "Flower", however, there are many more photos, and for almost the all species (483), from the 500 species. The scan-like pictures have the best example/species ratio, however there are many species for which scan-like leaf data is not available. From this point of view, we should surely focus on the classification using mainly the scan-like leaf and flower pictures, or when using the other categories, then at the score aggregation we should consider the number of examples in the training set for a category and species.

2.1 SheetAsBackground (Scan-like Leaf, LeafScan) Category

This category contains the most examples for a species, and it seems to be the easiest to preprocess, namely to get a mask of a leaf. We tried to find the less complex methods for masking, feature extraction and classification, because our aim is a fast method that can be used on a smartphone without waiting long for the classification or ranking results.

Masking: Our masking method is rather simple. After taking the blue channel of the color picture, a 3x3 blur is applied on it, and Otsu's thresholding method is applied. The result is obtained from this after the salt-and-pepper noise has been removed.

The masking method will fail in some cases, for example when the background has a similar intensity on the blue channel than the leaf, but in most cases it works well.

Features: In order to classify the species by their leaves, an important task is to extract features suitable for this. Since the leaves' orientation on the photo is not fixed, the appropriate features have to be scale and rotation invariant. Instead of starting from scratch, we based our features on two sets taken from the literature.

First, we implemented the vein and metric features introduced in [5]. These basic features contain area information of veins computed by an unsharp masking method. Vein density is computed at four levels with four different blur radii. Some of the metric features in [5] supposes, that the two endings of the leaves have been set by hand. We modified these features, such that our method finds the longest projection of a leaf, and computes the perpendicular projection, as well. This way, we don't need endings, and don't need a fixed leaf orientation. Metric features include $area/perimeter^2$, $perimeter/diameter$, $diameter/perpendiculardiameter$, etc.

More complex features are based on the TAR and TSL descriptors defined in [4], but in another way we used them to extract contour features. The idea behind it is that a given number of points are set uniformly on the contour of a leaf. At each point a given number of features can be computed from the data of triangles joining the neighbouring points. When the features are the ratio of side lengths, then these scale invariant measures are called TSL, and when the features are oriented angles of the triangles, then it is called TOA. For each of the N points on the leaf contour, a number of M rectangles are set, and each rectangle is described by two angles, or two side length ratios. This method gives a number of $2 \cdot N \cdot M$ values. The question is the way of representation for classification, because the feature vectors computed from the previous data have to be rotation invariant. In [4], the Locality Sensitive Hashing technique is proposed for the purpose of matching two vectors, and measuring similarity. Here we draw a simpler feature extraction method, which is also rotation invariant, as in [4]. In our solution only one angle is computed for a triangle, this being the angle at the actual control point for which the neighbouring triangles were set. The other two descriptors for a triangle are the ratios of side lengths, but in our representation the maximal side length is always the denominator, then the range of these values is between 0 and 1. After we compute the angle and the two side ratios for each triangle, we simply make three histograms, and since the histogram cumulates the values independently of the order of control points, the feature will be invariant under rotation. As the size of each histogram is 10, we will have 30 additional features.

Classification Method: We modeled the SheetAsBackground problem as a multi-class classification task. Many kinds of methods like these have been tried and tested from Weka and OpenCV, but at the end the best solution seemed to be the random forest implementation of OpenCV with the following parameters:

- Maximal tree depth=25
- Maximal iteration for termination=100, epsilon for termination=0.1
- It should also be mentioned that for the purpose of obtaining not just one class label, rather a ranked list, a voting scheme was implemented upon the basic classifier.

2.2 Other Categories

The segmentation of interesting parts in the task of the other categories is not simple, and we think that it should be the first step for getting good descriptors. However, for the flower task we can make some assumptions, for example the color of the flowers is very important. This is why we collected color-gradient features for all these tasks.

Features: First, from a test photo, 9 new pictures are created corresponding to different color intervals, and one picture for the grey part. Fig. 1 illustrates this step. After this decomposition, gradient-based features were computed on all the 9 picture channels, namely, an intensity histogram, a gradient magnitude histogram, and a gradient angle histogram. These histograms (containing 10 bins each) were also normalized before being used as features in the classification process.

Classification Method: Since in most cases very little training data is available per species, we used nearest neighbor classifier to generate test results.

3 Own Results

For the experiments, we used 10-fold cross validation method to evaluate our models. As mentioned earlier, for scan-like leaf photos a random forest was trained, while for all the other problems nearest neighbour classifiers were used.

3.1 Our results on the training database:

On the scan-like leaves task we achieved an accuracy score of 55.84%, which is competitive with the results of the ImageCLEF task's performance from the previous year [1]. The accuracy scores attained on the other tasks are listed in Table 2. These accuracy scores are much better than a trivial classifier's result would be, however, it should be taken into account, that when the cross validation method divided the data into training and test sets, many similar photos from the same plant taken at the same time may be placed into both sets, and this fact introduced a positive bias in accuracy score.

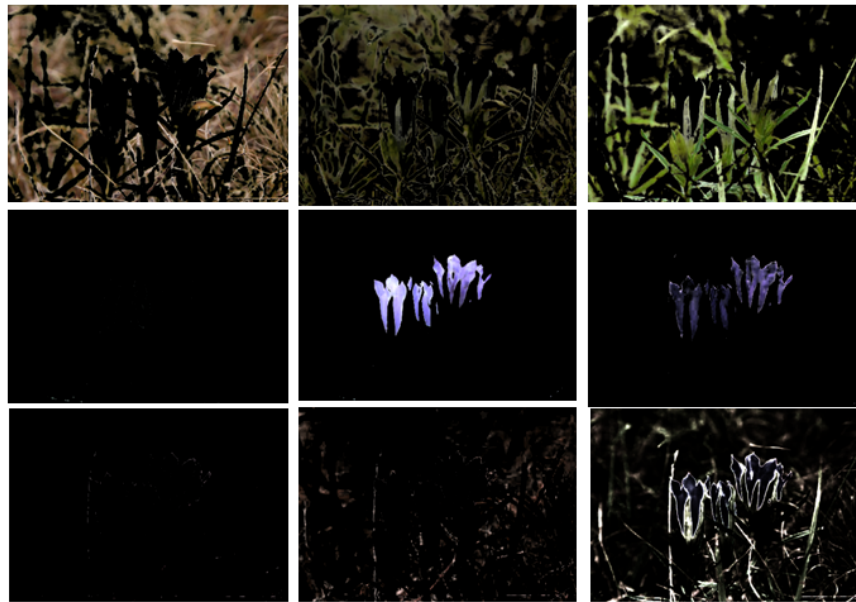


Fig. 1. Decomposing the original image into different color intervals, and into a grey part.

Table 2. Accuracies got for natural photos, using the training data and tenfold cross-validation. *However, with an observation oriented cross-validation, the accuracy values should be closer to the final results.*

Stem	Fruit	Flower	Branch	Entire	Leaf
39%	30%	28%	18%	20%	26%

3.2 Combination of The Scores

For the official evaluation, every participant had to aggregate the scores and ranks obtained for the different photos of the same plant. Our three aggregation strategies were very similar to each other, and they were all based on the following scheme:

- If there was a scan-like leaf photo, then only the result of this type was considered.
- If more than one scan-like image was available, then a weighted majority voting scheme was used to compute ranking scores.
- Otherwise, a score aggregation method was made from the nearest neighbor classifier’s results, using also some prior information.

The following prior information was used during the aggregating process:

- $N(s,c)$: the number of training data for a species and a category (stem, flower, etc.)

- $R(s,c)$: the ROC AUC value got from the training database by applying cross-validation for a species and a category
- $(S1,f1,S2,f2,\dots)(s,c)$: the species with which a specific species is frequently confused, and the frequency of this mistake (for a species and a category).

After the official evaluation, it turned out, that our aggregating strategies achieved almost the same results, hence it is enough to present only the first aggregation method of the three similar methods here. The aggregation rule is the following:

- When the NN classifier of a category "c" assigns an S0 class label, then a score of 1 is assigned to S0.
- Taking the prior information for $S0 \rightarrow (S1, f1, S2, f2, \dots)$ contains the frequent mistakes, S1 gets a rank of 1/4, S2 gets a rank of 1/5, etc.
- All the previous ranks are multiplied by a Trust (S0,c) factor, which is computed as: $(R(S0, c) - 0.5) * (1/(1 + \exp(5 - N(S0, c))))$. That means that the trust value is much higher when a ROC AUC value is much greater than 0.5, and the number of learning examples is much higher than 5.
- This is done for all the photos taken of a plant, and a weighted voting is applied to get the final scores and ranking.

4 Official Test Results

The results fit our expectations. Our main target was the scan-like leaf recognition, and the other categories were handled with a lower priority. Table 3 summarises our results on the different categories, Fig. 2 illustrates our results on the complex recognition task, and Fig. 3 shows our results for the category of LeafScan images, which are the most important for us. In the task of scan-like leaves we achieved a good result using features can be computed very efficiently and are described earlier.

Table 3. Accuracies of the test run.

LeafScan	Stem	Fruit	Flower	Branch	Entire Leaf
48.8%	7.7%	2.8%	8.8%	4.8%	2.9%

Acknowledgement

This work has been supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013).

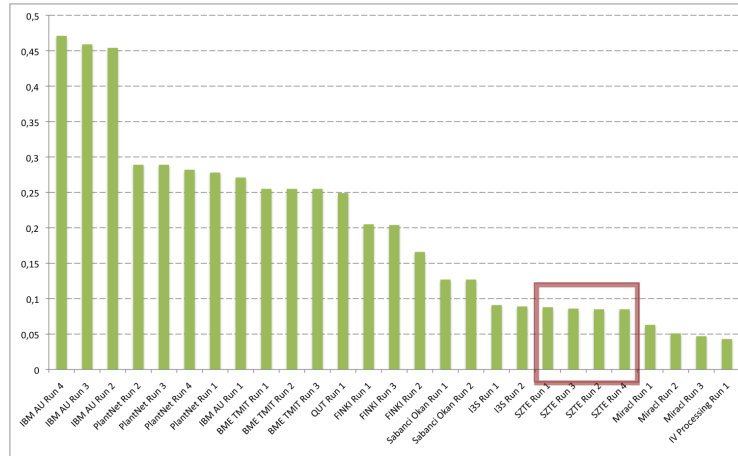


Fig. 2. Complex task: our solutions for the "non scanned leaf" tasks were very simple.

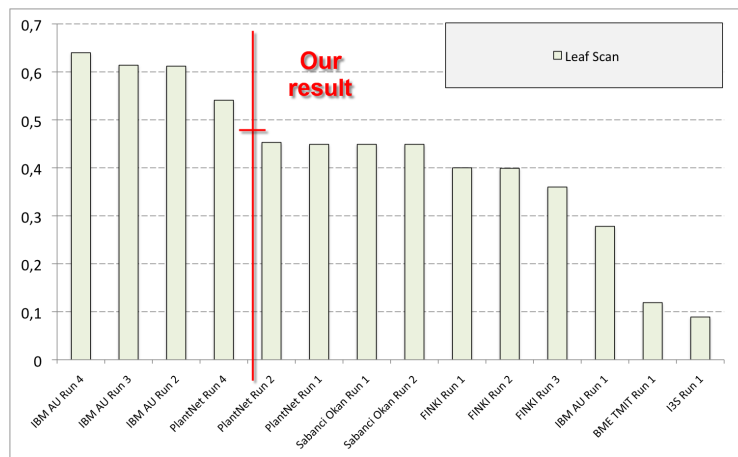


Fig. 3. Scan-like leaf task: since we were late to produce a bugless run, our result was not included in the official chart. However, our solution seems competitive (see the red bar at score=0.488).

5 Summary

Here, our goal was to define simple, but useful and powerful features for scan-like leaf images, and some features based on color and gradient for the other tasks, by which better results can be achieved than those for a trivial classifier. From the results of our experiments, we think that scan-like leaves can be classified quite well, and our aim is to improve the performance on this task by implementing an on-line learning solution based on a voting mechanism like intelligent crowdsourcing.

References

1. Goëau, H., Joly, A., Bonnet, P., Bakic, V., Barthélémy, D., Boujemaa, N., Molino, J.F.: The imageclef plant identification task 2013. In: Proceedings of the 2Nd ACM International Workshop on Multimedia Analysis for Ecological Data. pp. 23–28. MAED '13, ACM, New York, NY, USA (2013), <http://doi.acm.org/10.1145/2509896.2509902>
2. Goëau, H., Joly, A., Bonnet, P., Molino, J.F., Barthélémy, D., Boujemaa, N.: Lifeclef plant identification task 2014
3. Joly, A., Müller, H., Goëau, H., Glotin, H., Spampinato, C., Rauber, A., Bonnet, P., Vellinga, W.P., Fisher, B.: Lifeclef 2014: multimedia life species identification challenges
4. Mouine, S., Yahiaoui, I., Verroust-Blondet, A.: A shape-based approach for leaf classification using multiscale triangular representation. In: ICMR '13 - 3rd ACM International Conference on Multimedia Retrieval. ACM, Dallas, United States (Apr 2013), <http://hal.inria.fr/hal-00818115>, pl@ntNet (Agropolis fondation)
5. Wu, S., Bao, F., Xu, E., Wang, Y.X., Chang, Y.F., Xiang, Q.L.: A leaf recognition algorithm for plant classification using probabilistic neural network. In: Signal Processing and Information Technology, 2007 IEEE International Symposium on. pp. 11–16 (Dec 2007)