

Sabancı-Okan System at ImageClef 2013 Plant Identification Competition

Berrin Yanikoglu¹, Erchan Aptoula², and S. Tolga Yildiran¹

¹ Sabancı University, Istanbul, Turkey 34956

² Okan University, Istanbul, Turkey, 34959

{berrin,stolgay}@sabanciuniv.edu

erchan.aptoula@okan.edu.tr

Abstract. We describe our participation in the plant identification task of ImageClef 2013. We submitted one fully automatic run that uses different features for the uniform background (isolated leaves) and natural background (unconstrained photos) categories. Besides the category information, meta-data was only used in the natural background category. Our approach employs a variety of shape, texture and color descriptors. As in the previous years, we used shape and texture only for isolated leaves and observed them to be very effective. Our system obtained the best results in this category with a score of 0.607 which is the inverse rank of the retrieved class, averaged over all queried photos and users. As for the natural background category, we used a limited approach using a restricted set of features that were extracted globally due to lack of time, and obtained a score of 0.181.

Keywords: Plant identification, mathematical morphology, support vector machines.

1 Introduction

The ImageCLEF plant identification competition is organized every year since 2011 and aims to benchmark progress in the area of plant identification from photographs [3, 4, 2]. Similar to the previous years, the competition in 2013 consisted of identifying images of plants that were captured by different means: isolated leaves that were scanned or photographed on a uniform background comprised the *SheetAsBackground* category. Parts or full images of a plant taken on a natural background formed the *NaturalBackground* category. This category was further sub-divided as flower, fruit, entire, leaf and stem categories.

The organizers collected a large set of data from 250 different plant species over the course of several years. Part of this data formed the training set that was distributed to the participants along with the corresponding groundtruth. The remaining data was shared with the participants in order to collect their systems' responses, while the corresponding groundtruth was kept sequestered.

Submitted systems were scored in terms of the inverse average rank of the correct class for each submitted query. The details of this competition are described in [2].

2 Overview of the System

As a collaboration from two universities in Istanbul, we submitted a single fully automatic run (Sabanci-Okan-Run1) that uses different features for the uniform background (isolated leaves) category and natural background (unconstrained photos) sub-categories. The category information was obtained from the meta-data of the query image. This handling of queries in different categories was done to select the appropriate feature set for each group, but it also helped with the handling of this large task.

As in the previous years, we used shape and texture only for isolated leaves and observed them to be very effective. We had the best average score overall last year in both the automatic and manual categories [4] and this year we obtained the best score on the isolated leaf (uniform background) category.

For the natural background category, we used texture and color features for the flower, fruit and entire sub-categories; shape and texture for the leaf category; and only texture features for the stem category. The feature group selection was done based on our previous experiences in this problem and in order to increase generalization performance; it also helped reduce the time spent in feature extraction.

Meta-data was used only in the natural background category; specifically the month information was used to narrow down successfully the alternatives for fruit and flower categories.

3 Segmentation

Although segmentation is of crucial significance for content description, it has been used in our system only for isolated leaves and stems. In contrast, segmentation of photographs with a natural background is either not meaningful (i.e. the whole picture contains some part of the plant) or not an easy problem even though the background is well-defined (e.g. a plant photographed with the forest ground). In ImageCLEF 2012, we had used an approach where photos were aggressively segmented to leave only a single leaf in the image, in order to channel photographs to our successful isolated leaf recognition system [4]. While we believe that this is an interesting and complementary approach to one based on local invariants, it is limited in its potential as much information is discarded. This approach was skipped altogether this year due to lack of time.

Isolated leaves usually possess an uniform background, often with uneven illumination and sometimes shadow. Their segmentation has been conducted as in the past, using edge preserving morphological simplification by means of area attribute filters, followed by an adaptive threshold [9]. Moreover, contrary to flowers and fruit, it has been observed that the stem category contains mostly vertical or horizontal tree trunks that often occupy the majority of the image surface's center. Hence, in order to reliably obtain a background-free sub-image, we first determined the stem's orientation by controlling the horizontal and vertical derivatives' maxima, followed by cropping the corresponding central two third's of the image surface.

4 Preprocessing

Preprocessing stages were present only for the isolated leaves, in the form of size and orientation normalization. Specifically, we align the leaves' major axis with the vertical and normalize their height to 600 pixels, preserving the aspect ratio. Orientation normalization is realized through principal component analysis, with additional correction coming from the leaf petiole's location.

5 Features

Given the high visual variability of this year's dataset categories as well as the number of classes, feature extraction has become more challenging than ever before. Consequently, a large spectrum of descriptors has been evaluated, including shape, texture, color and local invariants. Moreover, considering the strong relation between seasons and image categories such as fruit and flowers, meta-data have also been exploited with great success. Here we summarize only the new descriptors, while the others have been explained in detail at the previous working notes [8, 9].

In particular, following the success of our past systems with scan and scan-like data (isolated leaves), it has been chosen not to greatly modify their descriptor set; instead we mainly optimized their parameters in order to cope with the higher class count. In addition, only one new descriptor was included in the feature extraction set: the *edge background/foreground histogram*. It is computed on the binary mask of its input and it consists in calculating the ratio of background to foreground pixels in a subwindow centered on each edge pixel. The normalized histogram of the said ratios constitutes the end feature vector.

As far as photographs are concerned, given the extreme variation of viewpoint and scale (especially w.r.t. the category "entire"), we resorted to using rather traditional, yet still reliable color descriptors. In particular, we employed the *color autocorrelogram* [6], computed in the LSH color space after a non-uniform subquantization to 63 colors (7 levels for hue, 3 for saturation and 3 for luminance). The color autocorrelogram describes the spatial correlation of colors. It consists of a table where the entry (i, j) denotes the probability of encountering two pixels of color i at a distance of j pixels.

We further employed the *saturation-weighted hue histogram* [5], where the total value of each bin $W_\theta, \theta \in [0, 360]$ is calculated as:

$$W_\theta = \sum_x S_x \delta_{\theta H_x} \quad (1)$$

where H_x and S_x are the hue and saturation values at position x and δ_{ij} the Kronecker delta function. As far as the color space is concerned, we have used LSH [1] since it provides a saturation representation independent of luminance.

And last, in order to exploit the effect of seasons on fruit and flowers, it has been decided to use the meta-data accompanying the visual samples, and specifically the month of acquisition.

6 Classifier Training and Evaluation

6.1 Data

The competition data consisted of a training set which was made available to all the participants, along with the corresponding groundtruth files, and a test set whose groundtruth was kept sequestered. The distribution of the data in each category and in each of these sets is shown in Table 1.

Table 1: Train and test dataset sizes.

| Category | Train | Test |
|--|--------|-------|
| SheetAsBackground (Isolated leaves) | 9,781 | 1,250 |
| NaturalBackground (Unconstrained photos) | 11,204 | 3,842 |
| Flower | 3,522 | 1,233 |
| Leaf | 2,080 | 790 |
| Entire | 1,455 | 594 |
| Fruit | 1,387 | 520 |
| Stem | 1,337 | 605 |
| All | 20,985 | 5,092 |

We split the available training data shown in Table 1 into train and validation subsets. The training set was used in training the corresponding classifier and the validation set was used as our internal test data for evaluating different features and algorithms. In order to help with the generalization capability, we tried to avoid having very similar images in the train and validation splits. Specifically, pictures from an individual plant were put in either the train or validation subset. The selection of the samples was done as described in [9]. As a result of this split, we obtained the train/validation subsets as shown in Table 2.

Table 2: Train and validation splits of the available training data.

| Category | Train | Validation |
|--|--------|------------|
| SheetAsBackground (Isolated leaves) | 7,867 | 1914 |
| NaturalBackground (Unconstrained photos) | 7,865 | 2,562 |
| Flower | 2,325 | 1197 |
| Entire | 1,455 | 594 |
| Fruit | 960 | 495 |
| Stem | 1,045 | 276 |
| All | 15,732 | 4,476 |

6.2 Classifiers

We used shape and texture only for isolated leaves in the *SheetAsBackground* category and observed them to be very effective. The length of the feature vector was 156 for this case, consisting of Fourier descriptors (50 of them),

in addition to various area and contour-based shape descriptors, and texture descriptors (106 altogether), many of them used in our previous system [9]. In the *NaturalBackground* category, we only used color features for the flower, fruit and entire sub-categories (autocorrelogram, saturation-weighted hue histogram and the month the picture was taken, for a total of 265 dimensions); shape and texture for the leaf category (same classifier as for isolated leaves); and only texture features for the stem category.

Feature extraction was done from the whole picture, except for the case of leaf images in the *SheetAsBackground* and the *NaturalBackground* categories, where segmentation step preceded feature extraction. The approach of using global features or using only color features is clearly not sufficient for unconstrained photos (e.g. flower, fruit, entire categories), however we did not have time to incorporate other methods based on local features.

The classifiers used for different categories were all trained with the training portion of the available data shown in Table 2, except for the leaf sub-category of *NaturalBackground* photographs. For this group, we used the same system developed for recognizing the *SheetAsBackground* category, after a simple segmentation of the image.

As classifier, we used a Support Vector Machine (SVM) classifier based on their good performance in many object recognition problem and used the SMO classifier inside the WEKA toolbox. The parameters for the SVM was set as $C = 10$ and a polynomial kernel of degree 2 after some limited tests with the validation set.

In Table 3, we give the cross-validation accuracy obtained while training a classifier using 10-fold cross-validation, as well as the accuracy of the same classifier on the validation subset. In the last column of this table, we also include the average inverse rank results published by the competition organizers for each category [2]. Here, a score of 1 indicates that all queries return the correct class as the top guess, while a score near 0 means the correct class is returned much later in rank.

Table 3: Cross-validation and validation set accuracies, along with the official test scores obtained by our system.

| Category | Features | Cross-Val. | Validation | Inverse Rank |
|-------------------|-----------------------|------------|------------|--------------|
| UniformBackground | Shape, texture | 93.77% | 70.64% | 0.607 |
| NaturalBackground | | | | 0.181 |
| Flower | Texture, color, month | 40.20% | 34.50% | 0.223 |
| Fruit | Texture, color, month | 51.33% | 43.64% | 0.194 |
| Entire | Texture, color, month | 34.23% | 29.50% | 0.174 |
| Stem | Texture | - | 9.30% | 0.106 |
| Leaf | Shape, texture | - | - | 0.049 |

7 Summary and Discussion

Participation into the ImageCLEF Plant Identification competition is an arduous task, especially when done in collaboration, with different people working in different parts of the problem. Last year we had to transfer partial results back and forth, since alternating steps of segmentation, preprocessing, feature extraction, and classification were done by different people in our small group. This year we streamlined this process a little better and concentrated on what we could accomplish the best. For that reason, we worked on isolated leaves the most, while some categories received minimal attention (e.g. leaves under the *NaturalBackground* category).

As the official results indicate, we obtained the best results in recognizing isolated leaves (*SheetAsBackground* category), with an average inverse rank of 0.607. This score roughly indicates that the correct class was returned as top-1 or top-2 alternative for the majority of queries, which is a promising result for the plant retrieval problem. In recognizing the unconstrained photographs in the *NaturalBackground* category, we started working on a system based on SIFT features [7]; although the initial results have been encouraging, the allocated time has not been sufficient for finalizing this module before the submission.

References

1. E. Aptoula and S. Lefèvre. On the morphological processing of hue. *Image and Vision Computing*, 27(9):1394–1401, August 2009.
2. B. Caputo, H. Muller, B. Thomee, M. Villegas, R. Paredes, D. Zellhofer, H. Goeau, A. Joly, P. Bonnet, J. Martinez Gomez, I. Garcia Varea, and M. Cazorla. Imageclef 2013: the vision, the data and the open challenges. In *Proc. CLEF 2013, LNCS*, 2013.
3. H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, J. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The clef 2011 plant image classification task. In *CLEF 2011 working notes*, Amsterdam, The Netherlands, 2011.
4. H. Goëau, P. Bonnet, A. Joly, I. Yahiaoui, D. Barthelemy, N. Boujemaa, , and J. Molino. The ImageClef 2012 plant identification task. In *CLEF 2011 working notes*, Rome, Italy, 2012.
5. A. Hanbury. Circular statistics applied to colour images. In *Computer Vision Winter Workshop*, pages 55–60, Valtice, Czech Republic, February 2003.
6. J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih. Image indexing using color correlogram. In *CVPR*, pages 762–768, San Juan, Puerto Rico, 1997.
7. David G. Lowe. Object recognition from local scale-invariant features. In *ICCV*, pages 1150–1157, 1999.
8. B. Yanikoglu, E. Aptoula, and C. Tirkaz. Sabanci-Okan system at ImageClef 2011: Plant identification task. In *CLEF (Notebook Papers/Labs/Workshop)*, 2011.
9. B. Yanikoglu, E. Aptoula, and C. Tirkaz. Sabanci-Okan system at ImageClef 2012: Combining features and classifiers for plant identification. In *CLEF (Notebook Papers/Labs/Workshop)*, 2012.