Late Information Fusion for Multi-modality Plant Species Identification[*]

Guillaume Cerutti[1,2], Laure Tougne[1,2], Céline Sacca[1,3], Thierry Joliveau[1,3], Pierre-Olivier Mazagol[1,3], Didier Coquin[4], Antoine Vacavant[5]

[1] Université de Lyon, CNRS
[2] Université Lyon 2, LIRIS, UMR5205, F-69676, France
[3] Université de Saint Etienne, EVS, UMR5600, F-42000, Saint Etienne
[4] LISTIC, Domaine Universitaire, F-74944, Annecy le Vieux
[5] Clermont Université , Université d'Auvergne, ISIT, F-63001, Clermont-Ferrand

Abstract. This article presents the participation of the ReVeS project to the ImageCLEF 2013 Plant Identification challenge. Our primary target being tree leaves, some extra effort had to be done this year to process images containing other plant organs. The proposed method tries to benefit from the presence of multiple sources of information for a same individual through the introduction of a late fusion system based on the decisions of classifiers for the different modalities. It also presents a way to incorporate the geographical information in the determination of the species by estimating their plausibility at the considered location. While maintaining its performance on leaf images (ranking 3rd on natural images and 4th on plain backgrounds) our team performed honorably on the brand new modalities with a 6th position.

Keywords: plant identification, confidence fusion, biogeographical information

1 Introduction

Considering the identification of a plant species, the study of its leaves is a common path, and undoubtedly the most fitting for an automatic system, given the relative visual stability of the planar, slowly growing objects that are leaves. However, they are not the organs on which was built the classification of the *Plantae* kingdom, which considers the temporal proximity of the species as a baseline grouping criterion. Flowers and fruits, representative of the plant's intrinsic reproduction properties are the most evident witnesses of this proximity, and have been used since the earliest classification works [12] as key features defining the species. Nevertheless, they constitute much touchier objects for automatic recognition, given their 3-dimensional complex structure, their high variability and short life cycles.

The main objective of the ReVeS project is to build a tree species identification system based on an explicit botanical description of leaf images. Unlike for the previous editions, the challenge laid by this year's identification task [7] was quite distant from this initial perspective, and required us to address different objects that are not suitable for the same analytic shape processing. The literature still provides numerous examples of methods dealing with flower [13,16], bark [19,9,6] and even fruit [15] images, using generally statistical descriptions

and rarely explicit modelling. Considering that most of the 5092 test images could be grouped by individual, producing sets of images of different organs for the same plant, our work focused especially on a strategy to combine those various sources of information, and on a way to make use of the GPS location provided to enhance the performance of the classification.

2   Image Processing and Feature Extraction

This year's database approximately doubled in size compared to last year, and it was separated into 2 big categories :

- **Sheet As Background** images, containing only leaves, and thus combining the categories **Scan** and **Pseudoscan** of previous editions [8], with the difference of containing some images that have previously fallen in the **Photograph** category, photographs of leaves on a plain but sometimes dark background.
- **Natural Background** images, being photographs of plant organs (or entire plants) in their natural environment. Only one type of content (**Leaf**, **Flower**, **Fruit**, **Stem** or **Entire**) is given for each image, and no distinction is made between images containing only one whole organ, part of an organ, several organs more or less in focus, or even different types of organs.

Among the 20985 images forming the Train database, we could find 250 plant species, but a distinction can be made between tree species (127 generally ligneous, taller than 1.2 m plant species) that were the ones considered in the previous editions and the ones represented in the SheetAsBackground category, and the 123 rather herbaceous remaining species. This distinction is important for the identification, the tree species presenting for example bark in the Stem images, that may be analyzed in a specific way, more accurately than a generic stem image.

The Test database contained 5092 images with the same information on the content, but with no indication whatsoever whether the species was a tree or not, which forced us to make assumptions and selections for the good running of our methods.

2.1   Manual Image Formatting

Before any actual processing, we had to perform some transformations on the Natural Backgroung images to convert them into suitable entry points for our algorithm. As mentioned previously, our objective is to develop a mobile application for identification, which allows a certain part of interaction with the user, by giving guidelines for the picture taking or directly asking for help in the image processing phase though the tactile interface. The procedure we applied on the images simply simulates what we would have asked of the supposed user of the application.

Most of the Leaf, Fruit and Flower images contained more than one example of the desired object. A first selection was done by placing all the images for which no object big and in focus enough could be found in the Entire category, and keeping the others. We subsequently cut those images to a region of interest containing only one object, and rotated them (90 or 180 rotations only) so that the resulting images correspond to what input a user asked to take a photograph
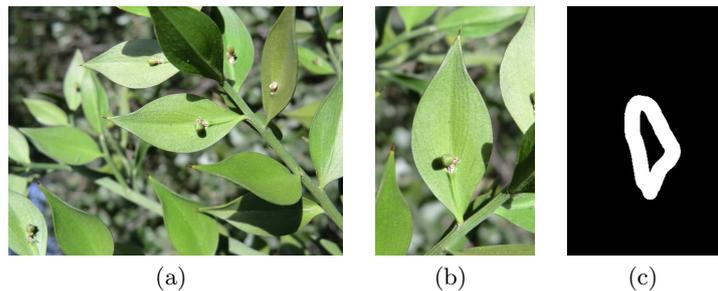
(a)   (b)   (c)

Fig. 1. Manual transformation of a natural background leaf image (a) : rotated and cropped (b) and user mark (c)

of one, vertically oriented, organ would have given. A typical example of this process can be found on the Figure 1 (b).

In addition to this first manipulation, we drew on top of each of the obtained images a rough mark inside the object that is used as an initialization of our segmentation algorithms. Note that in the case of leaves where the methods are the most developed, a distinction is made between simple and compound leaves : in the first case, one single mark is required whereas we require three distinct marks in three of the top leaflets for second type of leaves. An example of such quick colouring, as it would be done with the finger on a mobile device, is shown in Figure 1 (c).

For Stem images, where no segmentation process was considered, we simply rotated and cut the images to obtain a vertical view containing only bark pixels. The images clearly corresponding to non-tree species were discarded to the Entire category, as well as the non-tree leaves and fruits. Unfortunately, we did not have the time to do anything with the images of this extended category.

2.2   Leaf Understanding

Leaf images were processed with almost the same method as last year [2] with the idea of introducing some modelling of the leaf. This is a purpose that has proven to work quite efficiently [1,20]. The goal here is to represent faithfully the botanical properties of the leaf, in an attempt to model the analytic determination process used by botanists. For the recall, a leaf shape model is used as both a pre-segmentation and a description of the global shape [5], and after the segmentation by a constrained active contour, descriptors accounting for the margin [3], base and apex shapes are computed.
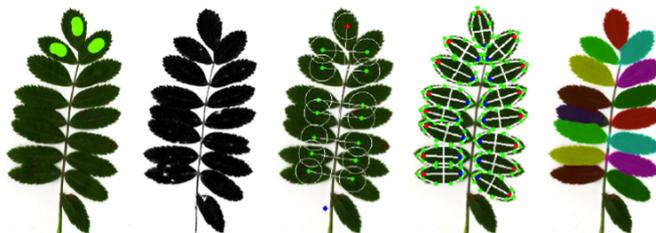

Fig. 2. The use of deformable templates to segment and describe compound leaves

The main differences concern the processing of compound leaves in which we try to describe all the leaflets at the same time. This is achieved by estimating the number and location of the leaflets, and jointly modelling their global shape, using highly variable deformable templates [4], and then segmenting each leaflet independently with multiple region-based active contours. Figure 2 illustrates the running of this new process.

Another improvement lies in the representation of the margin that now uses a string-like structure to keep track of the spatial repartition of teeth along the leaf contour. When our former descriptors used aggregative methods (computation of histograms, or averaging of curvature-scale measures) this new description is more discriminant as it can represent different levels of dentition or different orderings of teeth.

2.3   Fruit, Bark and Flower Image Analysis

The methods we used on the other organs were not nearly as dedicated and thought through as the ones elaborated for leaves, as they are not our main concern. We made the decision to rely on statistical shape descriptors and on vocabulary learning to produce color, texture and interest point based features. The only particularity came from the fact that in the case of Flower and Fruit images, a segmentation had to be performed first for these descriptors to be extracted correctly.

Fruits and flowers may be easier to isolate in an image given how their colors stand out, but they still require some advanced processing [16,18] or some modelling [13] to be segmented. The segmentation algorithm we used is derived from the one implemented for leaves. However, given the impossibility to model efficiently the shapes of fruits and flowers in the images, it relies only on the color dissimilarity part of the method. Starting from the user-made colouring, the average expected color is propagated throughout the image, leading to a map of the dissimilarity of image pixels to a local color model for the organ [2]. This dissimilarity map is then simply thresholded to produce a segmentation of the desired object. Some results of this algorithm are shown in Figure 3.
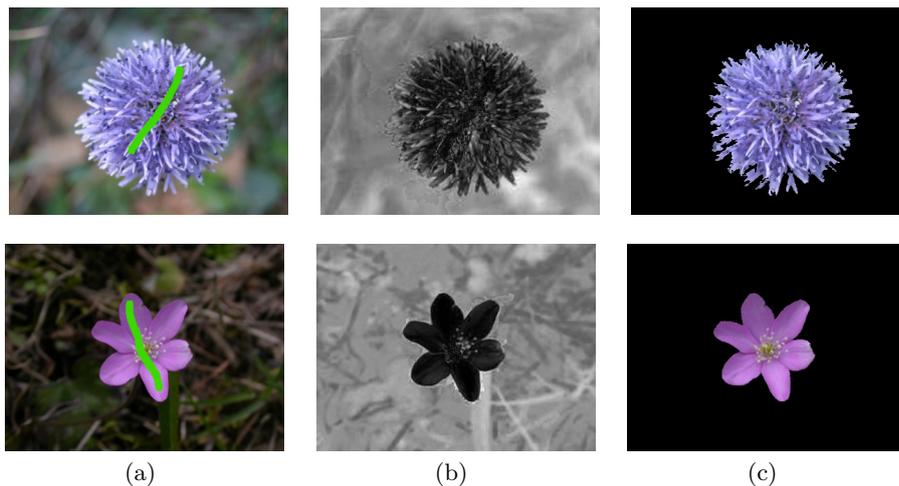

| (a) | (b) | (c) |

Fig. 3. Segmenting flowers : user mark (a), dissimilarity map (b) and segmentation (c)

Bagging features

To represent the color information, we used a bag-of-colors-like feature based on an organ-specific color vocabulary, much like what can be found in other works considering flowers [13]. This vocabulary was learned in a first phase by clustering the L*a*b* colors encountered in the considered (segmented) regions of all images into a predefined number of reference colors. Then the pixels for a given image are associated with one word of the color vocabulary, and the occurrences of each word are counted to form a histogram, later normalized by the size of the region. This generates fixed-sized, comparable, and even averageable features, that are displayed in Figure 4. Such color features were extracted only for Fruit and Flower images, with respective vocabulary sizes of 100 and 200 color words.
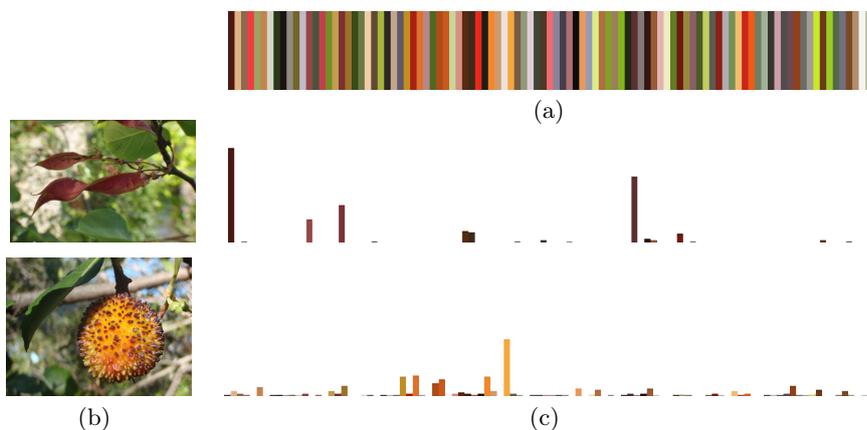

Fig. 4. Representing fruit colors : color vocabulary (a), original image (b) and bag of colors (c)

The same idea was adopted for texture description keeping for each pixel in the region, instead of a 3-channel L*a*b* vector, a vector of 30 values corresponding to the responses obtained by filtering the image with the commonly employed Gabor wavelets [9] of 6 different orientations and 5 different sizes. Then again a vocabulary is learned based on these features and bags of textures used to describe the image. This was performed on Stem and Fruit images, with a vocabulary size of 500 for each.

Finally, we computed the classical bag of visual words by repeating the same procedure again, this time on SURF descriptors obtained on detected keypoints (limited to 500 per image, for computational reasons). The only difference is that the histograms are not normalized by the number of keypoints extracted on the image, otherwise it is in every respect similar. These bag of words were used on Stem images only, with a vocabulary of 1000 visual words.

Shape descriptors

To capture the shape properties of the segmented organs, we decided to rely on well established, robust shape descriptors, commonly used in the literature for object recognition purposes, namely :

- Centered moments (8 values)
- Eccentricity (1 value)
- Hu moments (7 values)
- Zernike moments (9 values)

These values were computed on the binary segmentation images resulting from the algorithm described above. Therefore, they were only extracted on Fruit and Flower images.

3   Species Classification

We chose to view the problem of species identification as a classification problem rather than an image retrieval one, and subsequently trained classifiers to perform the recognition. Unfortunately, the time dedicated to this task was too short to reproduce the random forest classification that gave us the best results last year [2] and we had to rely on a more naive process. Furthermore, the content of the database somehow pushed us to reconsider the decision making, with the obvious interest of combining the different views and organs of the same plant individual to improve the accuracy.

3.1   Naive Distance-based Classification

For each set of the aforementioned descriptors, and for each considered species, we were able to model the distribution of the parameters by gaussians, evaluating the means and standard deviations from all the examples. This was also what we did for the histogram-like features that are maybe not the most suitable for this type of approach. The only difference concerns the strings describing leaf margins, non-vectorial structures that had to be addressed differently, simply by estimating an average string per species.

   The set of considered species, on which the training was performed and that are likely to appear in the classification results, was also different depending on the organ :

- **Leaf** (L) : 126 tree species, separated in simple-leaved and compound-leaved
- **Fruit** (F) : 44 tree species with enough examples
- **Flower** (f) : 139 species, essentially herbaceous
- **Bark** (B) : 78 species, only ligneous
- **Entire plant** (p) : No training, by lack of time

   To classify a new image, we simply evaluate the distances $d_o(x, S)$ of the descriptors extracted on it to the gaussian models of the retained species (S_o) for the corresponding organ o. The metric we used for histogram-like features is a simple $L2$ distance, for the string structures, a normalized Levenshtein, and for the vector features a so-called ellipsoid distance, designed to take variability into account, without distorting the parameter space as would the Mahalanobis distance. That distance represents the distance of a point x to the surface of the ellipsoid defined by the center μ and the axes of the covariance matrix Σ. Using the Mahalanobis distance $\|x - \mu\|_M = (x - \mu)\Sigma^{-1}(x - \mu)^T$, the metric we compute can be written as :

$$\|x - \mu\|_E = \|x - \mu\|_2 \cdot \max\left(1 - \frac{1}{\|x - \mu\|_M}, 0\right) \qquad (1)$$

   Each one of the N_D sets of descriptors extracted on the image thus produces a distance, that is then normalized by the average distance to the correct class (learned for each feature during the training phase). Once those terms are on the same scale, they are just summed, and the set of all considered species are ordered according to this last measure. This way, for each image, we produce a ranked list of species along with a distance value measuring their similarity.

3.2 Individual-based Information Fusion

This ranking could be enough to give a suitable answer per image, but it is much more interesting to try and make use of the different modalities available for a same individual plant. The images corresponding to a same value of the IndividualPlantID are actually different views of the same plant, taken at the exact same location. Consequently, the answer to each of these images needs to be the same, and every element that has to be taken into account to produce it with as much accuracy as possible.

To perform an efficient fusion of these various data sources, we used a weighted sum of confidence measures obtained from the distances to species issuing from the different classifiers. The fusion process is illustrated in Figure 5. An individual may be associated to a certain number of images (up to 24) of different organs. Each one having been processed independently by the according classifier, we retrieve for each of the n_o images of the organ o the distances $d_o\,(1..n_o)$ to the different species considered in the context of this organ.
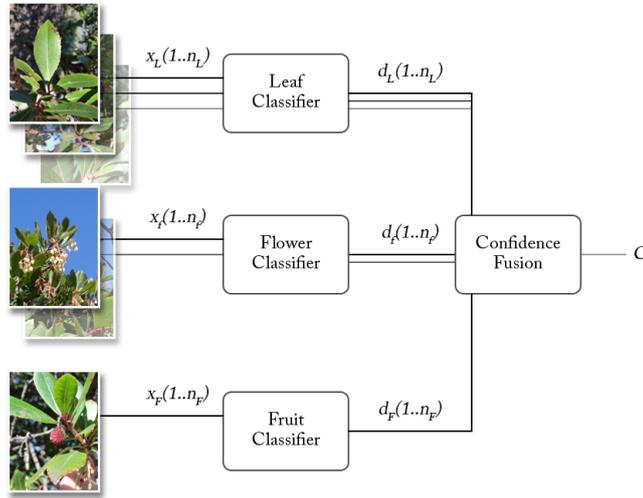

Fig. 5. Example of the individual-based fusion process

The distances are first converted into confidence measures. Each distance value being a sum of $N_{D,o}$ terms expected to be equal to 1 for the correct species, we define the confidence score associated to the distance as following :

$$C_o(x, S) = \begin{cases} \exp - \left(\dfrac{d_o(x, S)}{N_{D,o}} \right)^2 & \text{if } S \in (S_o) \\ 0 & \text{if } S \notin (S_o) \end{cases} \qquad (2)$$

In the end, even if distances were computed for only a subset of all the considered species, we now have for each image a set of confidences for each and every species. The value is 0 for those that were not learned but they are still present in the possible answer list. These confidence values all lie between 0 and 1 and the normalization of the distance by $n_{D,o}$ makes sure that they constitute comparable measures even for different organs. However, the results we obtained on the different modalities tended to show that some are more

reliable than others. With this in mind, the confidence measures are weighted by an organ-wise coefficient w_o derived from the classification rates we could observe on the training base, in a common attempt of giving more importance to reliable sources of information [17].

Finally, we used an additive process on these weighted confidences, in order not to penalize too much the species that were absent for a given organ. The sum of weighted confidences is then normalized again by the sum of weights to remain between 0 and 1. The final confidence score for a species S is then computed as :

$$C(S) = \frac{\sum_{o \in \{L,f,F,B\}} \left(\sum_{i=1}^{n_o} w_o \cdot C_o(x_i, S) \right)}{\sum_{o \in \{L,f,F,B\}} \left(\sum_{i=1}^{n_o} w_o \right)} \tag{3}$$

With these similarity measures, we can produce a new ranking of the species, this time all the 250 of them, that will be the same for all the images of the considered individual. The information available on each of the image is then taken into account, with a degree of trust depending on the modality, and the ordering that would have been obtained on one image only is then refined and hopefully improved by the content of the other associated images. Of course in the case when there is only one image per individual, the order remains unchanged, but when different organs are available, or different leaves of the same plant, the species appearing regularly as close will be put up front.

3.3   Introducing Biogeographical Knowledge

Geographical and environmental factors play a big part in the development (and survival) of plants, and not all species are likely to be found under every conditions. Knowing where a plant was observed, and what species may grow in that particular location might constitute a great help for identification. With this perspective, it was a core objective of our project to come up with methods to benefit from the location (easily accessible through GPS on mobile devices, for example) to enhance the identification by knowing in advance the plausibility of species.

In this respect, a large-scale work of digital cartography was undertaken (focusing on the French metropolitan territory) to map some environmental parameters that are related to the presence of species, with a resolution of 90 meters. Indeed, variables such as temperature, altitude, precipitation, humidity, exposition have a key influence on the repartition of species. Along with these raw explanatory variable, we also considered more global, integrative values elaborated by botanists and phytosociologists :

- Altitudinal zonation, or vegetation levels as defined in phytogeography [11], 9 zones for different altitudes, temperatures and orientations (modelled using CNRS vegetation maps at scale 1:200,000)

- GRECO (great ecological regions) from the French forest institute [10], 10 regions of homogeneous geological and environmental characteristics identified by letters

The images in the PlantView database were all referenced in terms of longitude and latitude, and most of them issued from the french territory, making
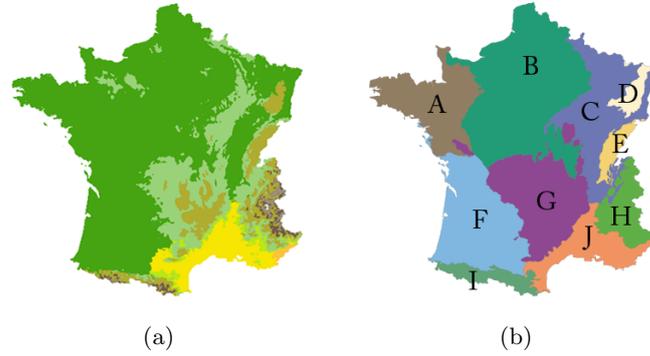
(a)   (b)

Fig. 6. Maps of two integrative geographical values on metropolitan France : Vegetation Levels (a) (▪ Snow, ▪ Alpine, ▪ Subalpine, ▪ Montane, ▪ Foothill, ▪ Planar, ▪ Supramediterranean, ▪ Mesomediterranean, ▪ Thermomediterranean,) and Ecological Regions (GRECO) (b)

them suitable to experiment this geographical approach. Unfortunately, there was a big incertitude about the precision of this location, whether it was the place where the image was actually acquired or simply uploaded. In the case of scanned leaves, is it the location of the tree or the scanner that is supplied? Therefore, trying to include local parameters like the raw measurements hardly made any sense, and we chose to take only levels and regions into account. The maps for those two variables can be seen in Figure 6.

To derive the plausibility of finding a species at a given location, we did not use the training base but included external botanical knowledge about the species. Based on the repartition knowledge and maps one can find in floras [14], we associated for each region and each zone a presence value of the species. However, this task was only performed for native tree species, leaving almost half of the challenge scope empty of information. The presence levels were simply 0 or 1 for the vegetation zones but 0, 1, 2, 3, 4 and 5 (in increasing order of plausibility) for the ecological regions.

Then, for given values of the region index g and of the vegetation level l, it is possible to compute a numerical value of plausibility for each species given its theoretical repartition. Knowing the presence values of the species in the region $p(S, g)$ and the level $p(S, l)$ we simply derive an additional plausibility score :

$$P(S,g,l) = \frac{1}{2} \begin{cases} 1 \text{ if } p(S,g) = 5 \\ 0.95 \text{ if } p(S,g) = 4 \\ 0.9 \text{ if } p(S,g) = 3 \\ 0.8 \text{ if } p(S,g) = 2 \\ 0.7 \text{ if } p(S,g) = 1 \\ 0.5 \text{ if } p(S,g) = 0 \end{cases} + \frac{1}{2} \begin{cases} 1 \text{ if } p(S,l) = 1 \\ 0.5 \text{ if } p(S,l) = 0 \end{cases} \tag{4}$$

For visualisation purposes, we computed this plausibility for different species all over the French territory, for which the values of vegetation levels and ecological regions were mapped. The result can be seen as a repartition map of the species, in which the places where the species is most plausible appear in darker green, on those where it will certainly not be found in lighter green. Examples of such maps are shown in Figure 7.
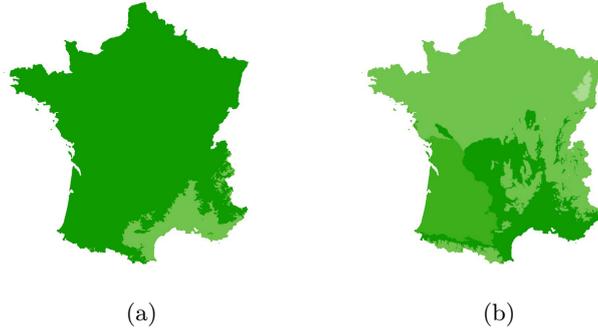
(a)   (b)

Fig. 7. Plausibility maps obtained for two tree species : *Fagus sylvatica*, L. (a) and *Acer monspessulanum*, L. (b)

For the species on which we had no prior repartition knowledge, we ensured the geography would not have a big influence on classification by setting the $P(S, l, g)$ to 1 for all values of l and g. This of course is not an optimal solution, and actually favours those species, since they will at least be as plausible as the other species, whatever the location.

The way we insert the geographical knowledge into the classification process is fairly simple. Given the coordinates (lon; lat) of the observation, we can extract from the mapped data the values $g(\text{lon}, \text{lat})$ and $l(\text{lon}, \text{lat})$ of the corresponding GRECO and level. Then for each species S the confidence score used to rank it simply becomes :

$$C_G\left(S, \text{lon}, \text{lat}\right) = P\left(S, g(\text{lon}, \text{lat}), l(\text{lon}, \text{lat})\right) \cdot C(S) \qquad (5)$$

The introduction of this geographical plausibility leaves unchanged the confidence score for the species that are perfectly adequate with the observation's location. However, it will penalize the species which are not, making them drop in the rankings and conversely letting the more plausible species make their way up to the first answers. It is a way to make a rearrangement of the top answers without actually overshadowing the information extracted from the image. It is in any case coherent with the idea that geography should not contradict the evidence found in the content of the image, but simply provide a lighting on the results of its analysis to make a better final decision.

4 Results and Comments

4.1 Training Results

The classifications of the different plant organs do not perform all the same, the most effective one naturally being the classification of leaves on plain backgrounds. As a matter of fact, leaves are the organ for which most work has been dedicated, and the methods used to process them were the most extensively tested. It is then only natural that the experimental results makes them appear clearly as the best performing modality.

The Figure 8 sums up the experimental classification rates obtained for the 4 modalities we addressed. Those scores were measured on the Train database, in a cross validation process (validation rates are displayed, with the rates obtained on the training data left in lighter color). If the Leaf category (126 classes, 4 descriptor sets) clearly benefits from the efforts put in the segmentation and the botany-inspired description (nearly 85% of presence of the correct species in the
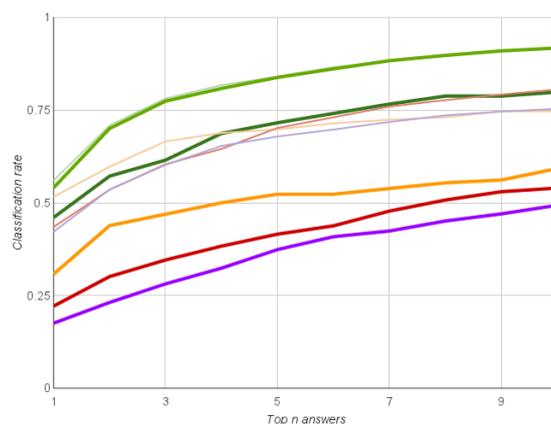
Fig. 8. Classification performance for the different modalities : ■ Leaf (SheetAsBackground), ■ Leaf (NaturalBackground), ■ Fruit, ■ Stem, ■ Flower

top 5 answers, up to 70% in the case of natural images), the other modalities are not as far behind as expected.

The Fruit category (44 classes, 3 descriptor sets) is the second best performing (more than 50% at 5 answers) but this performance has to be moderated by the fact the the number of classes is relatively small. On the other hand the Stem category (78 classes, 2 descriptor sets) performs honorably (more than 40% at 5 answers) given the number of species and the rawness of the description used. Same goes for the Flower modality (139 classes, 2 descriptors) for which it might be interesting to add that most of the accuracy (37% at 5 answers) actually comes from the simple color feature we used. All in all, these results underline the limits of the descriptions (hastily) implemented, and in the case of fruits and flowers the difficulties of the segmentation, given the low contribution of the shape description. It is also crucial to point out the greater abstraction power of the dedicated description created for leaves compared to the generic features used on the other organs ; this is visible in the blatant gap between the curves obtained on the training and validation sets.
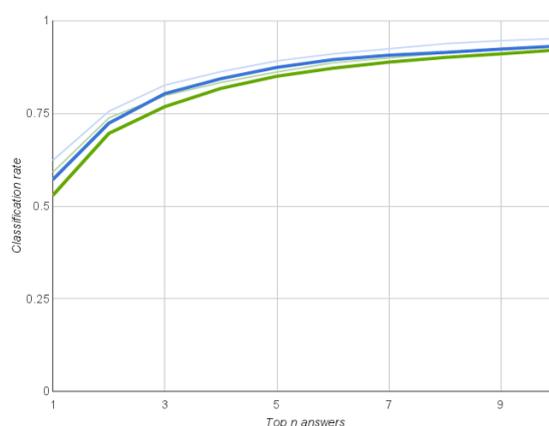

Fig. 9. Influence of geography on tree leaf classification : classification scores on plain background images of native species leaves ■ without and ■ with geographical plausibility

A very interesting point is to measure the contribution of geography in the performance of the classification. To do so, we considered only leaf images from tree species that are native in France, and on which we had the necessary theoretical geogaphical information. The Figure 9 shows the comparison between a classification where the geographical plausibility of the species is taken into account and one where it is not. More than the actual rates we obtain, it is clearly appearing that geographical knowledge leads to a better classification, and simply improves the ranking of the correct species. The Figure 10 shows this improvement by measuring the difference between the rank of the correct species in the classification without geographical information and with it, a difference that goes in the right direction for a major share of examples for which the ranking could still be improved. Even with unprecise location (obviously for scanned images) the region and level prove to give enough information to refine the decision in a more accurate way.
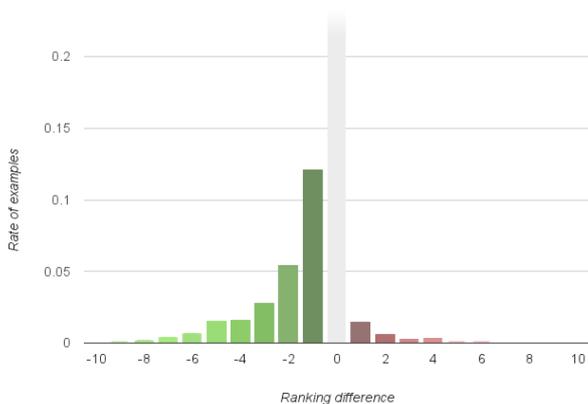
(a)

Fig. 10. Rank difference of the correct species when adding geographical plausibility

4.2 ImageCLEF 2013 Results

We submitted two distinct runs this year, both relying on the same visual analysis process but differing in the fusion phase. They both make use of the IndidualPlantID field, but the second run (LirisReVeS_run2) additionally used the longitude and latitude coordinates to compute the geographical plausibility of the species and use it to rearrange the results of the image processing phase.

Both runs were then mixed (visual + textual) runs, with an automatic method on the SheetAsBackground category and manual initalizing on the NaturalBackground category. To be exhaustive, the images remaining untreated, and not belonging to a processed individual were submitted as specimens of *Ginkgo biloba*, L. with an absolute confidence of 1.

As expected, the SheetAsBackground category, containing almost only leaves, is the one on which our leaf-focused algorithms gave the most satisfying results. As shown in Figure 11, we reach a score of 41% bringing us in 4th place among the 12 teams that have submitted runs.

Concerning the NaturalBackground images, the results are with no surprise lower, but our ranking on the Stem images (3rd place, yet with a score of 16%)
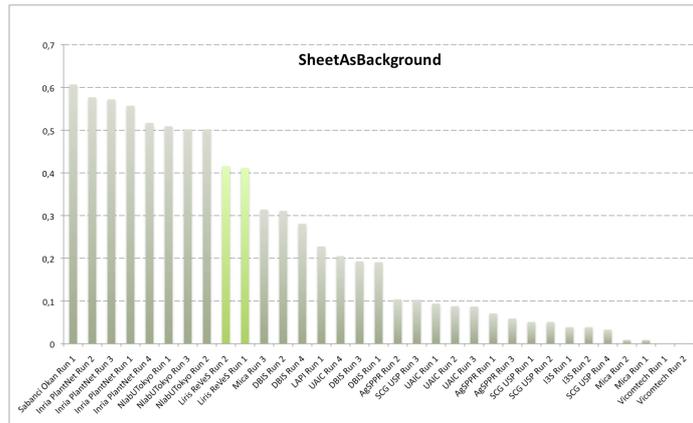
Fig. 11. Results on the Leaf (SheetAsBackground) images : 4th team on 12

was unexpectedly quite interesting. Our low performance on fruits and flowers (respectively 8% and 10%) is a sign, if needed, that the representation of such organs requires more dedicated descriptors, and would probably benefit from the introduction of some kind of model to focus on the discriminant criteria from a botanical point of view. And finally, our 3rd place on the Leaf images corresponds to our previous rankings, with again the same limitations (only broad-leaved tree species considered, some images out of our processing scope, learning performed on plain background images) that might explain the rather low raw score value of nearly 17 %.
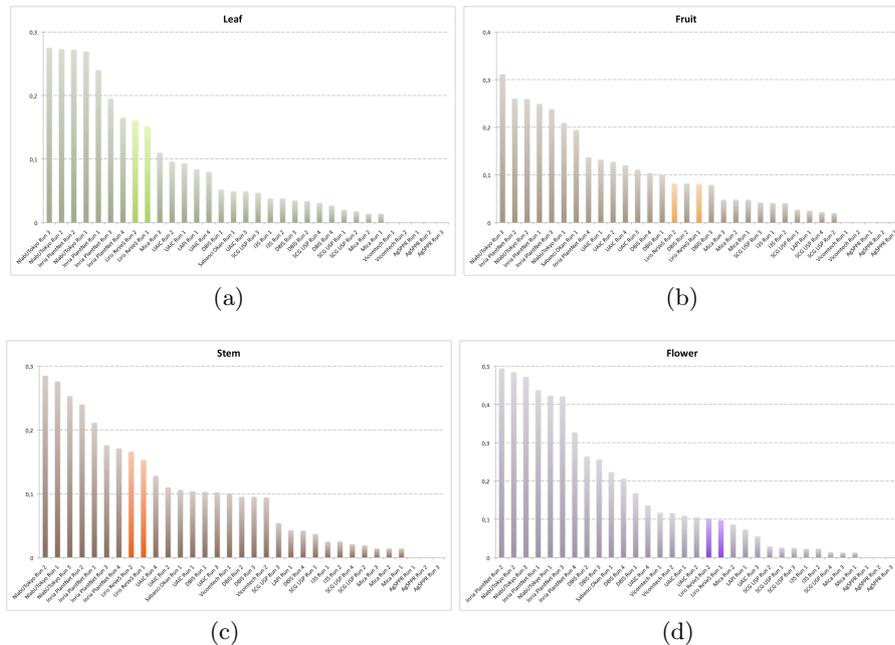

(a)

(b)

(c)

(d)

Fig. 12. Results on the NaturalBackground images : 3rd on Leaf images (a), 6th on Fruit images (b), 3rd on Stem images (c) and 7th on Flower images (d)

Concerning geography, the run LirisReVeS_run2 gave slightly better results than the run LirisReVeS_run1 from which the location information was completely left out. But the difference (almost 1% in every category) does not appear as clearly as it did on native species, and it is hard to draw any conclusion. There were however too many manifest flaws in the context of this task to make this a valid experiment for us. First, we had theoretical phytogeographical data at hand for hardly half of the 250 species, which made the plausibility values we computed only half-relevant. More importantly, because of the doubt about the precise location of the observation, we were unable to use the local geographical parameters that might be more accurate. The fact that many images come from cities, where it is more difficult to assess whether the plant grew in its natural environment, also made the problem a bit less clear. It is nevertheless an encouraging step towards the use of geographical information to help the identification, a direction that will in any case be further investigated.

5   Conclusion

This year's task constituted a real new challenge for our leaf-oriented project, and it came as no surprise that the generic, mechanical, methods we had the time to implement on fruits, barks and flowers were largely outperformed by the now mature leaf analysis process put into practice in the **Folia** application[§]. It was however a good chance to experiment with a late fusion mechanism which seemed to bear fruit. It was also the opportunity for a first testing of the novel approach proposed to combine the image information with geographical knowledge. The results of this last contribution are engaging, and we believe that with an adequate scope of species and a greater certainty on location it may constitute a real asset for field plant identification.

Acknowledgements

References

1. A. Arora, A. Gupta, N. Bagmar, S. Mishra, and A. Bhattacharya. A plant identification system using shape and morphological features on segmented leaflets. In *CLEF (Notebook Papers/Labs/Workshop)*, 2012.
2. G. Cerutti, L. Antoine, Violaine Tougne, L. Mille, Julien Valet, D. Coquin, and A. Vacavant. Reves participation - tree species classification using random forests and botanical features. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
3. G. Cerutti, L. Tougne, D. Coquin, and A. Vacavant. Curvature-scale-based contour understanding for leaf margin shape recognition and species identification. In *VISAPP 2013 - Proceedings of the International Conference on Computer Vision Theory and Applications*, 2013.

[§] *folia* : https://itunes.apple.com/app/folia/id547650203

4. G. Cerutti, L. Tougne, J. Mille, A. Vacavant, and D. Coquin. A model-based approach for compound leaves understanding and identification. In *IEEE International Conference on Image Processing*, 2013.
5. G. Cerutti, L. Tougne, A. Vacavant, and D. Coquin. A parametric active polygon for leaf segmentation and shape estimation. In *7th International Symposium on Visual Computing*, 2011.
6. S. Fiel and R. Sablatnig. Automated identication of tree species from images of the bark, leaves and needles. In *Proceedings of the 16th Computer Vision Winter Workshop*, 2011.
7. H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, and J.-F. Molino. The imageclef 2013 plant identification task. In *CLEF 2013 Working Notes*, 2013.
8. H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthelemy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The clef 2012 plant images classification task. In *CLEF 2012 Working Notes*, 2012.
9. Z.-K. Huang, C.-H. Zheng, J.-X. Du, Z.-H. Quan, and S.-B. Guo. Bark classification based on gabor filter features using rbpnn neural network. In *International Conference on Neural Information Processing*, pages 80–87, 2006.
10. Inventaire Forestier National. *Les Sylvoécorégions (SER) de France métropolitaine*. IGN, 2011.
11. P. Julve. Carte phytogéographique de la France. *Cahiers de Géographie Physique*, 13, 1999.
12. C. v. Linne. *Species plantarum : exhibentes plantas rite cognitas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas*. Holmiae : Impensis Laurentii Salvii, 1753.
13. M.-E. Nilsback and A. Zisserman. Delving into the whorl of flower segmentation. In *British Machine Vision Conference*, volume 1, pages 570–579, 2007.
14. J. Rameau, D. Mansion, G. Dumé, J. Timbal, A. Lecointe, P. Dupont, and R. Keller. *Flore forestière française: Guide écologique illustré*. Institut pour le Développement Forestier, 1989.
15. A. Rocha, D. C. Hauagge, J. Wainer, and S. Goldenstein. Automatic fruit and vegetable classification from images. *Computers and Electronics in Agriculture*, 70(1):96 – 104, 2010.
16. T. Saitoh and T. Kaneko. Automatic recognition of blooming flowers. In *International Conference on Pattern Recognition*, volume 1, pages 27–30, 2004.
17. M. Siegel and H. Wu. Confidence fusion. In *Proceedings of ROSE'2004 (2004 IEEE Workshop on Robotic Sensing)*, 2004.
18. L. Suta, M. Scuturuici, V.-M. Scuturici, and S. Miguet. Pathos: Part-based tree hierarchy for object segmentation. In *International Conference on Computer Analysis of Images and Patterns (CAIP)*, aug 2013.
19. Y.-Y. Wan, J.-X. Du, D.-S. Huang, and Z.-R. Chi. Bark texture feature extraction based on statistical texture analysis. In *International Symposium on Intelligent Multimedia, Video and Speech Processing*, pages 482–485, 2004.
20. B. A. Yanikoglu, E. Aptoula, and C. Tirkaz. Sabanci-okan system at imageclef 2012: Combining features and classifiers for plant identification. In *CLEF (Online Working Notes/Labs/Workshop)*, 2012.