Multi-Organ Plant Identification

Hervé Goëau
INRIA, Imedia Team
Rocquencourt, France
herve.goeau@inria.fr

Pierre Bonnet
INRA, UMR AMAP
Montpellier, France
pierre.bonnet@cirad.fr

Julien Barbe
INRA, UMR AMAP
Montpellier, France
julien.barbe@cirad.fr

Vera Bakic
INRIA, Imedia Team
Rocquencourt, France
vera.bakic@inria.fr

Alexis Joly
INRIA, Zenith Team
Montpellier, France
alexis.joly@inria.fr

Jean-François Molino
IRD, UMR AMAP
Montpellier, France
jean–francois.molino@ird.fr

ABSTRACT

This paper presents a new interactive web application for the visual identification of plants based on collaborative pictures. Contrary to previous content-based identification methods and systems developed for plants that mainly relied on leaves, or in few other cases on flowers, it makes use of five different organs and plant's views including habit, flowers, fruits, leaves and bark. Thanks to an interactive and visual query widget, the tagging process of the different organs and views is as simple as drag-and-drop operations and does not require any expertise in botany. All training pictures used by the system were continuously collected during one year through a crowdsourcing application that was set up in the scope of a citizen sciences initiative. System-oriented and human-centered evaluations of the application show that the results are already satisfactory and therefore very promising in the long term to identify a richer flora.

Categories and Subject Descriptors

H.5.2 [**Information Systems Applications**]: Information interfaces and presentation—*User Interfaces*

Keywords

Interactive plant identification, digital botany, plant image retrieval by organ, leaf, flower, fruit, bark, habit

1. BRIDGING THE TAXONOMIC GAP

The integration of life sciences and computer sciences has a major role to play towards managing and analyzing cross-disciplinary scientific data at a global scale. More specifically, building accurate knowledge of the identity, geographic distribution and uses of plants is essential if agricultural development is to be successful and biodiversity is to be conserved. Unfortunately, such basic information is often only partially available for citizens, professional stakeholders, teachers and even scientists. A noticeable consequence, expressed as the *taxonomic gap*, is that identifying plant species is usually impossible for the general public, and often a difficult task for professionals, such as farmers (who have to fight against weed species) or foresters, and even for the botanists themselves. The only way to overcome this problem is to speed up the collection and integration of raw observation data, while simultaneously providing to potential users an easy and efficient access to this botanical knowledge. In this context, content-based visual identification tools are considered as promising solutions to help bridging the taxonomic gap [4].

Figure 1: 6 plants sharing the same common name for *laurel* in French, belonging to distinct species.

Most methods proposed in the past for such automatic identification were actually based on leaf images [5, 1]. Leaves are far from being the only discriminant visual key between species but, due to their shape and size, they have the advantage to be easily observed, captured and described. Although not very new from the computer vision point of view [8], the problem of identifying plants is gaining more and more interest in the multimedia retrieval community, as illustrated by the ImageCLEF plant identification task [4] aimed at evaluating leaf-based identification systems worldwide. Another noticeable fact is the great success of the LeafSnap[1] iphone application, developed by Columbia University. Using uncluttered leaf images and state-of-the-art leaf shapes boundary features [2], it already allows to identify more than 250 northeastern American plant species with good identification performances. On the other side, the leaf is far from being the only useful organ for accurate identification [6]. As an example, the 6 species depicted in Figure

[1] http://leafsnap.com/

MAED'12, November 2, 2012, Nara, Japan.

1 share the same French common name of *laurel* (*"laurier"*) whereas they belong to different taxonomic groups (6 genera, 4 families). Main reasons for that are that these shrubs are often used in hedges and that theirs leaves have more or less the same-sized elliptic shape. Identifying a *laurel* can be very difficult for a novice by just observing leaves, while it is indisputably easier with flowers. Beyond identification performances, use of leaves alone has also some practical and botanical limitations. Leaves are not visible all over the year for a large fraction of plant species. Deciduous species, distributed from temperate to tropical regions, can't be identified by the use of their leaves over different periods of the year. Their leaves are indeed often too young or too much degraded to be exploited efficiently. Moreover, leaves of many species are intrinsically not enough informative or very difficult to capture (needles of pines, grass, huge leaves of banana trees...) This paper presents the first automated visual identification system dealing with multiple organs and views of plants [2]. It therefore allows to querying the system at any period of the year and it benefits from the complementarities of different views of the plant to improve identification performances. Thanks to an interactive and visual query widget, the tagging process of the different organs does not require any expertise in botany.

2. INTERACTIVE IDENTIFICATION

Figure 2 displays a snapshot of the query GUI of our web application. Inspired by computer assisted identification tools based on morphological attributes [7], each considered organ is represented by a graphical icon. When clicking on any of them, a browsing window pops up and allows the user to upload one picture of the selected organ (or to paste one picture url). The main advantage of this graphical widget is that it allows to implicitly tagging the submitted pictures without much effort and without any expertise on the underlying botanical terms. The version of the application presented and experimented in this paper is dedicated to the identification of trees and woody shrubs species of the French flora. To illustrate the reliability of the paradigm, we however mention that the application is currently being deployed on more specific flora with more complex organs and views such as anatomic woods cuts, banana *bract* and *buds*. Our web application GUI can be thus easily tuned to any set of attributes and corresponding icons in order to be adapted to other contexts or end-users.

After providing one or more pictures for one or more organ(s) and view(s), the user can finally launch the search by pressing the *identify* button on the right. Before searching, he also has the possibility to crop each submitted picture by clicking on it and selecting a window of interest. Finally, at any time, the user can refine his search with taxonomic filters (family, genus) and/or precise the taxonomic level of the returned results (family, genus or species).

Figure 4 displays the main result GUI of the application with an example of multi-organ query Figure 3 of a Judas Tree. The system actually returns a list of species ranked by decreasing confidence scores (or a list of families or genera if the taxonomic level has been activated). Each species is illustrated by a thumbnail of the best matched picture in the training set, across all organs. Clicking on one of the species brings a more detailed view where the whole list of retrieved species is still accessible. The user can access to additional pictures for each species, which are composed by the consist of other matching pictures across all organs. If the user clicks on the button *More of this species* he can visualize all pictures of the training set through a specific browsing window. Note that in any view and window of the application, the number of results and the size of the images are interactively adaptable by zoom-in and zoom-out operations. Finally, the user can access to very complete botanical description provided by a specialized website [3] when clicking on the button *More details* for any species.

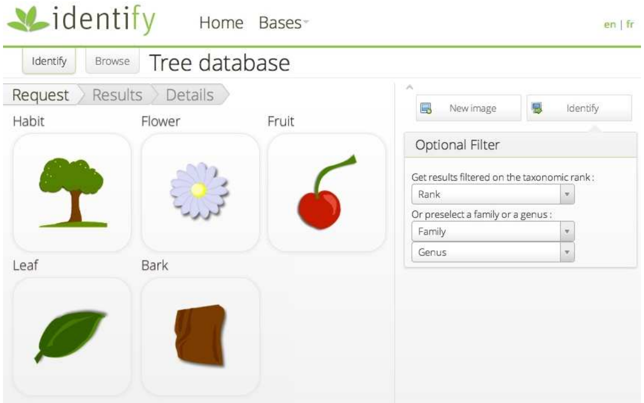

Figure 2: Query User Interface

Figure 3: Exemple of a multi-organ query: 5 pictures of a same Judas tree (*Cercis siliquastrum L.*)

3. COLLABORATIVE TRAINING DATA

The proposed application is built on top of a collaborative application allowing the training data to be continuously enriched thanks to the contributions of amateur and expert volunteers belonging to a botanical social network. The training pictures used in the experiments of this paper were collected during one year, in different regions and with distinct devices without any additional expertise than that of the contributors themselves. To our knowledge, there is no other multi-organ database meeting the needs of content-based identification, i.e. with a large number of species, explicit organ tags and a sufficient number of training pictures. Some initiatives like MorphBank[4] provide good illustrations for many plants but often with only one picture per organ and per species, making any training a very tricky task. On the other side, the ImageNet database [3] contains around thousands of pictures for a lot of several species, but without organ tags and a lot of noise. More generally, the large amount of plant images available on the web suffers from a

[2] `http://identify.plantnet-project.org/en/base/tree`

[3] `http://www.tela-botanica.org/site:botanique`

[4] `http://www.morphbank.net/`

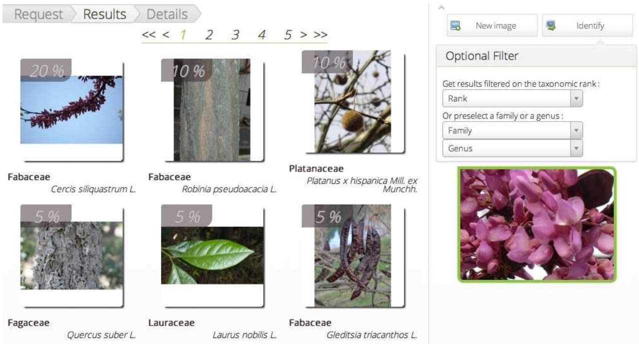

Figure 4: Result User Interface

long tail distribution, i.e. with very few species well represented and many species with very few images.
We believe that citizen sciences initiatives conducted in conjunction with botanists (such as the one launched in the background of this work) are the most promising tools towards speeding-up the integration of appropriate visual data. Potential contributors of the botanical social network have for instance access to some illustrations of the required organ views and acquisition protocol (reducing the noise). Some specific emphasized species are also advertised by monthly newsletters in order to focus the contributors on under-represented species or organs. At the time of writing, 3501 pictures were collected by 83 contributors, through the observations of 2846 trees covering 121 species over 71 districts[5]. The collection contains 229 pictures of *habit*, 1056 of *leaf*, 707 of *flower*, 477 of *fruit* and 1032 of *bark*.

4. VISUAL SEARCH ENGINE

Thanks to the organ-based tagging paradigm, pictures belonging to a given plant view category can be indexed and searched in a separate visual index. This allows reducing confusion between pictures of different parts and therefore increases identification performances. At query stage, the N_Q query pictures belonging to a query plant Q are searched separately in their respective visual index and the top-K most similar images are returned for each of them. Identification is then performed thanks to an instance-based classifier computed across all retrieved pictures (late fusion). Depending on the taxonomic level selected by the user, the voting process is applied either on species (selected by default) or genus or family. The final confidence score of each retrieved taxonomic group is computed as the result of the vote divided by the total number of retrieved pictures.
It is important to notice that each visual search index could be specialized for the targeted organ with specific visual features and similarity metrics. But after achieving a comparative study and in order to keep the whole application as generic as possible, we finally kept exactly the same visual search method for all organs. As suggested by the results of ImageCLEF2011 for leaves [4], it is based on local features and large-scale matching. Interest points are detected with a color Harris detector modified in order to favor points with a central position in the image and to reduce the impact of background features. Each interest point is then described with a SURF local feature and a HSV histogram.

[5] `http://www.tela-botanica.org/widget:cel:carto?tag=plantnet`

Organ	Habit	Flower	Fruit	Leaf	Bark
with tags	**0.124**	**0.189**	**0.088**	**0.105**	**0.16**
without tags	0.091	0.129	0.038	0.088	0.147

Table 1: Single-image top1 classification rates

5. EXPERIMENTS

Performances evaluation We first evaluated the identification performances of our application using a Leave-One-Out procedure, where *one* is not *one picture* but *one set of individual-plants* from the same species, observed by the same photograph at roughly the same place and time. This prevents from any bias between the training and testing data: plants observed by the same photograph, during a short period and in a restricted area are actually more likely to be similar between to each other (same lighting conditions and camera, similar neighboring trees, etc.). Performances are measured by a classical classification rate, i.e. how often the system gives the correct species at the top-1 position in the response list. At the time of writing, the database contains 121 species, so that performances are to be compared to a random rate of 0.008.
First evaluation presented in Table 1 concerns the identification performances obtained when using one single image of one single organ in the query. The contribution of the *part-based tagging* paradigm is evaluated by comparing the identification rate obtained when mixing all images in a single index to the one obtained when using independent indexes for each organ type. The Table shows that using tags improve performances from 9% for *bark* to 132% for *fruit*. The *flower* organ provides the best identification rate, which is not surprising according to botanists, since plant classification is mainly based on flower morphology. Ephemeral organs are actually more likely to be the most discriminant. More surprisingly *bark* tends to give a quite good classification rate. Botanists in tropical forest commonly use bark, since few other features are accessible. However, the general public does not commonly use it, as few structured description have been produced on this complex tissue. The second evaluation presented in Table 2 concerns multi-images and multi-organ queries. It provides the identification performances for different multi-query formulations. The notation $B_i Fl_j Fr_k H_l L_m$ means that the tested queries were composed of i pictures of the *bark*, j pictures of its *flowers*, etc. Classification rates were averaged on all query plants where the combination was actually possible. The two first rows of the table show the contribution of submitting several pictures of the same organ instead of one single picture as in Table 1 (on the best organ, i.e Flower).
The 4th and 5th rows provide the mean performances for the 5 best multi-organ combinations when using only one picture per organ selected at random. The results show that the simple late fusion strategy used so far in the application does not systematically improve the performances over the single image performances. Multi-organ classification rates are still better than when using no tag at all and better than the three worst organs used alone. The *Flower* and the *Bark* organ used alone however give equivalent performances which might be interpreted by their higher occurrence in the collaborative training data. As we will see in the human-centered experiments, these raw classification rates do however not reflect the overall performances

Multi-image	Fl_5	Fl_4	Fl_3	Fl_2	Fl_1
Perf. (top1)	**0.289**	0.263	0.219	0.183	0.188
Perf. (top10)	**0.636**	0.629	0.545	0.537	0.577
Multi-organ	Fl_1H_1	B_1Fr_1	L_1H_1	L_1Fr_1	B_1H_1
Perf. (top1)	**0.159**	0.130	0.129	0.108	0.106
Perf. (top10)	**0.613**	0.443	0.356	0.448	0.315

Table 2: Multi-query top-1 and top-10 performances

Evaluation criterion	**Best**	**Worst**	**Avg**
User Identification rate	1.0	0.64	0.85
Confidence score (/10)	9	5.57	7.49
Number of trials (/5)	1.21	2.21	1.62
Utility (/10)			
- Overall	10	4	7,93
- Crop	10	3	7.9
- Taxonomic filtering	10	4	7
- "More details"	10	3	7.67
- "More of this species"	10	8	8.67
Usability/ergonomics (/10)	9	5	7.6
Global appreciation (/5)	3.7		

Table 3: Human-centered evaluation results

of the interactive application. Browsing and visualization functionalities among the top-K returned species are actually essential for the human validation of the recommended species. So that submitting several organs still favors the disambiguation of several species (as illustrated in Fig. 1) whereas submitting one single organ might degrade the interpretability of the returned species. As a reference, 3rd and 6th row give top-10 classication rates, i.e. how often the correct species is among the 10 first species in the response, supposing that users will probably not browse the list over the 10th position.

Human-centered evaluation Beyond the raw identification performances presented above, we also achieved a human-centered evaluation to assess the utility and the ergonomy of its visualization and interactive functionalities. 10 non-expert users were asked to identify 14 plants randomly selected from two pools of about 200 multi-organ queries that were built outside the application. These photos were actually taken by a group of 24 people (not registered members), who participated to two one-day campaigns that we did organize at two different periods of the year. For that purpose, two botanists have identified more than 150 trees in a given area in order to permit to allow people to take pictures of identified specimens. Users were equipped with their own camera.

Now the evaluation itself worked as follows. For each of the 14 queries, the user can do anything he wants with the application to perform the identification. We limited to 5 the number of times he can click the identify button and to 5 minutes the identification of each of the 14 queries. At the end of each query, the user filled a form with the name of the species he chose and a confidence score from 1 to 10. At the end of the whole session, each user was asked to give a note on several aspects including: utility of the application for identifying plants, utility of optional functionalities (more details, crop, etc.), ergonomics and global appreciation. Results are summarized in Table 3. The average identification rate is 0.85, which is positively higher than the leave-one-out experiments, showing the effectiveness of our interactive application. The best user identified correctly all query plants and the worst one 64% of them. The number of times users press the identify button is on average 1.62 showing that they quickly understood how the application can give the best results. Most functionalities have been considered very useful to complete an accurate identification. Interestingly, the confidence scores show that some users still have some doubts even when they provide the correct identification. Any botanist would confirm that an identification is never 100% sure. Other evaluated criteria clearly show the very good acceptance and usability of the application.

6. CONCLUSIONS & PERSPECTIVES

This paper presented the first application allowing to identify plants from photos of various organs. The experiments did show that it is already a very useful and effective tool for identifying plants although the collected data are limited so far to 121 tree species and 3501 images. As all training data are continuously collected in a fully collaborative way, we believe it has a high potential for scaling up to much more species and therefore helps bridging the *taxonomic gap*. A challenging perspective is therefore to involve more people worldwide by launching new citizen sciences campaigns. Another perspective is to deploy the application on more specific flora addressed to scientists and professional stakeholders. From a technical point of view, the main perspective is to design effective feedback mechanisms allowing non-expert users to enrich the data without adding too much noise.

7. ACKNOWLEDGMENTS

This research has been conducted with the support of the Agropolis Fondation. Great thanks to all users of Tela Botanica social networks who shared their plant observations and who spent time tagging pictures.

8. ADDITIONAL AUTHORS

Daniel Barthelemy (CIRAD BIOS Direction and INRA UMR AMAP F-34398, Montpellier, France, daniel.barthelemy@cirad.fr), Nozha Boujemaa (INRIA, Saclay, France, nozha.boujemaa@inria.fr).

9. REFERENCES

[1] A. R. Backes, D. Casanova, and O. M. Bruno. A complex network-based approach for boundary shape analysis. *Pattern Recognition*, 2009.

[2] P. N. Belhumeur, D. Chen, S. Feiner, D. W. Jacobs, W. J. Kress, H. Ling, I. Lopez, R. Ramamoorthi, S. White, and L. Zhang. Searching the world's herbaria: A system for visual identification of plant species. In *ECCV08*.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR*, 2009.

[4] H. Goëau, P. Bonnet, A. Joly, N. Boujemaa, D. Barthélémy, J.-F. Molino, P. Birnbaum, E. Mouysset, and M. Picard. The ImageCLEF 2011 plant images classification task. In *ImageCLEF*, 2011.

[5] H. Goëau, A. Joly, S. Selmi, P. Bonnet, E. Mouysset, and L. Joyeux. Visual-based plant species identification from crowdsourced data. In *ACM Multimedia*, 2011.

[6] M.-E. Nilsback and A. Zisserman. Automated flower classification over a large number of classes. In *ICVGIP08*.

[7] C. Sarmiento, P. Detienne, C. Heinz, J.-F. Molino, P. Grard, and P. Bonnet. Pl@ntwood : a computer-assisted identification tool for 110 species of Amazon trees based on wood anatomical features. *Iawa Journal*, 2011.

[8] Z. Wang, Z. Chi, and D. Feng. Shape based leaf image retrieval. *IEE VISP*, 2003.