

Basics of Automated Plant Identification

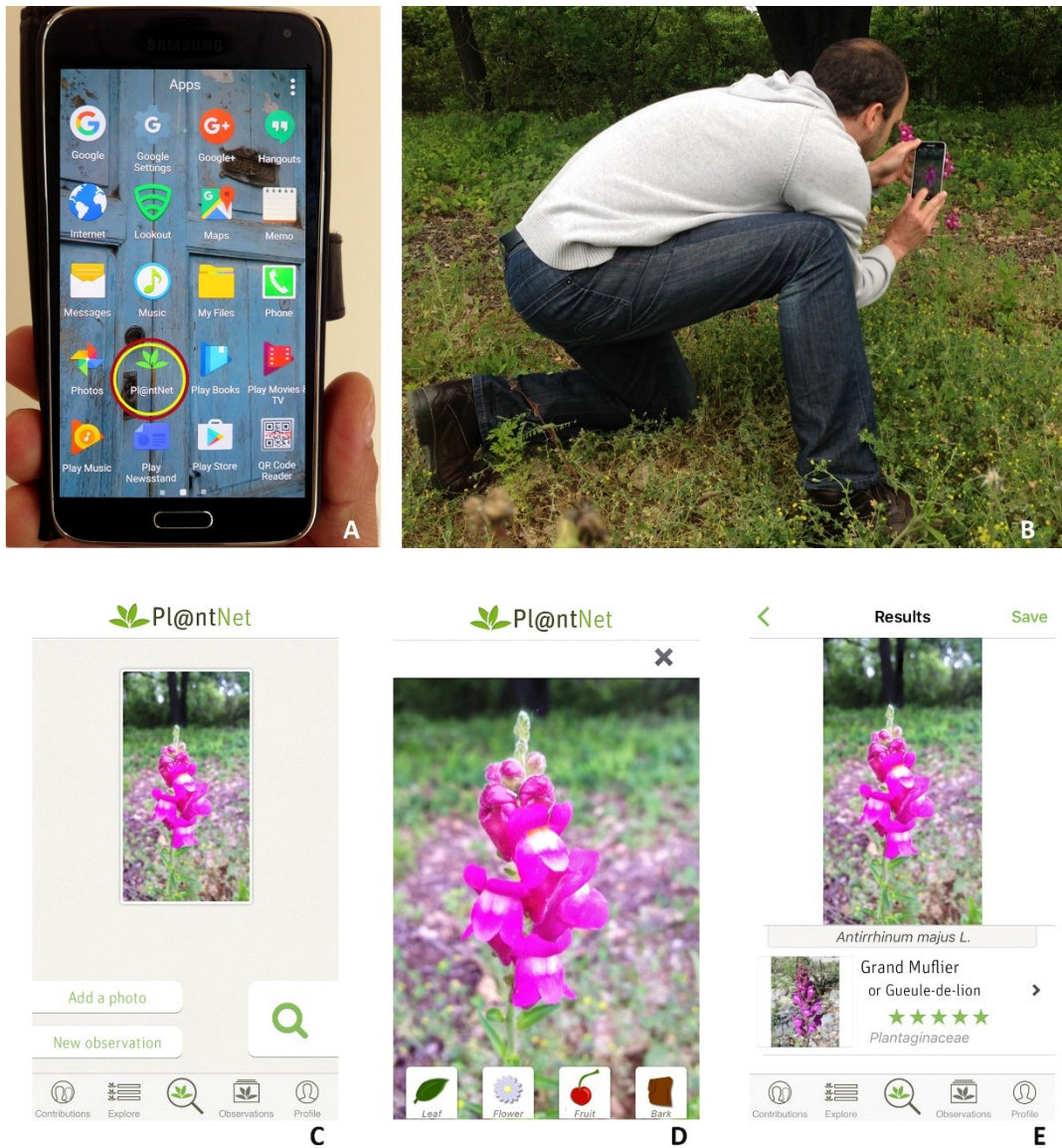


Figure 1. Using an automated plant identification application (Pl@ntNet). A, application displayed on a smartphone; B, plant of interest is photographed by user using smartphone; C, photo of plant as it appears in application; D, organ type depicted in photograph (flower) is manually chosen by user; E, identification results.

BASICS OF AUTOMATED PLANT IDENTIFICATION

Pierre Bonnet¹ and Dawn Frame²

ABSTRACT

Historically, image-based dichotomous plant identification keys precede text-based ones by nearly one hundred years. Having lain in conceptual torpor for over 300 years, the notion of image-based identification has experienced a revival as a result of the development of modern applications which depend upon recent technological advances in electronic hardware (e.g. image sensors, network bandwidth, computer storage capacity) and software (especially image recognition systems and efficient large file browsing). There are essentially two different approaches to automated image-based recognition of plant species: Leafsnap and Pl@ntNet. A brief discussion of the two approaches is here presented. Regardless of the approach, for successful automated plant identification, there are several dataset requirements and these are laid out in the following paper.

Key words: Image-based identification, social network, crowd-sourcing, citizen science, multi-organ, computer vision, mobile application, botany.

INTRODUCTION

The wide disparity between the reality, both qualitatively and quantitatively, of the species making up Earth's biodiversity and our knowledge of it, has been called the "Taxonomic Gap" (DUBOIS, 2010). Until recently, species identification has been carried out by the use of dichotomous keys. Lamarck has long been credited with the invention of the dichotomous key to species, which he presented in his *Flore Française* (1778). However, new evidence suggest that the first dichotomous identification key was proposed by Richard Waller in 1689 (GRIFFING, 2011), and unlike Lamarck's, which was text-based, Waller's was an image-based one, consisting of a series of water-colors of English herbs (GRIFFING, 2011). A great part of the problem of recognizing new species lies in knowing what has already been

¹ UMR AMAP "botAnique et bioinforMatique de l'Architecture des Plantes". TA A-51/PS2, Boulevard de la Lironde, 34398 Montpellier cedex 5, FRANCE. Email : Pierre.bonnet@cirad.fr.

² UMR AMAP "botAnique et bioinforMatique de l'Architecture des Plantes". TA A-51/PS2, Boulevard de la Lironde, 34398 Montpellier cedex 5, FRANCE. Email: framedm@gmail.com.

Basics of Automated Plant Identification

described. For the interested lay person, novice or confirmed taxonomist, text-based dichotomous keys have been the standard means of identifying species, and these are notoriously difficult to navigate even for professional taxonomists, and as such, they represent a major bottleneck to rapid species identification.

Early on, conceived as a labor-saving tool for the identification of common species cluttering up the taxonomist's workbench (GASTON & O'NEILL, 2004), automated species identification has come a long way in just over a decade. Hand in hand with technological improvements in hardware (e.g. image sensors, network bandwidth, computer storage capacity and memory) and software (especially image recognition systems and efficient browsing of large datasets), so too has the size and cost of equipment dramatically decreased leading to the democratization of computers and other portable devices such as telephones equipped with high resolution cameras. Consequently, there has been an explosion of web-scale multimedia data and with it the emergence of innovative processing and applications (CHANG et al., 2012). Moreover, there has been a paradigm shift, such that now automated approaches to species identification are designed not just to recognize common species, but potentially any species (Fig. 1). Today, there are essentially two different approaches to automated image-based recognition of plant species as embodied by the following two systems: Leafsnap (KUMAR et al., 2012) and Pl@ntNet (JOLY et al., 2014a). Typically, automated identification systems, be it plant identification from images taken by hand-held cameras, aerial photography or even hyperspectral sensors, comprise two separate processes, which may be conducted sequentially or in parallel. The first process involves the analysis of known entities such as images of organs of known (properly identified) species. This data is used to generate a "training" set from which differences between species are "learnt", that is to say, discriminated; the second process involves the analysis of unknowns to be automatically identified (GASTON & O'NEILL, 2004).

BUILDING A DATABASE: WHAT TO WATCH FOR

Image-based automated plant identification tools require substantial numbers of images in order to provide several examples of the same visual concept (i.e., a species). Kinds of examples to be included encompass both biological (e.g. intra-specific variation, different growth/developmental stages) and physical conditions (e.g. different lighting, taken from various angles), overall, whatever the project, plant image libraries typically suffer from one or more of the following problems, which often prevent their effective use as training data (JOLY et al., 2014a): -

1. Usually few images per organ per species.
2. A few species having many images, but most having very few.
3. “Noise”, i.e. cluttered or mixed images, errors in the metadata (tags, labels); illicit logging, coupled with weakness in monitoring.
4. High heterogeneity, in terms of acquisition protocols, views or tags.
5. High homogeneity — the result of few people taking images during a limited period of time in a restricted area, using an identical sensor and acquisition protocol.

The first problem is related to the fact that an image library is thought to be “good” when [A] it covers many species, so the interest is in obtaining high numbers of species, often with few photos each or alternatively [B] it covers only a few, usually common, species represented by many images, whereas for the vast majority there are only a few images for illustrative purposes. From a machine learning point of view, it is necessary to have many images of different organs taken under different conditions. Background noise (pt 3) can occur at the level of the image, affecting the machine’s ability to rapidly discern the queried plant in the image, or at the metadata level in terms of incorrect identification, organ name, view label and so on, that is, noise here is equivalent to degree of metadata error. Both high heterogeneity and homogeneity are sources of problems. High heterogeneity of photographing protocols (e.g. indoor, outdoor, studio, flat-bed scans) and variety of views (e.g. whole plants, portions thereof, landscapes or herbarium specimens) are problematic for image recognition and machine learning as is heterogeneous metadata such as the use of different terms for the same entity (e.g. leaf, leaves, foliage). The final weakness, oddly enough, is too much homogeneity. Some datasets have been built especially for computer vision and machine learning and they contain categories having many different (well populated), numerically balanced, homogeneous images without noise, the product of few specially trained people who took photos over a short time period in a restricted area following a stringent acquisition protocol using a single or few sensors. This leads to a dataset lacking diversity, which greatly limits its utility in the real world, but otherwise fulfils most of the requirements of a good dataset for computer vision and machine learning.

LEAFSNAP

Leafsnap was the first plant species mobile application (KUMAR et al., 2012) and is currently available for Apple mobile devices (iPod touch, iPhone and iPad). The dataset covers 185 tree species of the northeastern United States and Canada, and work is on-going to expand this to all tree species of the continental United States (<http://leafsnap.com/about/>). Recently, a United Kingdom version for iPhone has been developed called Leafsnap UK, which allows identification of 156 tree species (<http://www.nhm.ac.uk/take-part/identify-nature/leafsnap-uk-app.html>). Leafsnap is a visual recognition system designed for a single plant organ, the leaf. The requirements of this application are that the user must photograph on a solid light-colored background, a well-flattened leaf of the tree to be identified. The software then classifies the image (is it a leaf?), segments the image (separates background from leaf), extracts features (often involving compensation for curvature), compares resultant features with a labeled database, and then returns the species having the closest matches, totaling about five seconds for the whole process (KUMAR et al., 2012). Additionally, if the Global Positioning System (GPS) of the mobile device is turned on, the application allows a geographic information system (GIS)-based mapping of the study tree. The user can look through the returned results, which are associated with photos of other organs, views of the entire tree and a description, and make a final identification. Numerous specialists and trained volunteers took the photos forming the image database; it is composed of high quality laboratory images of back- and front-lit pressed leaves (23,147) and a lesser number of field images (7,719) mostly taken outdoors on iPhones. It is important that the database be populated with many images because of the high degree of within species (intraspecific) variability in leaf shape and considerable variation in lighting conditions under which the query photos will be taken. Despite its most notable drawbacks of single-organ recognition, coverage of relatively few species and necessity to photograph under fairly stringent conditions, all a reflection of the relatively homogeneous training dataset, Leafsnap has been adopted by over a million users and pedagogic materials using it have been developed to teach adolescents botany in the United States.

PL@NTNET

Pl@ntnet represents an alternative approach to automated plant identification. As mentioned above, homogeneity in training data severely limits wide applicability of an identification tool. In reaction to this, there has been a move towards

collecting crowd-sourced data. This method must be judiciously applied else it can be burdened with too much noise (pt 3 above), e.g., if raw research results of ImageNet were simply filtered and consolidated by a crowd-sourced interactive application, wherein images were validated by only a few users (JOLY et al., 2014a). In an inverse manner to Leafsnap, Pl@ntnet is a multi-organ system that derives its training data from images taken using mobile devices operated by experts, amateurs and novices (crowd-sourcing). These campaigns were and are coordinated through a thematic social network, Tela Botanica, the largest French-language botany network in the world, having over 29 000 registered users living in more than 70 countries. For such a project to be successful, it had to be carefully organized from inception, and a series of work-flows were devised by a team of experts in botany, image recognition and software development.

The Pl@ntNet initiative focused on the development of innovative digital tools, specifically: (i) visual aids for taxonomic identification, (ii) collaborative revision of data quality and (iii) management of large volumes of botanical observations. In 2009, the project team created a small dataset of images of Southern European tree species leaves. This dataset was enriched soon after by complementary images of other organs of these same species in addition to other species (initially the most common tree and weed species), collected during campaigns organized by Tela Botanica and employing network members, in this way, novices, amateurs and experts collected images used for training a content-based identification tool. The forays were guided by experts, either in person or through an illustrated newsletter to members with, for example, seasonal suggestions of what species views to collect. In an interactive, collaborative manner through the website and application, end-users can propose and verify identifications and there exists a weighting system related to botanical expertise so that, for instance, an expert's identification is higher valued than that of a novice. Similar to other crowd-sourced media, photos of species are voted on for quality, often a motivating force for some participants. Identification is guided by experts by means of illustrated booklets and web-based links to references and other identification resources. The growing dataset produced mostly by amateurs is tested each year by ImageCLEF and LifeCLEF campaigns (JOLY et al., 2014b) using different algorithms for plant identification based on real-world data, that is, collected by non-specialists. The dataset allowed the testing and evaluation of plant identifications made not just from leaves, but also from flowers, bark, fruit. The process of continuous data integration by Pl@ntNet allowed the development of a growing computational platform able to manage and benefit from thousands of contributions. This platform was first available on the web (2011) followed by an iPhone application in 2013 and Android in 2014. One of the major

Basics of Automated Plant Identification

innovations of the PI@ntNet platform has been the ability for the end-user to directly revise, in a collaborative manner, all visible data. Thanks to this continuous revision process, the application is able to cover an increasing number of species and has a growing number of images (Fig. 2.). By creating a structured dataset, developing innovative tools for data browsing, and building a community of volunteers, the PI@ntNet initiative made it possible to aggregate a huge volume of botanical observations (over 2 million observations are currently being analyzed) from the user community's identification requests. The created infrastructure has been used by more than a million and a half people, representing a wide range of users, from non-specialists to experienced botanical researchers, in over 150 countries in the world.

Characteristic	Leafsnap	PI@ntNet
Organs	leaves	leaves, flowers, fruits, bark & habit
Setting	homogenous background	natural conditions
Criteria	mono-image, -criterion	multi-image, -criteria
Database type	static	dynamic (daily updates)
Image recognition	segmentation	content-based image retrieval using data mining
Contributors	few, trained specialists	many thousands of lay-photographers
Species number & Flora	185 North-eastern US, 156 United Kingdom	ca. 6000 mostly France
Plant growth form	trees only	any

Table 1. Summary of the differences between Leafsnap and PI@ntNet.

PROOF OF CONCEPT

PI@ntNet and Leafsnap are similar in that they are image-based identification systems available as free mobile applications, which can use relatively low-resolution images to provide a list of probable species in a few seconds. However, these two applications differ in many ways (Table 1) largely related to the acquisition of the

training data. It is clear, whichever the system employed, that the future of automated plant identification lies in eschewing text-based dichotomous keys in favor of image-based applications. As Richard Waller in a letter to John Ray dated 5 April 1688 (DERHAM, 1718) aptly states when providing the rationale for his illustrated key "... my Design in these Tables being only to give an Idea of the Difference of Plants by *Pictures*, (the Representations of Beings) rather than by *Words* (the Representations of Pictures.) ...". Ray dismissed the idea and interestingly enough, his major opus *Historia Plantarum* (1686, 1688), which lacked illustrations, did not achieve the hoped-for success. Of course, designing watercolor illustrations is too time-consuming for rapid and efficient characterizations of numerous species, but with the vulgarization of inexpensive digital image sensors available in cameras and portable telephones, rapid and reliable plant identification is leaving the workbenches of the herbarium scientist passing through the hands of citizens-scientists and landing into the everyday life of ordinary people. We may never be able to completely bridge the taxonomic gap, but possibly with the aid of innovative identification tools in the hands of the many instead of the few, we will progress towards a better understanding of the natural world surrounding us.

ACKNOWLEDGEMENTS

The authors are deeply indebted to Stephen Elliott for the opportunity to present this paper at the Automated Forest Restoration workshop held in Chiang Mai, Thailand (2015).

REFERENCES

- CHANG, E., S. F. CHANG, A. HAUPTMANN, T. HUANG & M. SLANEY, 2012. Web-scale multimedia processing and applications. *Proc. Inst. Electrical & Electronics Engineers* 100: 2580-2583.
DOI: 10.1109/JPROC.2012.2204110.
- DERHAM, W., 1718. *Philosophical letters between the late learned Mr. Ray and several of his ingenious correspondents, natives and foreigners.* W. & J. Innys, London, United Kingdom.
- DUBOIS, A., 2010. Zoological nomenclature in the century of extinctions: priority vs. 'usage'. *Organisms, Divers. Evol.* 10: 259-274. DOI 10.1007/s13127-010-0021-3.
- GASTON, K. J. & M. A. O'NEILL, 2004. Automated species identification: why not? *Philos. Trans. R. Soc. Lond., B.* 359: 655-667. DOI: 10.1098/rstb.2003.1442.

Basics of Automated Plant Identification

- GRIFFING, L., 1923. Who invented the dichotomous key? Richard Waller's watercolours of the herbs of Britain. *Amer. J. Bot.* 98: 1911-1923.
DOI: 10.3732/ajb.1100188.
- JOLY A., H. GOËAU, P. BONNET, V. BAKIĆ, J. BARBE, S. SELMI, I. YAHIAOUI, J. CARRÉ, E. MOUYSET, J.-F. MOLINO, N. BOUJEMAA & D. BARTHÉLÉMY, 2014a. Interactive plant identification based on social image data. *Ecol. Informat.* 23: 22-34.
DOI: 10.1016/j.ecoinf.2013.07.006.
- JOLY A., H. GOËAU, H. GLOTIN, C. SPAMPINATO, P. BONNET, W.-P. VELLINGA, R. PLANQUE, A. RAUBER, R. FISHER & H. MÜLLER, 2014b. LifeCLEF 2014: Multimedia Life Species Identification Challenges. Pp. 229-249 in Kanoulas, E., Lupu, M., Clough, P., Sanderson, M., Hall, M., Hanbury, A. & E. Toms (eds), *Information Access Evaluation. Multilinguality, Multimodality and Interaction*. Berlin: Springer International Publishing. (Lecture Notes in Computer Science; 8685).
- KUMAR, N., P. N. BELHUMEUR, A. BISWAS, D. W. JACOBS, W. J. KRESS, I. LOPEZ & J. V. B. SOARES, 2012. Leafsnap: a computer vision system for automatic plant species identification. *Proceedings of the 12th European Conference on Computer Vision (ECCV)*
- LAMARK, J.-B. P. A. M., 1778. *Flore française, ou descriptions succinctes de toutes les plantes qui croissent naturellement en France* : 3v. 8°. Imprimerie Royale, Paris, France.
- RAY, J., 1868. *Historia plantarum: species hactenus editas aliasque insuper multas noviter inventas & descriptas complectens: In qua agitur primo de plantis in genere, earumque partibus, accidentibus & differentiis; deinde genera omnium summa tum subalterna ad species usque infimas, notis suis certis & characteristicis definita, methodo naturae vestigiis insistente disponuntur*. Vol. 1. Typis Mariae Clark, prostant apud Henricum Faithorne, London, England.
- RAY, J., 1688. *Historia plantarum: species hactenus editas aliasque insuper multas noviter inventas & descriptas complectens; in qua agitur primo de plantis in genere, earumque partibus, accidentibus & differentiis; deinde genera omnium summa tum subalterna ad species usque, infimas, notis suis certis & characteristicis definita, methodo naturae vestigiis insistente disponuntur*. Vol. 2 : 985 – 1944. Faithorne u. a., London, England.

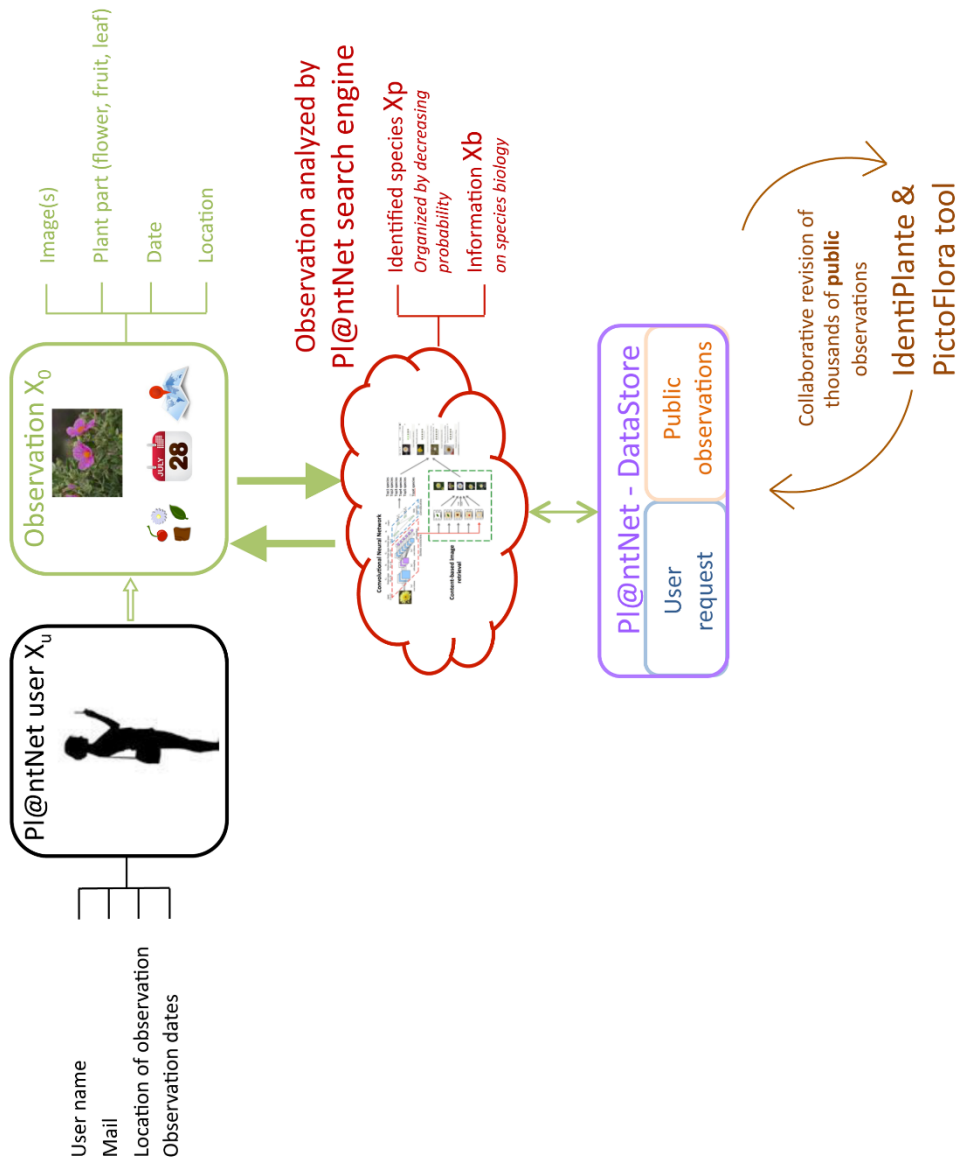


Figure 2. Outline of PI@ntNet interactive workflow.