Using a Deep Convolutional Neural Network for Extracting Morphological Traits from Herbarium Images

Yue Zhu[1], Thibaut Durand[2], Eric Chenin[3], Marc Pignal[4], Patrick Gallinari[2], Régine Vignes-Lebbe[5] (FRANCE)

(1) Ecole Polytechnique (Palaiseau)

(2) Laboratoire informatique de Paris 6 - UMR 7606 (Sorbonne Universités, CNRS, UPMC)

(3) Institut de recherche pour le Développement

(4) Muséum national d'Histoire naturelle

(5) Institut de Systématique, Evolution, Biodiversité - UMR 7505 (Sorbonne Universités, CNRS, MNHN, EPHE, UPMC)

Abstract

Natural history collection data are now accessible through databases and web portals. However, ecological or morphological traits describing specimens are rarely recorded while gathering data. This lack limits queries and analyses. Manual tagging of millions of specimens will be a huge task even with the help of citizen science projects such as "les herbonautes" (http://lesherbonautes.mnhn.fr). On the other hand, deep learning methods that use convolutional neural networks demonstrate their efficiency in various domains, such as computer vision (Krizhevsky et al. 2012, Azizpour et al. 2016), speech recognition (Abdel-Hamid et al. 2014) or face identification (Li et al. 2015, Freytag et al. 2016).

In a proof of concept project, we used a **Convolutional Neural Network (CNN)** in order to recognize automatically 4 morphological traits of leaves on images of the fully digitized collection of the Paris herbarium in the Muséum National d'Histoire Naturelle (MNHN, Paris, France).

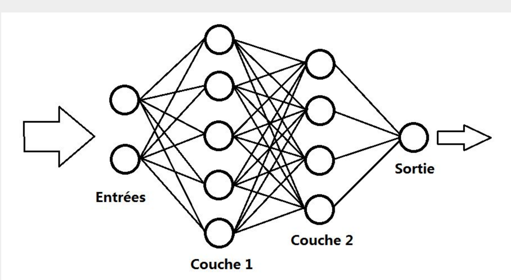
A second method visualizes the area of each image that was detected by the CNN as the most important for morphological character recognition (Durand et al. 2017). This method provides an explanatory view of the automatic recognition process.

Recolnat (https ://explore.recolnat.org/)

Joint project for the French naturalistic collections that is spread across a multitude of institutions of various sizes and status. It provides a virtual desk for consulting and working on all kinds of specimen images with a set of tools (annotation, measures …)

What is a neural network (NN) ?

An artificial neural network is a **supervised learning** system inspired from brain structure. A NN architecture includes the number of layers, the number of nodes for each layer, the activation function, and a threshold to classify the output signal. A back propagation modifies the weight of nodes to minimize the error. Parameters of the NN are optimized during the **training process**. The **validation step** determines the threshold and the number of cycles or epochs.

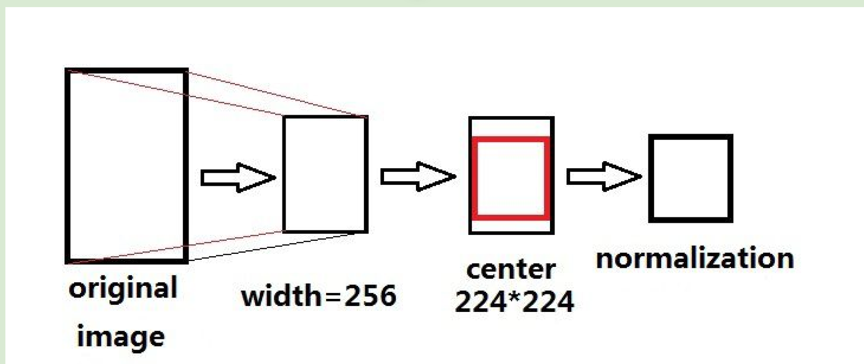


Selection of images and traits

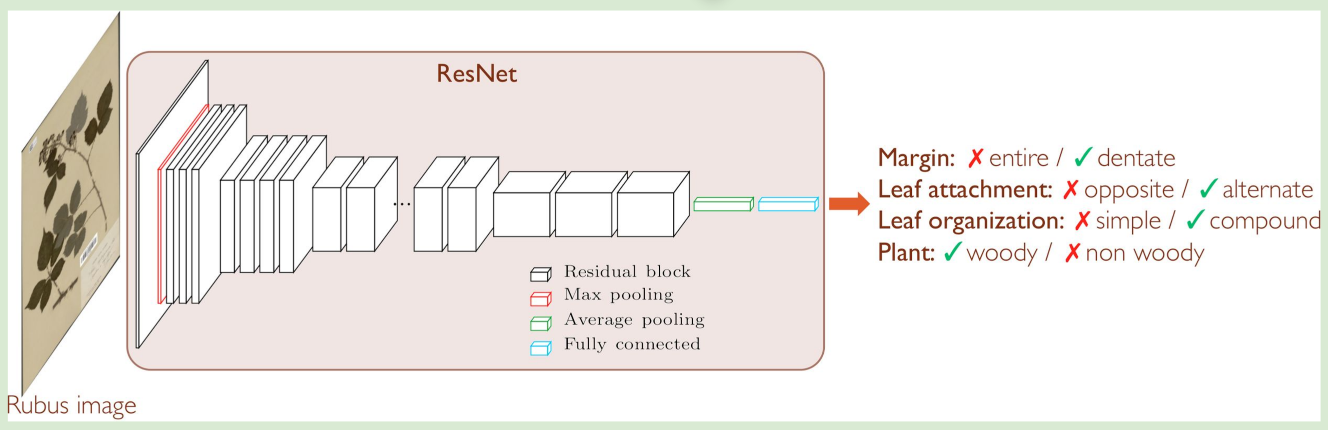
Taxon	Edge	Phyllotoxie	Leaf type	Wooden	Number
Convolvulaceae	Smooth	Alternate	simple	No	27835
Monimiaceae	Smooth	Opposite	simple	Yes	4227
Amborella	Smooth	Alternate	simple	Yes	121
Castanea	Toothed	Alternate	simple	Yes	591
Desmodium	Smooth	Alternate	simple	No	10890
Eugenia	Smooth	Opposite	Composite	Yes	12084
Laurus	Smooth	Opposite	simple	Yes	591
Litsea	Smooth	Alternate	simple	Yes	4437
Magnolia	Smooth	Alternate	simple	Yes	1630
Rubus	Toothed	Alternate	Composite	Yes	39573
Ulmus	Toothed	Alternate	simple	Yes	1795
Total	62015 S	87072 A	53511 S	65249 Y	103974
	41959 T	16902 O	50463 C	38725 N	

103,000 herbarium images from 11 taxa
4 binary characters:
- margin (entire / dentate)
- phyllotaxy (opposite / alternate)
- leaf type (simple / compound)
- plant type (woody / non-woody)

Pre-processing



Processing: Convolutional neural network



- We use **Resnet**, a NN pre-trained from ImageNet
- Data set : 70% for training / 10% for validating the hyper-parameters of the model / 20% for testing
- 9 epochs

Results

real / predict	entire	dentate	
entire	11936	425	**96.56%**
dentate	47	8344	**99.44%**
real / predict	**alternate**	**opposite**	
alternate	14567	2806	**83.85%**
opposite	23	3356	**99.32%**
real / predict	**simple**	**compound**	
simple	10057	485	**95.40%**
compound	381	9829	**96.27%**
real / predict	**not woody**	**woody**	
not woody	7511	234	**96.98%**
woody	563	12444	**95.67%**

TEST 1
Test and training are on the same taxa

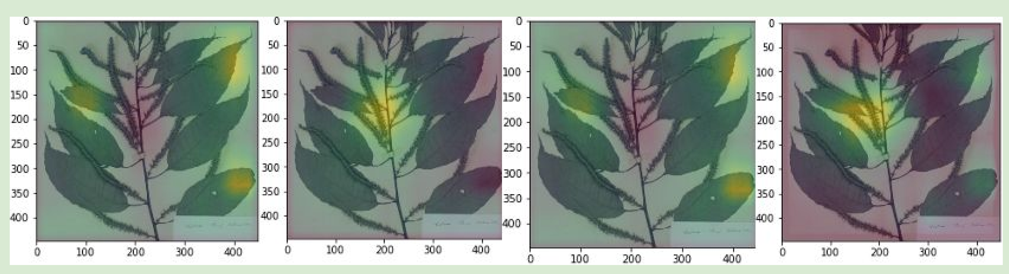
real / predict	entire	dentate	
entire	19162	2613	**88.00%**
dentate	954	1432	**60.02%**
real / predict	**alternate**	**opposite**	
alternate	2336	17007	**12.08%**
opposite	98	4720	**97.97%**
real / predict	**simple**	**compound**	
simple	10884	1796	**85.84%**
compound	9083	2398	**20.89%**
real / predict	**not woody**	**woody**	
not woody	7039	3851	**64.64%**
woody	5466	7805	**58.81%**

TEST 2
Test and training are on different taxa

First test:
- Training and test sets are homogeneous
- >95% good identification

Second test:
- Training and test sets are disjoint (7 taxa for testing, only 4 other taxa for training)
- Correct identification (58% to 98%) except for alternate phyllotaxy and compound leaves

Post-processing - Visualization (Weldon model)



Margin (**entire** / dentate) | phyllotaxy (**opposite** / alternate) | Leaf (**simple** / compound) | Plant type (**woody** / non woody)

A function selects areas where the factors are more numerous
Representative of a class: yellow spots indicate entire / opposite / simple / woody.

What is a Convolutional neural network (CNN) ?

A CNN is a kind of Neural network that is more efficient in dealing with data with local dependencies (such as images). The pixel image is filtered by a matrix (e.g 3*3) in order to convert it into a new matrix by sliding the filter over the image (convolutional layer). Several convolutional layers are used. The CNN learns the values of the filter during the training process.

WEakly supervised Learning of Deep cOnvolutional neural Networks (WELDON)

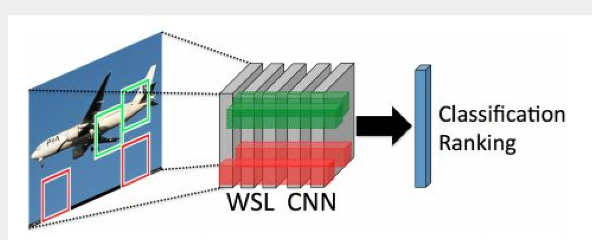


Figure 1. The WELDON model is a deep CNN trained in a weakly supervised manner. To perform image prediction, *e.g.* classification or ranking, WELDON automatically selects multiple positive (green) + negative (red) evidences on several regions in the image.

Discussion and perspectives

The first results are encouraging with over 80% success on the test set. In a second step, we test if the neural network does not overfit the training examples, and if it can be generalized to new taxa. If we restrict the training set to a small number of taxa (4 taxa containing 76% of images), the success rate on the 7 other taxa (unseen during training) decreases drastically. A good sample of the taxonomic diversity of plants appears crucial to train the neural network. A set unbalanced for the different values also has an impact on the results. This could explain the very bad result in the second test for alternate leaves.

The perspective is to provide automatic tagging of herbarium images and verbal description of images. In a first step, we suggest using the virtual desktop (Collaborator see https://lab.recolnat.org/) to facilitate tagging with limited terminology. The citizen science project Herbonaute is already used to edit the label of herbarium images. We began to mobilize the herbonaute community to validate automatic tagging and increase images for training.

Numerous recent publications on Plants and neural network show the promising interest of these methods (Sue Han Lee et al., 2017) (Barré et al., 2017) (Alfonso et al. 2017) (Heredia, 2017) (Carranza-Rojas et al., 2017).

- Abdel-Hamid O, Mohamed A, Jiang H, Deng L, Penn G, Yu D (2014) Convolutional Neural Networks for Speech Recognition. IEEE/ACM Transactions on Audio, Speech, and Language Processing 22 (10): 1533-1545.
- Azizpour H, Razavian AS, Sullivan J, Maki A, Carlsson S (2016) Factors of Transferability for a Generic ConvNet Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 38 (9): 1790-1802.
- Thibaut Durand, Nicolas Thome, and Matthieu Cord. WELDON : Weakly Supervised Learning of Deep Convolutional Neural Networks. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- Durand T, Mordan T, Thome N, Cord M (2017) WILDCAT: Weakly Supervised Learning of Deep ConvNets for Image Classification, Pointwise Localization and Segmentation. 29th IEEE Conference on Computer Vision and Pattern Recognition.
- Freytag A, Rodner E, Simon M, Loos A, Kühl HS, Denzler J (2016) Chimpanzee faces in the wild: Log-Euclidean CNNs for predicting identities and attributes of primates. In: Rosenhahn B, Bjoern A (Eds) Pattern recognition 38th German conference. Springer International Publishing AG, Hannover, Germany
- Krizhevsky A, Sutskever I, Hinton G (2012) ImageNet Classification with Deep Convolutional Neural Networks. NIPS 2012.
- Li H, Lin Z, Shen X, Brandt J, Hua G (2015) IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. https://doi.org/10.1109/cvpr.2015.7299170
- Sue Han Lee, Chee Seng Chan, Simon Joseph Mayo, Paolo Remagnino, How deep learning extracts and learns leaf features for plant classification, Pattern Recognition, Volume 71, November 2017, Pages 1-13, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2017.05.015.
- Pierre Barré, Ben C. Stöver, Kai F. Müller, Volker Steinhage, LeafNet: A computer vision system for automatic plant species identification, In Ecological Informatics, Volume 40, 2017, Pages 50-56, ISSN 1574-9541, https://doi.org/10.1016/j.ecoinf.2017.05.005.
- Carlos Affonso, André Luis Debiaso Rossi, Fábio Henrique Antunes Vieira, André Carlos Ponce de Leon Ferreira de Carvalho, Deep learning for biological image classification, In Expert Systems with Applications, Volume 85, 2017, Pages 114-122, ISSN 0957-4174
- Heredia, Ignacio. (2017). Large-Scale Plant Classification with Deep Neural Networks. 259-262. 10.1145/3075564.3075590.
- Jose Carranza-Rojas, Herve Goeau, Pierre Bonnet, Erick Mata-Montero, Alexis Joly, (2017) Going deeper in the automated identification of Herbarium specimens. BMC Evolutionary Biology, 17:181