InterFolia: an embedded, educational application to identify plant species in the wild

Maxime Langlade, Debaleena Misra, Carlos Crispim-Junior, Laure Tougne
Univ Lyon, Lyon 2, LIRIS F-69676
Lyon, France
{carlos.crispim-junior, laure.tougne}@liris.cnrs.fr

Abstract

This paper presents InterFolia, an iOS application dedicated to help the general public identify the plant species around them. The application can recognize up to 214 species of trees and shrubs found in France and western Europe. Its classification module contains three convolutional neural network models that run locally and are each associated with a plant organ (flower/fruit, leaf, or bark). A fusion module is integrated into the application to combine all available information to help improve the overall accuracy of the classification task. The main novelty of InterFolia is its use of embedded models that allow users to explore nature beyond network-covered zones.

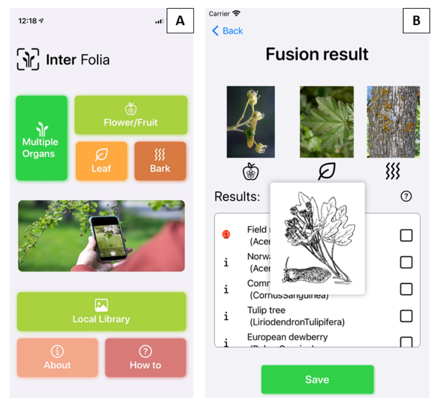

Figure 1. InterFolia application: A) homescreen, B) result screen

1. Introduction

Many mobile applications are devoted to identify plant species such as Pl@ntNet [5], Google Lens [1], PictureThis [3], LeafSnap [2], and PlantSnap [4]. However, all these applications require an Internet connection, limiting the range of nature exploration of their users. This paper presents InterFolia (freely avalaible on App Store[1]), a mobile application devoted to make botany accessible for everyone using deep neural networks and without the requirement of a network connection. To address this challenge, InterFolia focuses on a classification module based on embedded convolutional neural networks (CNN).

The main challenges associated with embedded deep learning are model inference time and memory requirements. The user must get an answer promptly, and the memory footprint should respect the device's capabilities.

2. InterFolia application

InterFolia has been designed following an user-in-the-loop approach. To carry out a plant classification task (Fig.1A), the user can use one or multiple organ images. For single-organ classification, he selects a plant organ and provides a picture (Fig.2A, single-organ path). The application loads the corresponding CNN model and infers the most likely species. The application also proposes a step-by-step tutorial to carry out a multi-organ classification (Fig.2A, multi-organ path). In this case, the user takes (or selects) a picture of each organ available (Fig.2A, multi-organ path). The user may skip any organ for which he cannot have an image. The fusion module is launched at the end of the tutorial to compute the final result based on the information of the organ-specific CNNs (Fig.2B). In all classification scenarios, the application predicts the top-5 most likely species (Fig. 1B) given the available information (1, 2 or 3 organs).

The plant classification step is based on three SqueezeNet models [7]. All models are pretrained on ImageNet dataset [6] and then finetuned for the target plant organ using the InterFolia training dataset.

The fusion module considers two factors to combine multi-organ information: the average performance of an organ model on the plant classification task and the confidence scores of models. The former factor seeks to compensate for the substantial gap of accuracy between organ models,

[1]https://apps.apple.com/tn/app/InterFolia/id1570104475

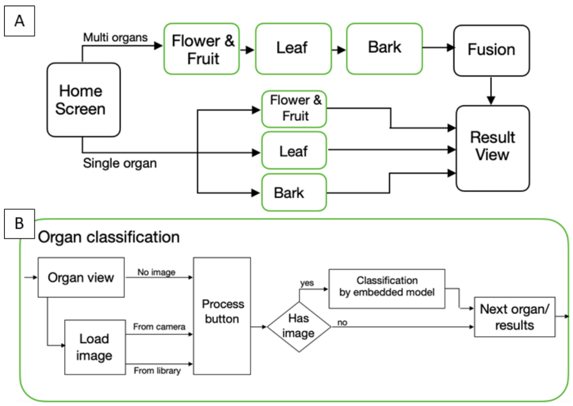

Figure 2. Classification process: A) overview of the system modules, B) typical single-organ classification module.

while the latter integrates the confidence of a model in its classes prediction. The average performance of each organ model is estimated experimentally on the validation set.

3. Experiments

The botanists of the InterFolia project have collected and labeled 9017 images of plant species present in France and western Europe (Table 1). One of the challenges of the dataset is that we cannot hypothesize that an image of each organ of a plant species is available every time. This dataset is used to train the CNN models embedded in the application. To evaluate the accuracy of the application, we created a test set composed of images collected from the Internet of 28 plant species (2 images per organ of each species).

Table 1. InterFolia training dataset - image distribution per organ

	Organ		
	Flower/Fruit	Leaf	Bark
Species	184	178	148
Images	5245	2536	1236

We carried out two evaluations to measure the performance of the InterFolia application in the different working settings. Firstly, we evaluated the individual performance of organ-specific CNN models. Secondly, we measured the performance of the fusion module at combining the results of two and three organ-specific CNN models, respectively.

4. Results and Discussion

To evaluate the performance of InterFolia on **single organ classification** (Table 2), we used the 56 images available of each organ and measured the performance of the corresponding CNN models. An image (or example) is correctly classified if the correct class appears among the five most probable species.

Results show that the "flower/fruit" model attains the highest performance. The bark model ranks lowest, probably because even human experts find it difficult to distinguish certain species using bark information alone.

Table 2. Top-5 accuracy on single-organ species recognition

	Organ		
	Flower/Fruit	Leaf	Bark
Accuracy	43%	32%	23%

The multi-organ classification (Table 3) consists in combining the performance of individual organ models to improve the overall classification results. An example here is composed of a pair (or triplet) of images of different organs of a plant species. For this evaluation, we created two examples per species by randomly combining the organ images of each species in the test set.

Table 3. Top-5 accuracy for multi-organ species recognition

	Organs			
	Leaf Bark	Leaf Flower/Fruit	Bark Flower/Fruit	All
Accuracy	33%	52%	43%	52%

Results indicate that the combination of "leaf" and "flower/fruit" information achieves the highest performance in the two-organ species classification task. The combination of "bark" with "leaf" information slightly improves over the individual performance of the "leaf" model. No improvement is observed by combining information of "flower/fruit" with "bark".

Finally, the use of all three organs' information achieves the same performance as the best combination of two plant organs. This observation suggests no additional information is obtained by integrating "bark" classification into the fusion of "flower/fruit" with "leaf". This behavior is also observed when combining "flower/fruit" information with "bark". These results indicate the fusion module can efficiently combined models with different levels of confidence.

5. Conclusions and Future Work

This paper has presented the InterFolia application, an academic effort to support non-expert users to recognize and learn more about the plant species in their surroundings. The application is based on three deep neural networks that run locally to enable users to keep exploring their vegetation environment beyond the network coverage area. Future work will focus on offering to InterFolia users the possibility of sharing with the project the pictures and the results of their classification tasks, once they are back on a network-covered zone. This functionality will allow us to extend the number of species covered by the application and possibly improve the application performance. Finally, we will also explore other modalities to refine the classification performance (*e.g.*, the date of the picture or the GPS localization).

References

[1] Google lens. https://play.google.com/store/apps/details?id=com.google.ar.lens&hl=fr&gl=US. Accessed: 2021-06-16.

[2] Leafsnap. https://plantidentifier.info. Accessed: 2021-06-16.

[3] Picturethis. https://play.google.com/store/apps/details?id=cn.danatech.xingseus&hl=fr&gl=US. Accessed: 2021-06-17.

[4] Plantsnap. https://play.google.com/store/apps/details?id=com.fws.plantsnap2&hl=fr&gl=US. Accessed: 2021-06-16.

[5] Pl@ntnet. https://plantnet.org. Accessed: 2021-06-16.

[6] J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[7] Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and $<$0.5mb model size. *arXiv:1602.07360*, 2016.