Expanding CNN-Based Plant Phenotyping Systems to Larger Environments

Chapter · January 2020

DOI: 10.1007/978-3-030-65414-6_29

3 authors, including:

Jonas Krause
University of Hawaiʻi at Mānoa

Lipyeow Lim
University of Hawai'i System

Expanding CNN-based Plant Phenotyping Systems to Larger Environments

Jonas Krause, Kyungim Baek, Lipyeow Lim

Dept. of Information and Computer Sciences, University of Hawai'i at Mānoa
1680 East-West Rd, Honolulu, HI, USA
krausej,kyungim,lipyeow@hawaii.edu

Abstract. Plant phenotyping systems strive to maintain high categorization accuracy when expanding their scopes to larger environments. In this paper, we discuss problems associated with expanding the plant categorization scope. These problems are particularly complicated due to the increase in the number of species and the inter-species similarity. In our approach, we modify previously trained Convolutional Neural Networks (CNNs) and integrate domain-specific knowledge in the fine-tuning process of these models to maintain high accuracy while expanding the scope. This process is the key idea behind our CNN-based expanding approach resulting in plant-expert models. Experiments described in this paper compare the accuracy of an expanded phenotyping system using different plant-related datasets during the training of the CNN categorization models. Although it takes much longer to train these models, our approach achieves better performance compared to models trained without the integration of domain-specific knowledge, especially when the number of species increases significantly.

Keywords: Plant Phenotyping, Convolutional Neural Networks, Integration of Domain-Specific Knowledge.

1 Introduction

The categorization of species and plant phenotyping are challenging problems relevant to both disciplines of Botany and Computer Science. Classifying plant images at the species level, considering specific characteristics of their phenotype, is called fine-grained categorization. Despite the availability of various applications, categorizing plants in an environment with a large number of species remains an unsolved problem. Furthermore, an automated system capable of addressing the complexity of this problem and handle larger environments has important implications, not only in preserving ecosystem biodiversity but also in numerous agricultural activities.

In this paper, we extend CNN-based phenotyping systems ([5], [6] and [7]) and present a new approach to maintain their high accuracy when applied to a larger environment. The proposed approach implements a novel scale-up process that adapts the CNNs so that it can handle environments with a broader range of plant species. In this process, we replace the top classification layers to accommodate a more significant number of plant species. However, due to the lack of training data, pre-trained weights have to be used to achieve satisfactory performance. Recently, Cui *et al.* [3], Xiangxi *et al.* [11], and Ngiam *et al.* [12] addressed this problem by introducing domain-specific models, fine-tuning their CNNs to other fine-grained categorization problems. Following their ideas, we implement the integration of domain-specific knowledge for plant phenotyping problems. Also called targeted fine-tuning, the pre-training process generates plant-related knowledge from multiple datasets to create expert CNN models. These fine-tuned CNNs take much longer to train but better performance is achieved when the number of species is significantly increased.

The contributions of our research are the process to expand the scope (i.e., the number of species to be categorized) of CNN-based phenotyping systems, the publication of a new plant species dataset, and the creation of plant-expert CNNs used to fine-tune our categorization models. More specifically, this paper describes modifications made to successfully deploy a plant categorization system (previously designed to categorize 100 species) in a larger environment, expanding its scope to 300 species without significant loss of performance.

We organized this paper as follows: Section 2 presents the related work by describing how most of the plant categorizing systems operate. In Section 3, we describe the adaptations made to expand the scope of the previously trained CNN models. Section 4 details the integration of domain-specific knowledge, a systematic pre-training process developed to integrate previously learned datasets. And Sections 5 and 6 present the experimental results and observations showing the prediction accuracy of our expanded approach and compares them with other commonly used training methods. Finally, Section 7 provides a conclusion and describes future work.

2 Related Work

Nowadays, applications created for categorizing plants implement different CNN models. Most of them are well-known CNN models adapted to work with plant images. However, few of them present new approaches designed to address specific aspects of plant phenotyping, such as the challenges of expanding the categorization scope.

Implementing a simple CNN model, Barré *et al.* [1] present the application called *LeafNet*. They used three datasets (*Flavia*, *Foliage*, and *LeafSnap*) to train and test their CNN-based plant identification system. By comparing CNN models with hand-designed feature methods such as the *LeafSnap* [9] application, Barré *et al.* show that learning features by using a CNN (even with a simple architecture) provides a better representation of leaf images and consequently better discrimination. However, their application does not inherit plant-related knowledge from other datasets, which limits the learning process to the extraction of discriminative features out of only one plant dataset at the time.

Adapting the CNN model called *AlexNet* [8] and using a dataset of 44 plant species, Lee *et al.* [10] also focus on the classification of preprocessed leaf images. They detect leaves in the images by using HSV (Hue, Saturation, and Lightness) color space information to extract the foreground pixels. For the initial experiments, they pre-trained a CNN using the ImageNet dataset and fine-tuned it using the segmented samples. Initial results are not as expected, so they decided to create a deconvolutional structure to visually verify what features the CNN has learned. During this process, Lee *et al.* noticed that the trained model is focusing almost exclusively on the contour and shape of leaves. Based on the low performance, they conclude that leaf shape is not a good choice to identify plants, which is not necessarily true. Morphological features of leaves have been heavily and successfully used for plant species categorization. In this case, non-satisfactory results may be a consequence of a poor pre-training process of the classification model, making the CNN focus almost exclusively on the counter of the leaves. Integrating domain-specific knowledge may assist with this problem.

Sun *et al.* [15] implement plant categorization models customized to 100 plant species. They use images collected from the Beijing Forestry University campus, available online in the BJFU100 dataset[1]. These images present a variety of backgrounds, different illumination conditions, shadows, and it is not always possible to identify a leaf of the plant. Knowing these challenges, Sun *et al.* implement a modified version of the Residual Network (*ResNet*) [4] to classify these images. In their model, a pre-trained *ResNet* works as a bottleneck structure between an initial convolutional block and the last layers of the network. Like this, they adapted the *ResNet* architecture to their needs, customizing this successful CNN model and fine-tuning it with their dataset. Despite implementing a successful customized CNN model, Sun *et al.* limit their work to 100 plant species. And its scope expansion may be harder to perform due to the customized model.

Using the BJFU100 dataset, the work of Krause *et al.* [5] explores multi-scale methods to improve the categorization process of plant species. They present better results when compared with the work of Sun *et al.* by implementing a CNN-based system called *WTPlant*. This system implements a scene parsing method to locate different plant organs and delimit the most representative areas in the images for the categorization of plant species. Additionally, Krause *et al.* present experiments using another dataset with 100 plant species called UHManoa100[2]. As a result, the plant phenotyping system designed by Krause *et al.* has a limited categorization scope. Because of that, our approach focuses on expanding CNN-based systems like the *WTPlant* to deploy it in larger environments. Furthermore, the proposed approach has to face the challenge of maintaining the high accuracy results that other systems reported when trained over small environments (with 100 species or less).

[1] https://pan.baidu.com/s/1jILsypS
[2] https://github.com/jonaskrause/UHManoa100

3 Expanding the Plant Categorization Scope

The accurate categorization of 100 plant species from multiple datasets suggests the work of Krause *et al.* [5] as an effective method for classifying natural images of plants. However, expanding its scope brings new problems: The first problem is to correctly identify which species exist in the target environment and collect representative images of the listed plants. This process requires the assistance of a botany specialist to define the number of species and annotate the training images. With the correctly annotated species, we collect images of plants in the wild to create a new dataset and define the expanded scope. Considering that this new dataset represents the flora biodiversity of a specific region of the globe, classification models trained over these images compose a plant categorization system with that particular scope.

The second problem is to include new species in the plant categorization scope while maintaining the knowledge previously learned and, consequently, minimizing the loss in performance. A simple solution to this problem would be training new CNN models from scratch using the target dataset. However, experiments performed by Krause *et al.* [5] using the ImageNet [13] pre-trained weights showed how valuable the knowledge integration is for the fine-tuning of these plant categorization models. Therefore, a solution that expands the plant categorization scope for more than 100 species must take into account the pre-training process of the classification models. To integrate previously learned knowledge, we implement a modification of the classification models by replacing the top layers of the CNNs with new extended ones. These new layers accommodate a more significant number of plant species but do not guarantee a high categorization accuracy. To address it, we create expert models by training the modified CNNs over domain-specific datasets and use their pre-trained weights to fine-tune the classification models over the new dataset. The pre-trained weights of these plant-expert CNN models are available online[3] and can be used by other researches to fine-tune their models. In this way, knowledge extracted from CNNs pre-trained over plant-related datasets assist in the fine-tuning of the models over new target datasets. And we expand the plant categorization scope by adapting the classification models to inherit powerful discriminative features previously learned during the training over multiple datasets.

Experiments performed using the proposed solution compare the accuracy of the CNN classification models when pre-trained with different datasets like the ImageNet, the UHManoa100 published by Krause *et al.*, and the iNat682 (a plant dataset from *iNaturalist*[4] with 682 species). We use them to pre-train the models before the fine-tuning process over the new target dataset. Although it takes much longer to train these CNNs, the resulting models with integrated domain-specific knowledge categorize plants more accurately throughout all experiments. For much larger scopes that encompass more than one environment (e.g., different regions of the globe, continents, or countries), multiple systems

[3] https://github.com/jonaskrause/Plant_Flower-Expert_CNN_Models

[4] https://www.inaturalist.org/

can operate in parallel using other guidance methods (such as geolocation of the testing image) to indicate which version to use. In this way, we can deploy multiple CNN categorization models to cover an even larger environment and categorize the entire flora of that specific region.

3.1 Creating a Dataset of Broader Range of Species

The first step to increasing the number of species analyzed by CNN-based phenotyping systems is to gather a representative dataset of the listed plants. So we invited a botanist specialist to perform a sanity check on all of our images and ensure that each one of them contains visible traits of the selected species. We also eliminate the incorrectly labeled images as well as the ones with poor quality and low-resolution (smaller than 400x400 pixels). This process is necessary due to the lack of datasets with annotated species available for the experiments conducted in this study.

Following this initial process, we organize the new dataset as a collection of 300 plant species with 50 natural images per species, totalizing 15,000 correctly annotated images. This new dataset, called UHManoa300[5], comprises diverse images with different sizes from 400x400 to 6000x4000 pixels, plants appearing in varying angles, scales, and stages of life. Consequently, it becomes more difficult to categorize this dataset as the appearance of plants changes considerably within the same species. For example, the 50 images of the species *Acacia koa* in Figure 1 illustrate how diverse the within-species plant appearance can be. In addition, similar to the UHManoa100, different plants may appear in the background or even in front of the dominant plant. The annotation of the plant in the images indicates the dominant species (*Acacia koa*) covering the most substantial areas of the images. Also, as shown in the Figure, images in this new dataset contain plants at various scales ranging from leaves and flowers to the entire tree.

3.2 Modifying CNN Models to Accommodate Expanded Scope

After constructing a new dataset containing the plants in the target environment, we adapt the classification models to work with a larger number of species. In this process, we remove the top classification layers of the CNNs initially trained to categorize 100 plant species, saving the weights of each CNN model without the top layers. For each model, pre-trained weights create a basic knowledge of what the model learned in previous training processes (also called base models). A new and larger classification layer added at the top of a base model creates a new CNN with similar architecture but adapted to work with an extended scope. Figure 2 shows this process by expanding the plant categorization scope from 100 to 300 species. Thus, we can load knowledge learned from previous experiments into the same models but with a larger classification layer at the top. Retraining the modified models over the new images, we fine-tune the CNNs to learn discriminative features between a more significant number of species using the pre-trained weights as a starting point for this process.

[5] https://github.com/jonaskrause/UHManoa300

Fig. 1. All 50 images of the *Acacia koa* in the UHManoa300 dataset.

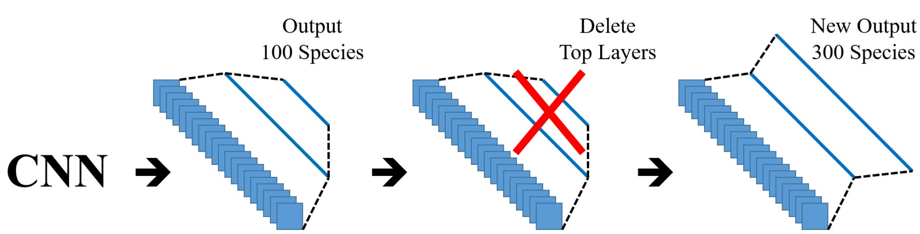

Fig. 2. Process of replacing the top classification layers of CNN models to expand the plant categorization scope.

The number of new species to be integrated into the scope determines the size of the new classification layer at the top of the model. In this paper, we expand the plant categorization scope from 100 to 300 species. For that, we exclude the two dense layers at the top of the CNN models and replace them with new ones that accommodate the expanded scope. The first layer has the same size as the previously excluded one, but the second layer (the last one) is customized to the exact number of classes in the new dataset (300 plant species). These two top layers work together and are responsible for producing the output predictions of the model. Thereby, modified CNN models have the same architecture as the previously trained ones but are loaded with pre-trained weights and ready to work with more classes (plant species).

The fine-tuning of modified models over the target dataset updates the parameter values for the entire CNN based on the pre-trained weights. Consequently, well-trained base models lead to a better fine-tuning process of CNNs over the target dataset. Implementing this adaptation of the classification models, we present a solution to expand the scope of CNN-based systems using multiple pre-trained models. The integration of knowledge from the continuous pre-training processes of the CNNs suggests the creation of domain-specific models. These classification models require an intensive computational effort for training but yield more accurate results.

4 Integration of Domain-Specific Knowledge

Due to the lack of training data for most of the fine-grained categorization problems, ImageNet pre-trained weights are frequently used to integrate knowledge during the training process of CNN models. These weights comprise a base model trained over a million images, and this knowledge is useful for most of the visual classification problems. As previously described, Cui *et al.* [3], Xiangxi *et al.* [11], and Ngiam *et al.* [12] recently introduced domain-specific models for fine-tuning their CNNs to different fine-grained categorization problems. Exploring these approaches, we expand the CNN-based plant categorization systems by adapting the classification models and searching for the best pre-trained weights to fine-tune the models over a new dataset with 300 plant species. Experiments using this new target dataset (UHManoa300) compare the performance of the plant categorization system with CNN models trained from scratch, with the ImageNet pre-trained weights, and with the knowledge integration using two different plant-related datasets (UHManoa100 and iNat682) with the ImageNet pre-trained weights as a starting point.

4.1 Training with Random Initialization of CNN Parameters

Initial experiments start by extracting samples from the training images of the UHManoa300 dataset and randomly dividing them into 80% for training and 20% for validation. We begin by training three CNN models (*Inception-v3* [17], *Inc-ResNet-v2* [16], and *Xception* [2]) using the random initialization of the model's weights. In this process, we initialize the training of each CNN model with random parameter values (weights and biases) and, consequently, no pre-trained knowledge. The training process runs for 100 epochs with hyperparameters (such as the number of convolutional filters, their sizes, padding, and stride) of each CNN set as suggested in the papers that presented the models. We use the backpropagation algorithm to propagate the error backward throughout the CNN and update its parameter values (weights). This process is performed for each training sample while the validation set is used to calculate the model's accuracy at each epoch. During the training process, CNN models are not over-fitting after 100 epochs. We observe this behavior by monitoring performance on the validation set and stop the training at 100 epochs because we are limited by computation. The final trained model is the one with the smallest validation error after completing the training process.

4.2 Training with ImageNet Pre-trained Weights

We extend previous experiments by using the ImageNet [13] pre-trained weights to initialize the training process of the CNN models. We use these pre-trained weights as base models for the next experiments over the UHManoa300 dataset. As suggested by previous results on datasets with 100 plant species (Krause *et al.* [5]), CNN models implementing inception modules (such as the *Inception-v3*, *Inc-ResNet-v2*, and *Xception*) take the most advantage of the ImageNet pre-trained weights. These CNN models have millions of parameters and are better fine-tuned over small datasets when using pre-trained weights as initial parameter values.

4.3 Training with ImageNet and UHManoa100 Pre-trained Weights

We use ImageNet pre-trained weights as a starting point for the training of the CNN models over different datasets to integrate domain-specific knowledge from plant images. The first one is the UHManoa100 dataset, a collection of plant samples from 100 species previously presented. The best performing CNN model after 100 epochs of training creates a new base model with domain-specific knowledge, enabling the extraction of the ImageNet+UHManoa100 pre-trained weights. Consequently, these base models are an integration of previously learned knowledge built over the ImageNet initial weights and what is learned during the training process of these models over the UHManoa100 dataset.

As described previously, we remove the top layers of these models and save their weights and biases (parameter values). In this way, knowledge learned during previous experiments can be used during the fine-tuning process of models adapted for the UHManoa300 dataset. It should be noted that the knowledge integration of the ImageNet pre-trained weights trained over the UHManoa100 dataset creates a domain-specific model that may provide better initial weights for fine-tuning models with new target datasets.

4.4 Training with ImageNet and iNat682 Pre-trained Weights

The iNat682 dataset is a collection of plant images from *iNaturalist* community with 682 species. We downloaded this dataset from the *iNaturalist* challenge website[6] for the classification of animals and plants. However, our training process selects only those images from the *Plantae* category, excluding other groups such as animals, insects, and fungus. The resulting training set is a highly unbalanced collection of 158,463 images over 682 plant species (and these images are not part of the UHManoa100 or UHManoa300 datasets). Ranging from 19 to 503, the number of images per class varies according to the endemic nature of each species worldwide. Furthermore, these images vary in resolution, orientation, and focus, making this dataset a very diverse collection of plant images.

In an attempt to create more robust domain-specific CNN models, experiments using the ImageNet+iNat682 pre-trained weights integrate the general knowledge from the ImageNet dataset and the knowledge from the iNat682 dataset to create plant-expert models. For this process, we initiate the training of the CNN over the iNat682 dataset with the ImageNet pre-trained weights for 100 epochs. In such a way, the iNat682 dataset contributes to the intermediate training process undertaken to create a powerful domain-specific CNN model that learns useful plant-related features for the fine-grained categorization of other plant datasets. We extract the learned knowledge by saving pre-trained weights of the CNN models with the best validation performance, creating the ImageNet+iNat682 pre-trained weights.

5 Experimental Results

Focusing on the UHManoa300 dataset, we implement a CNN-based approach to expand the scope of plant categorization systems to a larger environment of 300 species. Due to the clean-up previously performed in this dataset (Section 3.1), preprocessing these plant images creates highly representative species samples for training and testing. For this balanced dataset, the testing set comprises 10% of the data and the rest is the training set. We perform the fine-tuning process of the CNNs over the UHManoa300 dataset for 100 epochs. During this process, we divide the training set of extracted samples into training (80%) and validation (20%). It is important to reinforce that images selected for the validation set have all their samples for validation only. In this way, we use training and validation samples exclusively in their respective sets for fine-tuning the models. We evaluate the trained CNNs using the testing set of images unseen by the trained models, containing five different images of each plant species. As a metric for performance evaluation, we use the prediction accuracy, e.g., the percentage of images correctly categorized in the testing set. We consider an image is correctly categorized when the Top-1 prediction matches the annotated species of the plant.

[6] https://www.kaggle.com/c/inaturalist-challenge-at-fgvc-2017

Table 1 presents the Top-1 classification accuracy of CNN models pre-trained on different integrated datasets and fine-tuned over the UHManoa300 dataset. As shown in the Table, the CNNs' performance is improved by pre-training them multiple times to integrate domain-specific knowledge. This integration starts with the commonly used ImageNet weights and trains models initially loaded with these parameter values on different plant datasets. It includes pre-trained weights from 100 plant species (ImageNet+UHManoa100) and expert models trained with a larger domain-specific dataset (ImageNet+iNat682). The use of pre-trained weights allows the CNN-based systems to improve their accuracy for the categorization of 300 plant species. More specifically, this approach correctly categorized 84% of the testing images when the *Xception* is pre-trained as a plant-expert model and fine-tuned over the target dataset.

Table 1. Accuracy results with CNNs pre-trained on different dataset for classifying plant species in the UHManoa300 dataset.

CNN model	***Random Initialization***	***ImageNet Weights***	***ImageNet+ UHManoa100***	***ImageNet+ iNat682***
Inception-v3	51.93%	75.67%	76.07%	78.80%
Inc-ResNet-v2	52.00%	76.73%	77.07%	82.33%
Xception	52.40%	81.20%	81.40%	**84.00%**

As shown in Table 1, CNNs fine-tuned for the UHManoa300 achieved more accurate results when pre-trained as domain expert models. Initially, ImageNet pre-trained weights bring a general knowledge with models trained to classify 1,000 common objects. The ImageNet pre-trained models are commonly employed in numerous computer vision problems, but they are limited to the lack of domain-specific knowledge required for fine-grained categorization problems. In the process of creating plant-expert models, we use domain-specific datasets to train the CNNs before fine-tuning them over the target dataset. We collect the ImageNet+UHManoa100 pre-trained weights from CNN models that yielded the best performance over the UHManoa100 dataset. However, CNNs fine-tuned with these pre-trained weights resulted in just slightly more accurate models than those with no domain-specific knowledge integration (ImageNet). Although being a well-organized dataset with multi-scale samples of the original images, the UHManoa100 lacks diversity and is limited to 100 plant species. Hence, knowledge integration is not that evident since the CNN models inherited just a few discriminating features from the UHManoa100 dataset.

In the process of creating plant-expert models, we use a dataset covering much more variety of plant species to train the CNNs before fine-tuning them for the target environment. The training process over this domain-specific dataset helps the models to learn more discriminative features and better generalize for objects from that domain. Thus, we use natural images of different species in the iNat682 dataset to train plant-expert CNN models. These models produce plant-related pre-trained weights that serve as initial parameter values for the fine-tuning process over the UHManoa300 dataset. Consequently, the knowledge integration from an extensive dataset such as ImageNet and a domain-specific dataset such as the iNat682 resulted in better pre-trained CNN models for the plant categorization.

6 Observations and Discussions

The proposed approach described in this paper focuses on expanding the plant categorization scope of CNN-based systems to 300 plant species. As the first step, a botanist helped with the creation and organization of a new dataset (UHManoa300), ensuring a collection of good quality natural images of correctly annotated plants. Subsequently, classification CNN models need to be trained over this new dataset, and common methods such as retraining them from scratch resulted in poor performance. These state-of-the-art CNN models trained from scratch with randomly initialized parameters tend to overfit on training data with few images per class, resulting in incorrect predictions for test images. However, an adaptation on the top layers of the classification models accommodates a larger number of species and allows the use of pre-trained weights, improving the models' accuracy.

State-of-the-art CNN architectures implementing inception modules have shown improved results when using the ImageNet pre-trained weights (Krause *et al.* [5]). These base models trained to categorize the UHManoa100 dataset achieved highly accurate results, serving as plant-related pre-trained weights for the experiments with the UHManoa300 dataset. Furthermore, we integrate domain-specific knowledge using a much larger dataset (iNat682) into the categorization CNNs to create plant-expert models. The integration of domain-specific knowledge also helps to avoid the overfitting problem often encountered when training large CNN models over small datasets. Presented in Table 1, experimental results show the improvement of the CNN-based system when the classification models are pre-trained with domain-specific datasets. This accuracy gain is more evident when a large dataset, such as the iNat682, is used to generate the base models. Training over a big plant-related dataset, base models become plant-expert CNNs and assist on the fine-tuning of the classification models over the UHManoa300 dataset. As a result, the adaptation of previously trained CNNs accommodates an extended categorization scope and allows the CNN-based system to upgrade its models for target datasets with a larger number of classes.

Even with the ease of downloading ImageNet pre-trained weights[7], the creation of plant-expert models demands the learning of thousands of plant images. For instance, integrate domain-specific knowledge from the iNat682 dataset requires a lot of additional computational effort. The ImageNet+iNat682 plant-expert models required the full training of each CNN over 158,463 images before the final fine-tuning process. As an example, for the experiments presented in

[7] https://keras.io/applications/

this paper, we used three GPUs (two *GeForce RTX 2070* and one *GeForce GTX 1080*), and it took almost two months to completely train all CNN models. Among those, the best performing model (*Xception*) took three weeks to be fully trained and fine-tuned.

Experimental results presented in Table 1 show that CNN models improved their predictions when fine-tuned using domain-specific pre-trained weights. The results also show that the *Xception* [2] is the most effective CNN model for classifying 300 plant species. Focusing on the predictions of this model, this fine-tuned CNN can correctly categorize an average of 2.7% (81 testing images) more when using the ImageNet+iNat682 pre-trained weights (compared to other pre-trained weights). In particular, the *Xception* model using the ImageNet+iNat682 pre-trained weights correctly categorize ten images that are not even listed in the Top-5 predictions of models using other pre-trained weights. Figure 3 presents some of these images showing possible discriminative features that this CNN (*Xception*) has inherited from its respective plant-expert model. Visually reviewing these images, they all resembled the shape of small trees with voluptuous treetops. The iNat682 dataset has multiple annotated images of trees (not including any images from the UHManoa300 dataset) that create transferable knowledge related to these types of plants. Representative images of entire trees are not common for some species in the target dataset, causing categorization errors. As suggested by experiments performed in this paper, these errors can be remediated by integrating plant domain-specific knowledge.

Expanding the plant categorization scope brings the challenges of gathering a new target dataset and adapt the classification models to work with the new scope. This paper addresses both of these challenges, but one problem is noticeable in all plant species categorization system: the categorization of different plant components such as leaves, flowers, fruits, barks, etc. As shown in Figure 4, some incorrectly categorized plants do not have their leaves completely visible, while the flowers of that species are evident in the images. In these cases, CNNs trained specifically to classify flowers may successfully categorize the plant species. Therefore, we designed our expandable CNN-based approach to be capable of adapting multiple categorization models, even if they focus on different objectives – i.e., new expert CNN models can be created focusing on each one of the plant's components. In the literature, previously implemented solutions using multiple CNN models to analyze different plant components ([5], [7], and [14]) suggest an improvement in accuracy when categorizing plant images. Therefore, the addition of multiple expert CNN models may be a suitable alternative to better handle natural images of plants. Future experiments will consider expanding the flower categorization scope as well as integrating domain-specific knowledge from flower-related datasets.

7 Conclusion

In this paper, we study the problem of expanding the categorization scope of CNN-based plant phenotyping systems. Amongst the many challenges of this

Fig. 3. Images correctly categorized by the expanded CNN-based approach using the *Xception* model with ImageNet+iNat682 pre-trained weights.

problem, we address the particular challenges of gathering a new representative dataset for a larger environment, adapting the CNN classification layers to accommodate the larger scope, and integrating domain-specific knowledge to maintain a high categorization accuracy. The contributions of this paper include the publication of a new dataset (UHManoa300) with 50 correctly annotated images for each of the 300 plant species, the adaptation process implemented in the top layers of the CNN classification models to accommodate an expanded scope, and the integration of domain-specific knowledge to create plant-expert models (also available online). The pre-trained weights of these expert models can be used by other researchers as a starting point for the fine-tuning process over their new datasets.

Among the challenges, the creation of plant-expert models by integrating domain-specific knowledge is the most demanding one. We implement this inte-

Hibiscus rosa-sinensis *Thunbergia grandiflora*

Fig. 4. Images incorrectly categorized by the expanded CNN-based approach.

gration process by repeatedly training the classification models over plant-related datasets to maintain the high categorization accuracy over a more significant number of species. By dedicating an enormous computational effort to train and fine-tune the modified models, we successfully expand the plant categorization scope to a broader environment while maintaining high accuracy.

In summary, the proposed approach provides a scalable solution for the problem of species categorization and expanding an existing plant phenotyping system to a larger environment. The future work of this research includes developing a unified approach to consolidate various plant phenotyping systems and integrating the knowledge previously learned by them to create better trained CNN expert models.

References

1. Barre, P., Stover, B., Muller, K., Steinhage, V.: LeafNet: A Computer Vision System for Automatic Plant Species Identification. Ecological Informatics **40** (2017)
2. Chollet, F.: Xception: Deep Learning with Depthwise Separable Convolutions. CoRR **abs/1610.02357** (2016)
3. Cui, Y., Song, Y., Sun, C., Howard, A., Belongie, S.J.: Large Scale Fine-Grained Categorization and Domain-Specific Transfer Learning. CoRR **abs/1806.06193** (2018)
4. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. CoRR **abs/1512.03385** (2015)
5. Krause, J., Baek, K., Lim, L.: A Guided Multi-Scale Categorization of Plant Species in Natural Images. In: CVPR Workshop on Computer Vision Problems in Plant Phenotyping (CVPPP 2019). IEEE Press (2019)

6. Krause, J., Sugita, G., Baek, K., Lim, L.: What's That Plant? WTPlant is a Deep Learning System to Identify Plants in Natural Images. In: BMVC Workshop on Computer Vision Problems in Plant Phenotyping (CVPPP 2018). BMVA Press (2018)
7. Krause, J., Sugita, G., Baek, K., Lim, L.: WTPlant (What's That Plant?): A Deep Learning System for Identifying Plants in Natural Images. In: Proceedings of the International Conference on Multimedia Retrieval (ICMR 2018). ACM Press (2018)
8. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 25, pp. 1097–1105. Curran Associates, Inc. (2012)
9. Kumar, N., Belhumeur, P.N., Biswas, A., Jacobs, D.W., Kress, W.J., Lopez, I.C., Soares, J.V.B.: Leafsnap: A Computer Vision System for Automatic Plant Species Identification. In: Fitzgibbon, A.W., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV (2). pp. 502–516 (2012)
10. Lee, S.H., Chan, C.S., Wilkin, P., Remagnino, P.: Deep-plant: Plant Identification with Convolutional Neural Networks. 2015 IEEE International Conference on Image Processing (ICIP) pp. 452–456 (2015)
11. Mo, X., Cheng, R., Fang, T.: Pay Attention to Convolution Filters: Towards Fast and Accurate Fine-Grained Transfer Learning. CoRR **abs/1906.04950** (2019)
12. Ngiam, J., Peng, D., Vasudevan, V., Kornblith, S., Le, Q.V., Pang, R.: Domain Adaptive Transfer Learning with Specialist Models. CoRR **abs/1811.07056** (2018)
13. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV) **115**(3), 211–252 (2015). https://doi.org/10.1007/s11263-015-0816-y
14. Rzanny, M., Mäder, P., Deggelmann, A., Chen, M., Wäldchen, J.: Flowers, leaves or both? How to obtain suitable images for automated plant identification. Plant Methods **15**(77), 1746–4811 (2019)
15. Sun, Y., Liu, Y., Guan, W., Zhang, H.: Deep Learning for Plant Identification in Natural Environment. Computational Intelligence and Neuroscience **2017**(7361042), 6 pages (2017)
16. Szegedy, C., Ioffe, S., Vanhoucke, V.: Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. CoRR **abs/1602.07261** (2016)
17. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the Inception Architecture for Computer Vision. CoRR **abs/1512.00567** (2015)