

Article

Urban Plants Classification Using Deep-Learning Methodology: A Case Study on a New Dataset

Marina Litvak ^{*,†} , Sarit Divekar and Irina Rabaev [†] 

Software Engineering Department, Shamoon College of Engineering, 56 Bialik St., Be'er Sheva 8410802, Israel; saritdi@ac.sce.ac.il (S.D.); irinar@ac.sce.ac.il (I.R.)

* Correspondence: marinal@ac.sce.ac.il

† These authors contributed equally to this work.

Abstract: Plant classification requires the eye of an expert in botanics when the subtle differences in stem or petals differentiate between different species. Hence, an accurate automatic plant classification might be of great assistance to a person who studies agriculture, travels, or explores rare species. This paper focuses on a specific task of urban plants classification. The possible practical application of this work is a tool which assists people, growing plants at home, to recognize new species and to provide the relevant caring instructions. Because urban species are barely covered by the benchmark datasets, these species cannot be accurately recognized by the state-of-the-art pre-trained classification models. This paper introduces a new dataset, Urban Planter, for plant species classification with 1500 images categorized into 15 categories. The dataset contains 15 urban species, which can be grown at home in any climate (mostly desert) and are barely covered by existing datasets. We performed an extensive analysis of this dataset, aimed at answering the following research questions: (1) Does the Urban Planter dataset provide enough information to train accurate deep learning models? (2) Can pre-trained classification models be successfully applied on Urban Planter, and is the pre-training on ImageNet beneficial in comparison to the pre-training on a much smaller but more relevant dataset? (3) Does two-step transfer learning further improve the classification accuracy? We report the results of experiments designed to answer these questions. In addition, we provide the link to the installation code of the alpha version and the demo video of the web app for urban plants classification based on the best evaluated model. To conclude, our contribution is three-fold: (1) We introduce a new dataset of urban plant images; (2) We report the results of an extensive case study with several state-of-the-art deep networks and different configurations for transfer learning; (3) We provide a web application based on the best evaluated model. In addition, we believe that, by extending our dataset in the future to eatable plants and assisting people to grow food at home, our research contributes to achieve the United Nations' 2030 Agenda for Sustainable Development.



Citation: Litvak, M.; Divekar, S.; Rabaev, I. Urban Plants Classification Using Deep-Learning Methodology: A Case Study on a New Dataset. *Signals* **2022**, *3*, 524–534. <https://doi.org/10.3390/signals3030031>

Academic Editors: Francisco Martínez González and Mohammed K. A. Kaabar

Received: 24 April 2022

Accepted: 1 August 2022

Published: 3 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: plant classification; deep-learning neural networks; transfer learning; plants dataset

1. Introduction

Often, plant classification requires the eye of an expert in botanics. Subtle differences in leaves or petal forms might differentiate between different species. On the contrary, there may be high intra-class variability, where species belonging to the same class exhibit very different visual characteristics. Therefore, an accurate automatic plant classification might be of great assistance to a non-expert person who studies agriculture, travels, or grows plants at home.

Plants classification from their images is just an application of a more general task of image classification. In order to train supervised models for this task, one needs a large volume of high-quality training data. However, not many datasets with plant images categorized by species are publicly available for research, and those which are publicly available are far from covering all plant species over the world.

This paper introduces a new dataset, Urban Planter, for plant species classification with 1500 images categorized into 15 categories. The motivation behind the Urban Planter dataset was to collect datasets of plant species growing in our district, which are barely covered by existing datasets. This research may have a practical application in the form of a tool that helps people who cultivate plants at home recognize new species and provide appropriate care recommendations. Because urban species are hardly covered by benchmark datasets, state-of-the-art pre-trained classification methods cannot reliably recognize them. Furthermore, we hope that, by expanding our dataset to include edible plants in the future and supporting people in growing food at home, our research will contribute to the UN's 2030 Agenda for Sustainable Development. The dataset contains 15 house and garden plant species that can be grown mostly in a desert climate and are barely covered by existing datasets. The dataset was collected to develop a mobile application assisting urban planters in identifying plants and discovering best growing methods. However, this paper does not describe the application but focuses on the analysis of Urban Planter. We report the results of experiments performed on the new dataset. The experiments aimed at testing the quality of Urban Planter. We explored whether accurate classification models can be trained on an Urban Planter. In addition, we explored transfer learning with two other datasets: ImageNet (<http://www.image-net.org/> (accessed on 24 April 2022)), which is a classical choice for pre-training, and Oxford102—a much smaller dataset but more relevant to plants classification. We also experimented with two-steps transfer learning, where models, pre-trained first on ImageNet and then on Oxford102, were trained and applied on Urban Planter. The results show that, although Oxford102 is more related to plants classification, the size and the rich diversity of ImageNet are advantageous for accurate classification.

2. Related Work

During the last two decades, various computer vision techniques were employed for plant species classification. Earlier methods used manually defined features, mostly based on the combination of shape, color and texture descriptors, and statistical information [1–8]. Local features descriptors, e.g., HOG [9] and SIFT-based, have also been applied for flower analysis [10–12]. Manually defined and local feature descriptors are fed to the traditional machine learning models, e.g., k-NN [10] and SVM [9]. Machhour et al. [13] introduced a method for plant classification by analyzing leaf images. The method extracts invariants from the shifted Legendre–Fourier moments feed them to a fully connected artificial neural network. The comprehensive survey on the studies of computer vision approaches for plant species identification can be found in [14].

With the advances of hardware, especially with the incorporated use of GPUs, deep neural networks (DNNs) have achieved new standards in many research frontiers. The main advantage of the DNNs is that they do not require manual feature extraction, and the features are learnt within a DNN framework. However, DNNs require a large amount of training data, which is not always available. Pearline and Kumar [15] compared between a deep learning model (VGG19) and conventional machine learning methods, and showed that DNN yielded a higher accuracy for all (four) datasets of plant images. For small or moderate size datasets, transfer learning can help to overcome dataset size limitation [16]. In the case of plant classification, ImageNet is a common choice for a transfer learning scenario [17–20]. Xia and Xu [17] based their flower classifier on an Inception-v3 model, trained on ImageNet. Wu et al. [19] explored the effect of transfer learning for flower classification using four deep-learning models, VGG16, VGG19, Inception-v3, and ResNet. They showed that pre-training the models avoids over-fitting and improves the recognition accuracy. Hiary et al. [18] presented a two-step deep-learning classifier, where first a flower is segmented from the background and then is classified. This framework is based on the VGG16 model with pre-trained weights. Ref. [21] introduced a deep learning system for diverse plants classification in agriculture applications. The experiment, performed on the Plant Seedlings Dataset, aimed to determine which of three pre-trained models—Inception-v3,

VGG16, and Xception—reaches the best accuracy. Results determined that Xception is the best performing model.

With the extensive use of mobile devices, lightweight DNNs are desirable. The MobileNet model was specially designed for usage in mobile applications. MobileNet significantly reduces the number of parameters compared to other conventional deep-learning models, such as VGG16 or Inception-v3. Gavai et al. [20] experimented with MobileNet models for flower classification on the Oxford102 dataset. They showed that MobileNets give comparable performance accuracy while being much smaller and requiring 2.5 times less computation. In [22], a feature extraction using different DNNs—ResNet50-v2, Inception ResNet-v2, MobileNet-v2, and VGG16—was explored in large-scale plant classification methods. A comparative evaluation on the PlantCLEF2003 dataset showed the superiority of the SVM classifier with MobileNet-v2 as a feature extractor. In [23], a comparative evaluation of four deep convolutional feature extraction models—MobileNet-v2, VGG16, ResNet-v2, and Inception-ResNet-v2—tested with the SVM classifier, showed the superiority of MobileNet-v2 on the Vietnamese plant image dataset. MobileNet combination with Logistic Regression was also the best performing system for leaf classification in [24], based on comparative evaluation on two botanical datasets—Flavia (32 classes) and Leafsnap (184 classes).

In most works on plant classification [17–20], the experiments were performed on the Oxford102 dataset [9,10], as one of the largest available datasets of plant images classified into 102 categories. Some works also explored other available datasets, such as the Plant Seedlings Dataset [21], containing images of approximately 960 unique plants belonging to 12 species at several growth stages, and the PlantCLEF2003 dataset [22], consisting of 51,273 images from 609 plant species. The authors of [23] described the Vietnamese plant image dataset, collected from an online encyclopedia of Vietnamese organisms and the Encyclopedia of Life, and containing a total of 28,046 environmental images of 109 plant species in Vietnam.

However, none of these datasets fully covers desert urban plants that our research focuses on. We see an important mission in our task—an accurate classifier for plants that can be grown at home in desert conditions can be of much help to small business, farmers, and individual planters that deal with growing decorative, medicinal, or edible plants. Therefore, we decided to collect our own dataset for training supervised classifiers.

The aim of this paper is threefold. First, we introduce a new dataset, Urban Planter, of desert plants, which are unique in their own way. Second, we analyze the quality of this dataset for the plant classification task. Third, we extend further the study of ImageNet and Oxford102 pre-training for plant classification task by exploring a larger number of deep-learning models.

3. Urban Planter Dataset

The dataset was especially collected and annotated by our research team. We photographed the plants in the countryside and public gardens, while trying to choose the underrepresented species not covered by the existing dataset.

The Urban Planter dataset covers 15 species of houseplants images, 100 images per each. Some species have a unique visual appearance, for example, Begonia Maculata; others have a very similar appearance, for example, House Leek and Paddle Plant (see Figure 1). There are large viewpoint, scale, and illumination variations. The large intra-class variability and sometimes small inter-class variability make this dataset very challenging for the plant classification task. The plant categories are deliberately chosen to have some ambiguity on each aspect. For example, some classes cannot be distinguished on color alone (e.g., moon cactus, nerve plant, poinsettia), others cannot be distinguished on shape alone (e.g., coleus), as illustrated in Figure 2. The majority of the images were photographed by our team. In addition, the plant images were retrieved from multiple sources, including numerous websites (USDA Plants Dataset, Missouri Botanical Garden Database, Better Homes & Gardens, the Urban Nursery, the National Gardening Association Database, House Plants Expert, RHS, and ASPCA), social networks, and self-made photographs of

house plants. Table 1 contains a summary of 15 species covered by the Urban Planter dataset. For experiment, the dataset was traditionally split into 70% training, 10% validation, and 20% test sets. According to [25], all three sets are necessary for fitting a classifier to a new domain. Namely, *training set* is used for learning, that is, to fit the parameters of the classifier; *validation set* is used for fine-tuning, that is, to tune the parameters of a classifier (for example, the number of hidden units in a neural network), while a *test set* is used to evaluate the performance of a fully-specified classifier.

Table 1. The Urban Planter Dataset summary.

| ID | Class | Scientific Name | Higher Classification | Habitat |
|----|--------------------|----------------------------|----------------------------|--------------------------------------|
| 0 | Begonia Maculata | Begonia maculata | Begonia | Brazil |
| 1 | Coleus | Coleus | Ocimeae | Southeast Asia and Malaysia |
| 2 | Elephant’s Ear | Colocasia | Aroideae | Pacific Islands |
| 3 | House Leek | Sempervivum | Stonecrops | Sahara Desert and Caucasus |
| 4 | Jade Plant | Crassula ovata | Pigmyweeds | South Africa |
| 5 | Lucky Bamboo | Dracaena sanderiana | Dracaena | Southeast Asia |
| 6 | Moon Cactus | Gymnocalycium mihanovichii | Gymnocalycium | tropical and subtropical America |
| 7 | Nerve Plant | Fittonia albivenis | Fittonia | South America |
| 8 | Paddle Plant | Kalanchoe luciae | Kalanchoideae | South Africa |
| 9 | Parlor Palm | Chamaedorea elegans | Chamaedorea | Southern Mexico and Guatemala |
| 10 | Poinsettia | Euphorbia pulcherrima | Euphorbia subg. Poinsettia | Central America |
| 11 | Sansevieria Ballyi | Sansevieria Ballyi | Asparagaceae | Africa, Madagascar and southern Asia |
| 12 | String Of Banana | Senecio rowleyanus | Ragworts | South Africa |
| 13 | Venus Fly Trap | Dionaea muscipula | Dionaea | Carolinas |
| 14 | Zebra Cactus | Haworthia attenuata | Haworthiopsis | South Africa |



Figure 1. Examples of the fifteen species from the Urban Planter dataset. Note the similarity between Paddle Plant and House Leek (top row, right), and between Coleus and Poinsettia (the middle and last images in the last column).

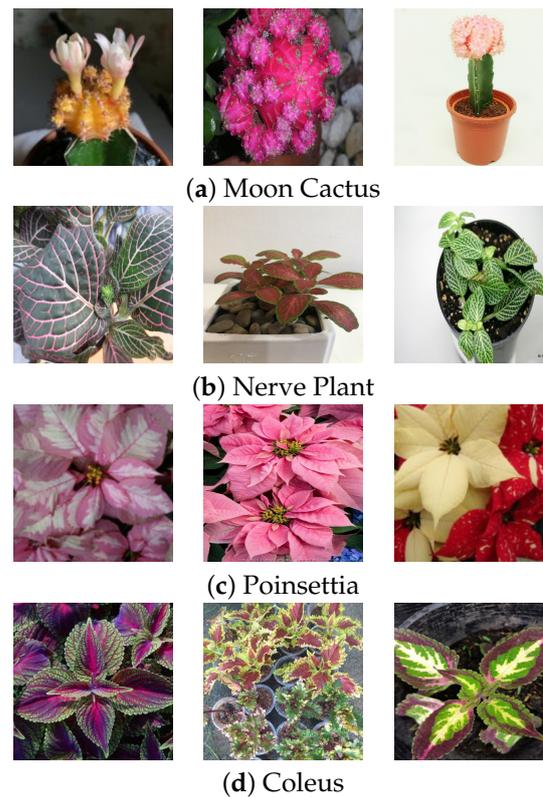


Figure 2. Classes that cannot be distinguished on color alone (moon cactus, nerve plant, poinsettia) or shape alone (coleus).

4. Case Study

4.1. Methods

All recent most successful models for image classification are CNN-based. It has been shown that shallow layers extract simple (low-level) features of an image, and deeper layers can extract more complex (high-level) features. Thus, to make CNN more accurate, researchers mainly increase their depth by adding more layers. Table 2 summarizes all the networks applied in our study—including their architecture, size, and number of parameters—in the chronological order of their introduction. The size of the networks addresses the models pre-trained on ImageNet. Below, we briefly explain about each model.

Table 2. Summary of networks used in our study.

| Network | Used Models | Architecture | Size | Params |
|-----------|---------------------|--------------------------------------|--------|--------|
| VGGNet | VGGNet16 | 13 convolution and 3 fully connected | 113 MB | 138 M |
| | VGGNet19 | 16 convolution and 3 fully connected | 153 MB | 144 M |
| Inception | Inception-v3 | 48 layers | 169 MB | 24 M |
| | Inception-ResNet-v2 | 164 layers | 419 MB | 56 M |
| Xception | Xception | 36 convolution layers | 160 MB | 24 M |
| DenseNet | DenseNet201 | 5 convolution (201 total) layers | 144 MB | 20 M |
| MobileNet | MobileNet-v2 | 3 convolution (20 in total) layers | 19 MB | 13 M |

One of the first successful inventions that demonstrated that the representation depth is beneficial for the classification accuracy is VGGNet [26], introduced in 2014 for large-scale image classification. VGGNet is composed of a sequence of convolutional and pooling layers, followed by three dense layers. In this work, we use VGG16 and VGG19. The main difference between them is a number of convolutional layers.

The Inception [27] deep convolutional architecture was introduced in 2015 (Inception-v1). Later, the Inception architecture was refined, first by the introduction of batch normalization [28] in Inception-v2, subsequently by additional factorization ideas [29] in Inception-v3. Next year, this architecture was refined again, in [30], and several architectures for Inception-ResNet, including Inception-ResNet-v2, were proposed. In our case study, we use two Inception-based models: Inception-v3 and Inception-ResNet-v2. The difference between them is that Inception-v3 is a deep CNN not utilizing residual connections, while Inception-ResNet-v2 is Inception style networks that utilize residual connections instead of filter concatenation.

The Xception [31] model was proposed in 2017. It is an extension of the Inception architecture which replaces the standard Inception modules with depthwise Separable Convolutions.

The Xception architecture has the same number of parameters as Inception-v3, but its performance is better due to more efficient use of these parameters.

While CNNs go deeper and the path from the network input layer to its output layer becomes longer, the chance of information to reach the other side gets lower. DenseNet [32], introduced in 2017, solves this problem by ensuring maximum information flow with connections from each layer to every other layer in a feed-forward fashion. DenseNet has several compelling advantages: it alleviates the vanishing-gradient problem, strengthens feature propagation, exploits the potential of the network through feature reuses, and substantially reduces the number of parameters.

MobileNets [33], also introduced in 2017, are based on a streamlined architecture that uses depthwise separable convolutions to build light-weight deep neural networks. The inventors of MobileNets empirically showed that smaller and faster MobileNets can be built, using width and resolution multipliers by trading off a reasonable amount of accuracy to reduce size and latency.

We used open-source implementations of all mentioned models using Keras and TensorFlow, which are provided as part of the Keras Applications module <https://keras.io/api/applications/> (accessed on 24 April 2022). All the networks we used come with predefined parameters and weights (pre-trained), only the number of epochs during fine-tuning was set to a number with which a convergence of the loss function was reached.

4.2. Datasets for Transfer Learning

We used two external datasets for transfer learning:

- ImageNet. ImageNet was used for transfer learning in many tasks and domains in computer vision. All DNNs that we applied are available with weights pre-trained on ImageNet. The dataset contains about 1.2 M images classified into 1 K categories;
- Oxford102. In contrast to ImageNet, this dataset is much closer to the plants domain. It contains about 8000 images of flowers assigned into 102 species.

The motivation behind using these datasets was threefold: (1) to see whether transfer learning is efficient for the plants classification in our dataset, (2) to compare the gain of pre-training on a large general dataset with the gain of pre-training on a more specific and relevant but much smaller dataset, and (3) to check whether both datasets can be used for optimizing pre-trained models.

4.3. Experiment Scenarios

We performed the following experiments:

1. Training the models from scratch, i.e., with random initialization, on the Urban Planter dataset (denoted by 0-TL);
2. One-step transfer learning using Oxford102 (denoted by 1-TL-Ox), where models, pre-trained on Oxford102, are trained and applied on Urban Planter.
3. One-step transfer learning using ImageNet (denoted by 1-TL-IN), where models, pre-trained on ImageNet, are trained and applied on Urban Planter.
4. Two-steps transfer learning (denoted by 2-TL), where models, pre-trained first on ImageNet and then on Oxford 102, were trained and applied on Urban Planter.

The transfer learning was implemented as follows: we use models' weights pretrained on ImageNet/Oxford102. Then, we performed fine-tuning by replacing the last FC layer with a softmax activation function. This training scheme follows the conventions of transfer learning, where the pre-trained early layers of the network are frozen, and the newly added layers are trained. Then, the early layers are unfrozen, and the model is further trained as a whole. (Each model was trained for 50 epochs. We used the following hyper-parameters: optimizer = `tf.keras.optimizers.RMSprop` ($lr = 0.0001$), loss = "sparse_categorical_crossentropy", metrics = ["Sparse Categorical Accuracy"]).

4.4. Data Preprocessing

In each experiment, the input images are resized to input dimensions of the respective network (see Table 2). During the training, the batches are generated with real-time data augmentation by `tf.keras.preprocessing.image.ImageDataGenerator` Keras method. Except for this, no other adaptation or pre-processing was applied.

5. Results and Discussion

Table 3 contains accuracy scores for each experiment scenario. We can note that the results of the experiment 0-TL are the lowest. This can be attributed to the size of the Urban Planter dataset, i.e., more data are needed to accurately train DNNs with a very large number of parameters. Both Oxford102 and ImageNet pre-training (experiments 1-TL-Ox and 1-TL-IN) improve the results. We have three exceptions—both Inception models and VGG16 perform worse in 1-TL-Ox than in 0-TL. However, all models, including these three, perform significantly better when pre-trained in ImageNet (1-TL-IN).

Table 3. The accuracy rates on the Urban Planter dataset in different experiments.

| | 0-TL | 1-TL-Ox | 1-TL-IN | 2-TL |
|---------------------|---------------|---------------|---------------|---------------|
| Xception | 66.33% | 69.67% | 94.67% | 95.00% |
| Inception-ResNet-v2 | 74.33% | 72.00% | 93.33% | 93.67% |
| Inception-v3 | 71.67% | 62.33% | 91.67% | 90.00% |
| DenseNet201 | 63.33% | 66.33% | 96.00% | 94.67% |
| MobileNet-v2 | 6.67% | 41.67% | 86.33% | 83.67% |
| VGG19 | 51.67% | 59.00% | 70.67% | 75.00% |
| VGG16 | 62.00% | 57.67% | 80.67% | 81.00% |

In general, the results of transfer learning with the ImageNet are higher than those of transfer learning with the Oxford102. This outcome can be explained by the size of the datasets used for pre-training: ImageNet contains over 1.2 M million images, while Oxford102 contains only about 8000 images. It bears evidence that a training size plays a more crucial role for fitting networks' parameters than a training domain.

The results of the experiment 2-TL—two-steps transfer learning—are very close to the results of the experiment 1-TL-IN—transfer learning with ImageNet pre-training only. As such, we can conclude that the following pre-training on Oxford102 has almost no influence on the quality of the classification models pre-trained on ImageNet and their ability to classify plants in Urban Planter. The possible explanation for this outcome is also the superior size of the ImageNet and its rich class diversity. Many features and patterns are shared between different categories, and it seems that Oxford102 hardly adds any new information to it.

In addition, we can note a significant gap in the performance of the MobileNet in the experiment 0-TL. After training for 100 epochs, its classification accuracy reached only 6.67%. However, the accuracy rates of the MobileNet in other experiments are comparable to the other models, and even higher than the accuracy rates of both VGGs in the experiments 1-TL-IN and 2-TL. The MobileNet architecture is much more shallow than the architecture of advanced deep models. Therefore, to learn complex features needed for fine-grained classification, it requires a larger amount of data.

We can see that Xception, Inception-ResNet-v2, and DenseNet201 are the best-performing systems in most scenarios. Their advanced architectures can explain their superiority. Both the Xception and Inception-ResNet-v2 models are the extensions of the Inception architecture. The Inception architecture is characterized by Inception modules—modules that combine multiple layers with their output filter banks concatenated into a single output vector, which forms the next layer’s input. In the DenseNet architecture, each layer receives direct input from all preceding layers; consequently, it can learn very diverse features and patterns.

In comparison to the previous works, the following can be concluded:

(1) No direct comparisons between accuracy scores can be made because our results are obtained on a new—Urban Planter—dataset. Previous results range from 25% to 99% [14] on different datasets;

(2) The majority of studies proposed approaches for plant classification that are based on the analysis of only one part of a plant’s structure. Leaf followed by flower was the most widely studied part [14]. Scientists focused on leaves in plant classification because leaves are available for examination throughout most of the year, they are easy to find and to collect, and they can easily be imaged compared to other plant morphological structures, such as flowers, barks, or fruits. Therefore, while collecting images for Urban Planter, it was important to us to cover full images of the same plants through different seasons and not to focus on leaves or other parts. Our approach to image classification obtains the entire plant’s picture as an input;

(3) Most of the works applied traditional ML methods with feature engineering. In contrast, our approach does not require feature engineering because it applies DNNs which encode the input images automatically;

(4) To the best of our knowledge, our study is the first one that compares between multiple DNNs and number of stages in transfer learning for plants classification. Multiple works applied DNNs for the plant classification in the last couple of years, both as feature extractors [15,22] and as end-to-end classifiers [17–19,19–21,23,24], achieving very good results. Some of them experimented with one deep architecture for plant classification [17,18,20], some compared between multiple networks [19,21–23], and some applied transfer learning [19,24], but neither experimented with both, including different configurations and data sources for transfer learning.

6. Limitations of our Study

One of the limitations of our study is that the Urban Planter dataset is small. Currently, we are extending the dataset to include more images of each plant class.

In addition, the introduced study does not take into consideration a hierarchical structure of a species taxonomy. It is quite intuitive that distinguishing between categories of species is a much easier task than between individual species in the same categories. Therefore, we expect that fine-grained classification [34], which focuses on differentiating between hard-to-distinguish object classes, will improve plant’s classification accuracy. We plan to experiment with fine-grained classifiers in the future.

Another unaddressed problem in our study is that not every species always has enough images in the training set. Currently, our dataset contains the same amount of images per species. However, in the future, we plan to experiment with few-shot learning, which aims at learning a classifier to recognize classes with limited training samples [35]. Given such classifiers, researchers will be less dependent on a coverage of the available training datasets.

7. Conclusions and Future Work

This paper introduces a new dataset—Urban Planter—and reports the results of an extensive empirical case study for urban plants classification with transfer learning with multiple DNNs.

The Urban Planter dataset represents species of houseplants, which can be grown at home in any climate (mostly desert) and barely covered by existing datasets, thus contributing to the diversity of the available plants' datasets.

The study aims at evaluation of multiple DNNs, which are state-of-the-art classifiers used in computer vision, and different configurations of transfer learning with the help of two benchmark datasets. We show that the pre-trained on ImageNet models can classify Urban Planter with high accuracy (94–96% for the best models), and that ImageNet pre-training achieves much higher accuracy rates than pre-training on the smaller Oxford102 dataset. We also show that two-steps transfer learning (ImageNet pre-training followed by Oxford102 pre-training) has almost no effect on the classification score. Thus, we can conclude that pre-training on an extensive general dataset is enough for fitting parameters of a fine-grained classifier. The results of experiments prove that a training size plays a more important role for fitting networks' parameters than a training domain in our case.

The study was conducted as a part of the undergraduate project with the aim to develop a mobile application assisting urban planters. The Urban Planter dataset, the video demo, and the code of the web app for urban plants classification are available on <https://github.com/UrbanPlanter/urbanplanterapp> (accessed on 24 April 2022) (The current version must be installed on a local host. We are currently deploying the server to the Heroku platform <https://www.heroku.com/> (accessed on 24 April 2022)).

To pursue a practical benefit from our study and contribute to the United Nations' 2030 Agenda for Sustainable Development, (<https://sdgs.un.org/2030agenda> (accessed on 24 April 2022)), we plan to extend our dataset and study to contain more species, including edible plants in the future. In addition, we intend to work on a new version of our application, which will provide more helpful information to the end user in addition to the plant's category. This can be done by linking the categorization results with external data sources.

Author Contributions: Conceptualization, M.L.; methodology, M.L. and S.D.; software, S.D.; validation, S.D.; formal analysis, I.R. and M.L.; investigation, M.L.; resources, S.D.; writing—original draft preparation, M.L. and I.R.; writing—review and editing, M.L. and I.R.; supervision, M.L.; project administration, S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The installation code of the alpha version and the demo video of the app can be found on <https://github.com/UrbanPlanter/urbanplanterapp> (accessed on 24 April 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tan, W.N.; Sem, R.; Tan, Y.F. Blooming flower recognition by using eigenvalues of shape features. In Proceedings of the Sixth International Conference on Digital Image Processing (ICDIP 2014), Athens, Greece, 5–6 April 2014; Volume 9159, pp. 344–348.
2. Tan, W.N.; Tan, Y.F.; Koo, A.C.; Lim, Y.P. Petals' shape descriptor for blooming flowers recognition. In Proceedings of the Fourth International Conference on Digital Image Processing (ICDIP 2012), Kuala Lumpur, Malaysia, 7–8 April 2012; Volume 8334; pp. 693–698.
3. Phyu, K.H.; Kutics, A.; Nakagawa, A. Self-adaptive feature extraction scheme for mobile image retrieval of flowers. In Proceedings of the 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, Sorrento, Italy, 25–29 November 2012; pp. 366–373.
4. Hsu, T.H.; Lee, C.H.; Chen, L.H. An interactive flower image recognition system. *Multimed. Tools Appl.* **2011**, *53*, 53–73. [[CrossRef](#)]
5. Hong, S.W.; Choi, L. Automatic recognition of flowers through color and edge based contour detection. In Proceedings of the 2012 3rd International conference on image processing theory, tools and applications (IPTA), Istanbul, Turkey, 15–18 October 2012; pp. 141–146.
6. Cho, S.Y.; Lim, P.T. A novel virus infection clustering for flower images identification. In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, 20–24 August 2006; Volume 2; pp. 1038–1041.

7. Cho, S.Y. Content-based structural recognition for flower image classification. In Proceedings of the 2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), Singapore, 18–20 July 2012; pp. 541–546.
8. Apriyanti, D.H.; Arymurthy, A.M.; Handoko, L.T. Identification of orchid species using content-based flower image retrieval. In Proceedings of the 2013 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Jakarta, Indonesia, 19–21 November 2013; pp. 53–57.
9. Nilsback, M.E.; Zisserman, A. Automated flower classification over a large number of classes. In Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, Bhubaneswar, India, 16–19 December 2008; pp. 722–729.
10. Nilsback, M.E.; Zisserman, A. A visual vocabulary for flower classification. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 2; pp. 1447–1454.
11. Qi, W.; Liu, X.; Zhao, J. Flower classification based on local and spatial visual cues. In Proceedings of the 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE), Zhangjiajie, China, 25–27 May 2012, Volume 3; pp. 670–674.
12. Zawbaa, H.M.; Abbass, M.; Basha, S.H.; Hazman, M.; Hassenian, A.E. An automatic flower classification approach using machine learning algorithms. In Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Delhi, India, 24–27 September 2014; pp. 895–901.
13. Machhour, A.; Zouhri, A.; El Mallahi, M.; Lakhliai, Z.; Tahiri, A.; Chenouni, D. Plants Classification Using Neural Shifted Legendre-Fourier Moments. In Proceedings of the International Conference on Smart Information & Communication Technologies, Oujda, Morocco, 26–28 September 2019; pp. 149–153.
14. Wäldchen, J.; Mäder, P. Plant species identification using computer vision techniques: A systematic literature review. *Arch. Comput. Methods Eng.* **2018**, *25*, 507–543. [[CrossRef](#)] [[PubMed](#)]
15. Anubha Pearline, S.; Sathiesh Kumar, V.; Harini, S. A study on plant recognition using conventional image processing and deep learning approaches. *J. Intell. Fuzzy Syst.* **2019**, *36*, 1997–2004. [[CrossRef](#)]
16. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? *arXiv* **2014**, arXiv:1411.1792.
17. Xia, X.; Xu, C.; Nan, B. Inception-v3 for flower classification. In Proceedings of the 2017 2nd International Conference on Image, Vision and Computing (ICIVC), Chengdu, China, 2–4 June 2017; pp. 783–787.
18. Hiary, H.; Saadeh, H.; Saadeh, M.; Yaqub, M. Flower classification using deep convolutional neural networks. *IET Comput. Vis.* **2018**, *12*, 855–862. [[CrossRef](#)]
19. Wu, Y.; Qin, X.; Pan, Y.; Yuan, C. Convolution neural network based transfer learning for classification of flowers. In Proceedings of the 2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP), Shenzhen, China, 13–15 July 2018; pp. 562–566.
20. Gavai, N.R.; Jakhade, Y.A.; Tribhuvan, S.A.; Bhattad, R. MobileNets for flower classification using TensorFlow. In Proceedings of the 2017 International Conference on Big Data, IoT and Data Science (BIGDATA), Pune, India, 20–22 December 2017; pp. 154–158.
21. Diaz, C.A.M.; Castaneda, E.E.M.; Vassallo, C.A.M. Deep learning for plant classification in precision agriculture. In Proceedings of the 2019 International Conference on Computer, Control, Informatics and its Applications (IC3INA), Tangerang, Indonesia, 23–24 October 2019; pp. 9–13.
22. Van Hieu, N.; Hien, N.L.H. Recognition of Plant Species using Deep Convolutional Feature Extraction. *Int. J. Emerg. Technol.* **2020**, *11*, 904–910.
23. Van Hieu, N.; Hien, N.L.H. Automatic Plant Image Identification of Vietnamese species using Deep Learning Models. *arXiv* **2020**, arXiv:2005.02832.
24. Beikmohammadi, A.; Faez, K. Leaf classification for plant recognition with deep transfer learning. In Proceedings of the 2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, 25–27 December 2018; pp. 21–26.
25. Ripley, B.D. *Pattern Recognition and Neural Networks*; Cambridge University Press: Cambridge, UK, 2007.
26. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
27. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 8–10 June 2015; pp. 1–9.
28. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
29. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
30. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
31. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
32. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–27 July 2017; pp. 4700–4708.
33. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.

-
34. Wei, X.S.; Song, Y.Z.; Mac Aodha, O.; Wu, J.; Peng, Y.; Tang, J.; Yang, J.; Belongie, S. Fine-Grained Image Analysis with Deep Learning: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**. [[CrossRef](#)] [[PubMed](#)]
 35. Chen, W.Y.; Liu, Y.C.; Kira, Z.; Wang, Y.C.F.; Huang, J.B. A Closer Look at Few-shot Classification. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.