

Tree Species Identification from Bark Images Using Convolutional Neural Networks

Mathieu Carpentier, Philippe Giguère and Jonathan Gaudreault

Abstract—Tree species identification using bark images is a challenging problem that could prove useful for many forestry related tasks. However, while the recent progress in deep learning showed impressive results on standard vision problems, a lack of datasets prevented its use on tree bark species classification. In this work, we present, and make publicly available, a novel dataset called *BarkNet 1.0* containing more than 23,000 high-resolution bark images from 23 different tree species over a wide range of tree diameters. With it, we demonstrate the feasibility of species recognition through bark images, using deep learning. More specifically, we obtain an accuracy of 93.88% on single crop, and an accuracy of 97.81% using a majority voting approach on all of the images of a tree. We also empirically demonstrate that, for a fixed number of images, it is better to maximize the number of tree individuals in the training database, thus directing future data collection efforts.

I. INTRODUCTION

The ability to automatically and reliably identify tree species from images of bark is an important problem, but has received limited attention in the vision and robotics communities. Early work in mobile robotics has already shown that the ability to recognize trees from non-trees in combined LiDAR+camera sensing can improve localization robustness [1]. More recent work on data-efficient semantic localization and mapping algorithms [2], [3] have demonstrated the value of semantically-meaningful landmarks; In our situation, trees and the knowledge of their species would act as such semantic landmarks. The robotics community is also increasingly interested in flying drones in forests [4]. In terms of forestry applications, one could use this visual species identification to perform autonomous forest inventory. In the context of autonomous tree harvesting operations [5], the harvester or forwarder would be able to sort timber by species, improving the operator’s margins. Similarly, sawmill processes such as debarking could be fine-tuned or optimized based on the species knowledge of the currently processed log.

For tree species identification, relying on bark has many advantages when compared to other attributes, such as the appearance of its leaves or fruits. First of all, bark is always present despite seasonal changes. It is also present on logs long after the trees have been cut and stored in a lumber yard. In the case of standing tree inventory, bark tends to be visually accessible to most robots, as foliage is not prevalent at the robot’s height in forests of commercial value. However, tree species classification using only images of the bark is a challenging task that even trained humans struggle to do, as some species have only very subtle differences in their bark structure. For example, two human experts obtained

respectively 56.6% and 77.8% classification accuracy on the *Austrian Federal Forests* (AFF) dataset [6].

Recent progress in deep learning have shown that neural networks are able to surpass human performance on many visual recognition tasks [7]. One significant drawback of deep learning approaches is that they generally require very large datasets to obtain satisfactory results. For instance, the ImageNet database contains 14 millions images separated in almost 22,000 synsets.

In the literature, there is no equivalent database for bark recognition, in terms of size or variety. For example, the largest one is the AFF dataset [6], with only around 1,200 images covering 11 species. This dataset is also private, making it difficult to use in an open, scientific context. This lack of data might explain why the majority of research on bark recognition has been mostly centered around hand-crafted features such as Gabor filters [8], [9], SIFT [6] or Local Binary Pattern [10], [11], as they can be trained using smaller datasets.

To address this issue, we gathered a novel bark dataset specifically designed to train deep neural networks. It contains 23,000 high-resolution images of 23 different tree species found in forests and parks near Quebec City, Canada, from which over 800,000 unique crops of 224x224 pixels can be extracted. The species are typical trees present on the eastern seaboard forests of Canada, most of which have commercial value. In addition to providing the species annotation, we also collected the tree diameter at breast height (DBH), a commonly-used metric in forest inventories. The DBH captures in some sense the age of the tree, thus having the possibility to provide auxiliary information to the network during training. Indeed, bark appearance can change drastically with age, which might help a network optimizer in finding solutions that exhibit better generalization performance. Moreover, having this extra label opens up the possibility to experiment with multi-task learning approaches, for which few datasets exists in the literature [12].

The contributions presented in this paper are as follow:

- We collected and curated a novel bark image dataset¹, named *BarkNet 1.0*, that is compatible with deep learning research on fine-grained and texture classification problems. This dataset can also be used in the context of multi-task benchmarking.
- We demonstrated that using this dataset, we can perform visual tree recognition of 20² species, far above any

¹Available at <https://github.com/ulaval-damas/tree-bark-classification>

²For three species, there was an insufficient number of images to perform training and testing.

other work. We also quantify the difficulty of differentiating between certain species, via confusion matrices.

- We performed experiments in order to determine the impact of several key factors on the recognition performance (number of images used during training, use of a voting scheme on classification during testing.)

This paper is organized as follows. In Section II, we review existing methods and datasets used to accomplish bark image classification. Section III introduces our dataset, and details on how it was collected. Section IV describes the network architecture used to perform classification. Section V presents the results obtained for various test cases. Finally, Section VI concludes this paper.

II. RELATED WORK

Bark classification has most frequently been formulated as a texture classification problem, for which a number of hand-crafted features have historically been employed. For instance, some works based their approaches on Local Binary Patterns (LBP) [10], [11], [13] and others [6] used SIFT descriptors combined with a support vector machine (SVM) to obtain around 70% accuracy on the AFF dataset. Meanwhile, [14] extracted four statistical parameters (uniformity, entropy, asymmetry and smoothness) used in texture classification on trunk images, and employed a decision tree for classification. Furthermore, [15] developed a custom segmentation algorithm based on watershed segmentation methods, extracted saliency, roughness, curvature and shape features and fed them to a Random Forest classifier.

Interestingly, some early works used neural networks for bark classification. For instance, [8] extracted texture features based on Gabor wavelet and used a radial basis probabilistic network as the classifier. With their method, they obtained close to 80% accuracy using a dataset containing around 300 images. This work predates, however, the advent of deep learning approaches, spearheaded by AlexNet [16].

With respect to the more general task of tree classification, some did apply deep learning methods. For instance in the *LifeCLEF* competition, which attempts to classify plants using images of different parts such as the leaves, the fruit, or the stem, the best performing methods all employed deep learning [17], [18], [19], [20]. For our purpose however, the number of images with significant bark content in their training database is too small. Less related to the work described herein, work on leaf classification by [21] extracted features from deep neural networks, in order to determine what were the most discriminating factors.

Deep learning has also been employed for tree identification from bark information, but using a different type of image. In their work, [22] used LiDAR scans instead of RGB images. They used a point cloud with a spatial resolution of 5 mm at a 10 m distance, from which they generated a depth image of size 256x256. For the classification, they fine-tuned a pre-trained AlexNet [16] on around 35,000 scans. This allowed them to obtain around 90% precision on their test set containing 1,536 scans. However, they only used

Reference	Number of classes	Number of images	Public
[9]	8	200	No
[8]	17	300	No
Trunk12[27]	12	393	Yes
[14]	5	540	No
[28]	23	920	No
AFF[6]	11	1183	No

TABLE I
EXISTING BARK IMAGE DATASETS

two different species, Japanese Cedar and Japanese Cypress, making the problem significantly less challenging.

Finally, some authors have started exploring deep learning on RGB images of textures. By leveraging extracted features from CNNs pre-trained on ImageNet and different region segmentation algorithms, [23] used an SVM to classify texture materials, notably on the *Flickr Material Dataset* [24]. They also improved the state-of-the-art by at least 6 % on all of the datasets on which they tested. Also, [25] modified the standard convolutional layer to learn rotation-invariant filters. They did this by grouping filters into groups and by tying the weights of each filter within the same group so that they would all correspond to a rotated version of each other. They tested their layer on the three *Outex* [26] texture classification benchmarks and improved the state-of-the-art on one of these benchmarks and obtained similar results on the other two.

III. BARK DATASET (*BarkNet 1.0*)

A. Existing bark datasets

One significant hurdle when trying to use deep learning for bark classification is the lack of existing datasets for training purposes. Table I shows datasets that were used in previous work for the bark classification task. Note that most of these datasets contain only a very small number of images as well as a limited number of classes. Moreover, only one of those datasets is publicly available, hindering the global research effort on this problem.

B. Image collection and annotation

To solve the dataset issue, we collected images from 23 different species of trees found in parks and forests near Quebec City, Canada. We hired a forestry specialist to identify the species on site. Indeed, tree identification is much easier and reliable when relying on extra cues such as leaf shape or needle distribution. To accelerate the data collection process, we used the following protocol. First, a tree was selected and its species and circumference written on a white board by the forestry specialist. While the specialist moved to another tree, a second person took a picture of the white board as the first picture of the tree. It was then followed by 10-40 images of the bark at different locations and heights around this tree, depending on its circumference. Images were captured at a distance between 20-60 cm away from the trunk. This

distance was highly variable, depending on the conditions in which the photos were taken (due to obstacles, tree size, etc.). Having this kind of variability prevents overfitting to a particular distance of camera. Finally, all images were taken so as to have the trunk parallel to the vertical axis of the image plane of the camera.

We also gathered the images under varied conditions, to ensure that the dataset would be as diversified as possible. First, we used four different cameras, some of which were cellphones: Nexus 5, Samsung Galaxy S5, Samsung Galaxy S7, and a Panasonic Lumix DMC-TS5 camera. To increase the illumination variability, we took the pictures under a number of weather conditions which ranged from sunny to light rain. Finally, we selected trees from a number of different locations, such as in open areas like the university campus or parks and in the forest. This can greatly affect the appearance of the bark, especially in high vegetation density locations where the reflection of the canopy can add different shades of green to the bark color. In total, we gathered pictures during 15 outings, which took place during the summer.

From the picture of the white board, we obtained the species and circumference information to annotate the subsequent pictures. This means that each photo in our database contains a unique number identifying the tree, its species, its DBH, the camera used and the date and time at which it was taken. We also curated the dataset by removing approximately 25 % of the pictures, most of them corresponding to blurred images due to camera motion. Each remaining picture was then manually cropped, so as to only keep the part of the image where bark was visible. This had the side effect that younger trees yielded very narrow pictures (Fig. 2 (1)), while mature trees were full-sized pictures (Fig. 2 (3)). Table II shows the composition of our dataset. We aimed at keeping the dataset as balanced as possible, while maximizing the number of different trees used for each class. The data collection strategy was also modulated based on initial classification results. Indeed, we increased the number of trees collected for species that were found to be difficult to separate. One can see this as a loose form of active learning, but implemented with humans in the loop.

We also aimed at having a wide distribution on the DBH which is shown in Fig. 1. Most of the trees have a DBH between 20 and 30 *cm*, but we also have a few trees near 100 *cm*. This can have an impact on the classification since the size of the tree can greatly affect the appearance of the bark. Fig. 2 shows an example of this, with the younger tree having a relatively smooth bark while the older ones are covered with ridges and furrows.

IV. EXPERIMENTS

A. Architecture

As is commonly done in image recognition tasks, we employed networks that have been pre-trained on ImageNet. Moreover, we used the ResNet architecture [29], as it is both powerful and easy to train on standard classification problems.

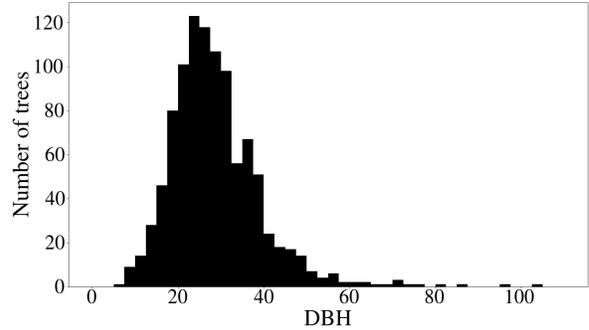


Fig. 1. DBH distribution of the dataset

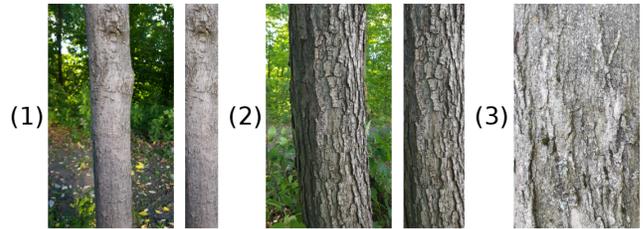


Fig. 2. Three images of *Acer Sacharricum*, to illustrate the impact of DBH on crop size. (1) has a DBH of 8.9 *cm*, (2) has a DBH of 25.8 *cm* and (3) has a DBH of 68.1 *cm*. Note that no cropping was needed for (3)

B. Training Details

We used PyTorch 0.3.0.post4 [30] for all experiments and downloaded the weights of the `resnet18` and `resnet34` networks pre-trained on ImageNet. As commonly-accepted practice, we froze the first layer, since our problem is very different from ImageNet, and then fine-tuned the networks using an initial learning rate of 0.0001. We reduced the learning rate at fixed epochs (16 and 33) by a factor of 5, and trained for a total of 40 epochs. We used Adam as the optimization method, with a weight decay of 0.0001.

Since the photos are high definition, we resized them to half of their original size. This allowed for a faster loading and image processing of the images when creating the mini-batches. It also takes into account the Bayer filter pattern on color cameras, which only samples colors for every other pixel on the imaging element. For each mini-batch, we uniformly sampled a random tree species (class), from which we sampled a random image from a random tree. This allowed us to mitigate the problems of having an unbalanced dataset, similarly to the class-aware sampling used in [31]. Then, we augmented the data using random horizontal flips and finally, we took a random crop of 224x224 pixels in the resulting image. Recall that during the data gathering process, a fair amount of randomness in terms of illumination and scale was present, so we did not perform color, scale or contrast jittering.

V. RESULTS

In our experiments, we compared the effect of network depth (18 vs 34) on classification precision. We also tested

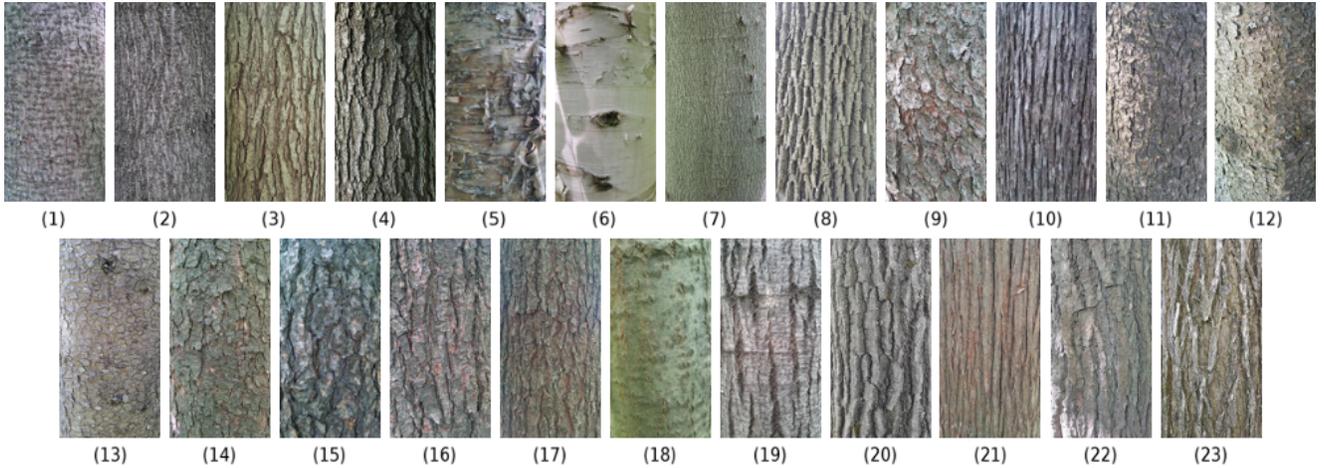


Fig. 3. Example image for each species. The number refers to the ID of each species, detailed in Table II. Some of the pictures have a greener tint, due to the improper white balance of the camera caused by the canopy in the forest.

ID	Species	Common name	Number of trees	Number of images	Number of potential unique crops
1	<i>Abies balsamea</i>	Balsam fir	41	922	28235
2	<i>Acer platanoides</i>	Norway maple	1	70	2394
3	<i>Acer rubrum</i>	Red maple	64	1676	48925
4	<i>Acer saccharum</i>	Sugar maple	81	1999	68040
5	<i>Betula alleghaniensis</i>	Yellow birch	43	1255	37325
6	<i>Betula papyrifera</i>	White birch	32	1285	33892
7	<i>Fagus grandifolia</i>	American beech	41	840	23904
8	<i>Fraxinus americana</i>	White ash	61	1472	53995
9	<i>Larix laricina</i>	Tamarack	77	1902	114956
10	<i>Ostrya virginiana</i>	American hophornbeam	29	612	28723
11	<i>Picea abies</i>	Norway spruce	72	1324	35434
12	<i>Picea glauca</i>	White spruce	44	596	19673
13	<i>Picea mariana</i>	Black spruce	44	885	43127
14	<i>Picea rubens</i>	Red spruce	27	740	22819
15	<i>Pinus rigida</i>	Pitch pine	4	123	2264
16	<i>Pinus resinosa</i>	Red pine	29	596	14694
17	<i>Pinus strobus</i>	Eastern white pine	39	1023	25621
18	<i>Populus grandidentata</i>	Big-tooth aspen	3	64	3146
19	<i>Populus tremuloides</i>	Quaking aspen	58	1037	63247
20	<i>Quercus rubra</i>	Northern red oak	109	2724	72618
21	<i>Thuja occidentalis</i>	Northern white cedar	38	746	19523
22	<i>Tsuga canadensis</i>	Eastern Hemlock	45	986	27271
23	<i>Ulmus americana</i>	American elm	24	739	27821
Total	-	-	1006	23616	817647

TABLE II

BarkNet 1.0 DATASET COMPOSITION. ALTHOUGH WE USED RANDOM CROPS DURING TRAINING, WE INDICATE IN THE LAST COLUMN THE NUMBER OF UNIQUE CROPS (WITHOUT ANY OVERLAP) THAT CAN BE THEORETICALLY GENERATED FOR TRAINING PURPOSES

for different batch sizes, to evaluate its regularization effect [32]. For the evaluation, we used a 5-fold cross-validation method using 80% of the trees for the training and the remaining for testing. Care was taken in performing the split on the trees instead of the image, to avoid positively biasing results due to the network learning to recognize individual trees instead of the species. We report the average accuracy on the 5 folds for two different scenarios: (i) one where we evaluate all the image individually as if they were all from different trees and (ii) one where we classify each tree by using all of its images. Note that we did not use *Acer platanoides*, *Pinus rigida* and *Populus grandidentata*

since we did not collect enough images in these categories to obtain meaningful results.

A. Test results when using individual images

Table III contains the results of evaluating the two models on each image individually, for a number of batch sizes. We report both single crop (random) and multiple crop results. For the latter, we split the test image into multiple non-overlapping 224x224 crops and classified each one individually. Then, we performed majority voting to determine the final outcome. As can be seen from Table III, progressing from single crops (87.04%) to multiple crops (93.88%) on a complete image significantly improves the accuracy, which is

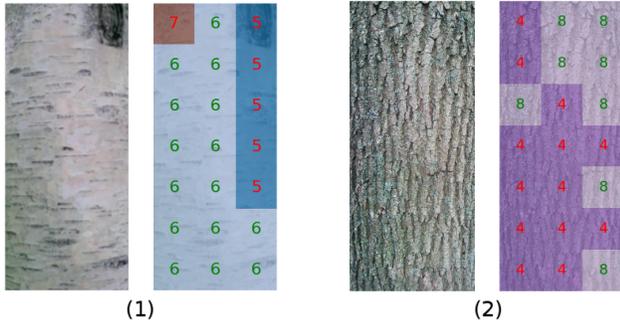


Fig. 4. Examples of multiple crop majority classification. (1) shows a successful classification on *Betula papyrifera* and (2) shows a classification error on *Fraxinus americana*. The green number indicates the correct class and red numbers indicate the incorrect classes. The numbers refer to the ID of each species, detailed in Table II

Images			
Network	Batch size	Single crop	Multiple crops
resnet18	8	85.00	92.93
	16	85.89	93.16
	32	85.87	93.23
	64	86.03	93.04
resnet34	8	85.24	93.15
	16	86.75	93.50
	32	87.04	93.88
	64	86.96	93.43

TABLE III

CLASSIFICATION ACCURACY ON THE *individual image* SCENARIO, FOR DIFFERENT BATCH SIZES. SINGLE CROP AND MULTIPLE CROPS RESULTS ARE REPORTED

expected. Fig. 4 displays two examples of classification using the multiple tiled crops, showing the spatial distribution of the classification. It also displays the ID label for each crop.

Fig. 5 shows the average confusion matrix of our multiple crops voting on individual image experiments using a resnet34 and a batch size of 32. As one may suspect, trees from the same family are more difficult to differentiate. For instance, *Betula papyrifera* and *Betula alleghaniensis* as well as *Acer rubrum* and *Acer saccharum* are often confused with one another. Fig. 5 also shows some other difficult combinations, such as *Fraxinus americana* and *Acer saccharum*.

B. Test results when using all images of a tree

We were interested in seeing if the use of images taken at several different locations along the trunk would improve the classification results. We thus performed majority voting across all of the images of a given tree, both for single and multiple crops per pictures. Note that the number of available images per tree was variable, as stated in Section III-B. Table IV contains the results of this evaluation, again for a number of batch sizes. The results indicate that we are able to further improve the classification results (97.81%). More interestingly, we did not see any real difference between using a single or multiple crops in each image. This seems

Trees			
Network	Batch size	Single crop	Multiple crops
resnet18	8	97.10	97.51
	16	97.21	97.50
	32	96.92	97.21
	64	97.31	97.70
resnet34	8	97.81	97.22
	16	97.50	97.11
	32	97.50	97.51
	64	97.30	97.31

TABLE IV

CLASSIFICATION ACCURACY ON *all images per tree* SCENARIO, FOR DIFFERENT BATCH SIZES. SINGLE CROP AND MULTIPLE CROPS RESULTS ARE REPORTED

to indicate that having a greater variety of locations along a trunk is more beneficial than having a large number of crops that are closely located. This can probably be explained by anecdotal observations in the field, where we noticed that the bark appearance changed significantly from one trunk region to another.

C. Effect of dataset size on training performance

A common question arising when developing new classifier systems is: how much data do we need for training purposes? To answer this, we empirically evaluated the impact of the size of the training dataset on the classification accuracy. Moreover, we performed this evaluation for two cases that are particular to our classification problem: *a*) reduced number of images and *b*) reduced number of individual trees. To accomplish this, one fold from the previous experiment in Section V-A was taken and 9 smaller training datasets were created from the training set per case. For case *a*), we randomly sampled images from the training set until we hit a target goal of images. For case *b*), instead of sampling the images, we sampled the individual trees directly until we hit a target number of trees. Fig. 6 shows the results obtained. Note that we used the same testing set in both cases.

As can be seen, the general trend is that an increase in the number of images for the training leads to better results. However, the network is much more sensitive to the number of trees in the training dataset, rather than to the overall number of pictures. Indeed, when the number of overall images is randomly reduced by 90%, only about 5% of accuracy is lost. On the other hand, when the number of trees is randomly reduced by 90%, the results fall by more than 30%. This indicates that it is much more important to collect training data over a large number of trees, rather than taking a large number of pictures per tree. In other words, only a fairly limited number of pictures per tree are required to obtain a good performance.

VI. CONCLUSION AND FUTURE WORK

In this paper, we have empirically demonstrated the ability for ResNets to perform tree species identification from the pictures of bark, for 20 Canadian species. On our collected

1 - <i>Abies balsamea</i> - Balsam fir	169	0	0	1	0	0	0	0	0	0	1	7	0	0	0	0	4	0	0	0	92.86 %	
3 - <i>Acer rubrum</i> - Red maple	0	276	32	0	0	2	4	0	2	0	0	1	0	0	1	2	4	0	1	1	84.66 %	
4 - <i>Acer saccharum</i> - Sugar maple	0	15	364	0	0	0	13	0	0	0	0	0	0	0	0	0	6	0	0	0	91.46 %	
5 - <i>Betula alleghaniensis</i> - Yellow birch	3	1	0	230	13	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	92.74 %	
6 - <i>Betula papyrifera</i> - White birch	0	0	0	9	244	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	95.69 %	
7 - <i>Fagus grandifolia</i> - American beech	0	2	1	0	0	164	0	0	0	0	1	0	0	0	0	0	0	0	0	0	97.62 %	
8 - <i>Fraxinus americana</i> - White ash	0	7	5	0	0	0	273	0	1	0	0	3	0	0	0	0	2	0	0	1	93.49 %	
9 - <i>Larix laricina</i> - Tamarack	0	0	0	0	0	0	0	377	0	0	0	1	0	0	0	0	0	0	0	0	99.74 %	
10 - <i>Ostrya virginiana</i> - American hophornbeam	0	4	1	0	0	0	0	0	110	0	0	0	0	0	0	0	0	1	0	5	90.91 %	
11 - <i>Picea abies</i> - Norway spruce	0	0	0	0	0	0	0	0	0	258	4	0	0	2	0	0	0	0	0	0	97.73 %	
12 - <i>Picea glauca</i> - White spruce	0	0	0	0	0	0	0	0	6	113	0	0	0	0	0	0	0	0	0	0	94.96 %	
13 - <i>Picea mariana</i> - Black spruce	1	0	0	0	0	0	0	0	0	0	175	1	0	0	0	0	0	0	0	0	98.87 %	
14 - <i>Picea rubens</i> - Red spruce	0	0	0	1	0	0	0	9	0	2	0	2	132	0	1	0	0	0	0	0	89.80 %	
16 - <i>Pinus resinosa</i> - Red pine	0	0	0	0	0	0	0	0	0	1	0	0	0	112	4	0	0	0	0	0	95.73 %	
17 - <i>Pinus strobus</i> - Eastern white pine	7	0	0	1	0	0	0	0	1	0	0	0	0	1	190	0	1	0	1	0	94.06 %	
19 - <i>Populus tremuloides</i> - Quaking aspen	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	205	0	0	0	1	99.51 %	
20 - <i>Quercus rubra</i> - Northern red oak	1	4	20	0	0	0	1	0	0	0	0	1	0	0	1	4	510	0	1	0	93.92 %	
21 - <i>Thuja occidentalis</i> - Northern white cedar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	146	0	0	100.00 %	
22 - <i>Tsuga canadensis</i> - Eastern Hemlock	0	2	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	190	1	96.94 %	
23 - <i>Ulmus americana</i> - American elm	0	7	1	0	0	0	1	0	14	0	0	0	0	0	2	1	1	0	0	130	82.80 %	
		1	3	4	5	6	7	8	9	10	11	12	13	14	16	17	19	20	21	22	23	

Fig. 5. Average confusion matrix for multiple crops voting on whole images using a `resnet34` and a batch size of 32

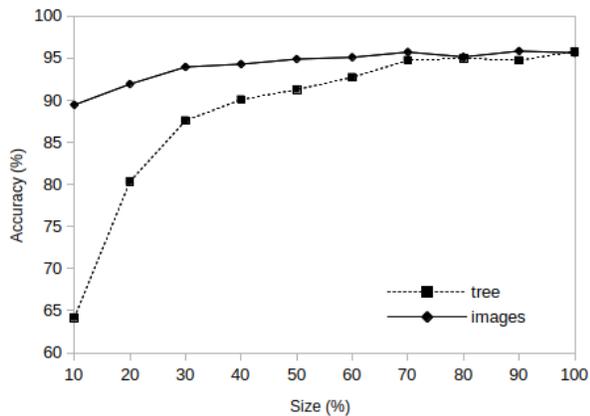


Fig. 6. Results obtained with a `resnet34` when only using a smaller percentage of *a*) images or *b*) trees from the dataset with multiple crops per image

dataset, the accuracy of the method ranges from 93.88% (for multiple crops on a single image) to 97.81% (using all trunk images), far above the 5% chance classification. We have found empirically that training is significantly more susceptible to the number of trees in our database rather than the overall number of images. This result will help tailor further data gathering efforts on our side.

In the process, we have also created a large public dataset (named *BarkNet* 1.0) containing labeled images of tree barks. This database can be used to accelerate research on bark classification for robotics or forestry applications. It can

also contribute in helping the computer vision community develop algorithms on the challenging problems of fine-grained texture classification.

Nevertheless, more work is needed to adapt the architecture of the network specifically to this task. As future work, we aim to leverage the DBH into a multi-task approach [33]. The use of multi-scale classifications will also be studied in an effort to determine the optimal scale at which to perform bark image classification. Moreover, we will explore the use of novel deep architectures that have been tailored to texture classification. We also plan on testing the approach on a sawmill floor, where we will have access to thousands of logs for data gathering. A new challenge will be to ensure that damages to bark due to logging operations do not adversely affect classification performances.

ACKNOWLEDGMENT

The authors would like to thank Luca Gabriel Serban and Martin Robert for their help in creating this dataset.

REFERENCES

- [1] F. T. Ramos, J. Nieto, and H. F. Durrant-Whyte, "Recognising and modelling landmarks to close loops in outdoor slam," in *Proceedings 2007 IEEE International Conference on Robotics and Automation*, April 2007, pp. 2036–2041.
- [2] N. Atanasov, M. Zhu, K. Daniilidis, and G. J. Pappas, "Localization from semantic observations via the matrix permanent," *The International Journal of Robotics Research*, vol. 35, no. 1-3, pp. 73–99, 2016.
- [3] A. Ghasemi Toudeshki, F. Shamshirdar, and R. Vaughan, "UAV Visual Teach and Repeat Using Only Semantic Object Features," *ArXiv e-prints*, Jan. 2018.
- [4] N. Smolyanskiy, A. Kamenev, J. Smith, and S. Birchfield, "Toward low-flying autonomous MAV trail navigation using deep neural networks for environmental awareness," *CoRR*, 2017.

- [5] T. Hellström, P. Lärkeryd, T. Nordfjell, and O. Ringdahl, "Autonomous forest vehicles: Historic, envisioned, and state-of-the-art," *International Journal of Forest Engineering*, vol. 20, no. 1, 2009.
- [6] S. Fiel and R. Sablatnig, "Automated Identification of Tree Species from Images of the Bark, Leaves and Needles," *Proceedings of the 16th Computer Vision Winter Workshop*, pp. 67–74, 2011.
- [7] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 11-18-Dece, 2016, pp. 1026–1034.
- [8] Z.-k. Huang, D.-S. Huang, J.-X. Du, Z.-h. Quan, and S.-B. Gua, "Bark Classification Based on Contourlet Filter Features," in *Intelligent Computing*, pp. 1121–1126, 2006.
- [9] Z. Chi, L. Houqiang, and W. Chao, "Plant species recognition based on bark patterns using novel Gabor filter banks," in *International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003*, vol. 2, dec 2003, pp. 1035–1038 Vol.2.
- [10] S. Boudra, I. Yahiaoui, and A. Behloul, "A comparison of multi-scale local binary pattern variants for bark image retrieval," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, vol. 9386, pp. 764–775.
- [11] M. Sulc, "Tree Identification from Images," 2014.
- [12] Y. Zhang and Q. Yang, "A Survey on Multi-Task Learning," *ArXiv e-prints*, July 2017.
- [13] M. Sulc and J. Matas, "Kernel-mapped histograms of multi-scale lbps for tree bark recognition," in *Image and Vision Computing New Zealand (IVCNZ), 2013 28th International Conference of*. IEEE, 2013, pp. 82–87.
- [14] A. Bressane, J. A. F. Roveda, and A. C. G. Martins, "Statistical analysis of texture in trunk images for biometric identification of tree species," *Environmental Monitoring and Assessment*, vol. 187, no. 4, 2015.
- [15] A. A. Othmani, C. Jiang, N. Lomenie, J. M. Favreau, A. Piboule, and L. F. C. L. Y. Voon, "A novel Computer-Aided Tree Species Identification method based on Burst Wind Segmentation of 3D bark textures," *Machine Vision and Applications*, vol. 27, no. 5, pp. 751–766, 2016.
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*. Curran Associates, Inc., 2012, pp. 1097–1105.
- [17] J. Champ, T. Lorieul, M. Servajean, and A. Joly, "A comparative study of fine-grained classification methods in the context of the LifeCLEF plant identification challenge 2015," in *CEUR Workshop Proceedings*, vol. 1391, 2015.
- [18] M. Šulc, D. Mishkin, and J. Matas, "Very deep residual networks with maxout for plant identification in the wild," *Working notes of CLEF*, 2016.
- [19] N. Sunderhauf, C. McCool, B. Upcroft, and P. Tristan, "Fine-grained plant classification using convolutional neural networks for feature extraction," *Working notes of CLEF 2014 conference*, pp. 756–762, 2014.
- [20] H. Goëau, P. Bonnet, and A. Joly, "Plant identification based on noisy web data: the amazing performance of deep learning (LifeCLEF 2017)," *CLEF working notes*, vol. 2017, 2017.
- [21] S. H. Lee, C. S. Chan, S. J. Mayo, and P. Remagnino, "How deep learning extracts and learns leaf features for plant classification," *Pattern Recognition*, vol. 71, pp. 1–13, 2017.
- [22] T. Mizoguchi, A. Ishii, H. Nakamura, T. Inoue, and H. Takamatsu, "Lidar-based individual tree species classification using convolutional neural network," *Proc.SPIE*, vol. 10332, pp. 10 332 – 10 332 – 7, 2017.
- [23] M. Cimpoi, S. Maji, and A. Vedaldi, "Deep Filter Banks for Texture Recognition and Segmentation," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, jun 2015.
- [24] L. Sharan, R. Rosenholtz, and E. Adelson, "Material perception: What can you see in a brief glance?" *Journal of Vision*, vol. 9, no. 8, pp. 784–784, Aug 2009.
- [25] D. Marcos, M. Volpi, and D. Tuia, "Learning rotation invariant convolutional filters for texture classification," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, dec 2016, pp. 2012–2017.
- [26] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, and S. Huovinen, "Outex - new framework for empirical evaluation of texture analysis algorithms." 2002, proc. 16th International Conference on Pattern Recognition, Quebec, Canada, 1:701 - 706.
- [27] M. Švab, "Computer-vision-based tree trunk recognition," 2014.
- [28] L. J. Blaanco, C. M. Travieso, J. M. Quinteiro, P. V. Hernandez, M. K. Dutta, and A. Singh, "A bark recognition algorithm for plant classification using a least square support vector machine," in *2016 Ninth International Conference on Contemporary Computing (IC3)*, aug 2016, pp. 1–5.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 770–778.
- [30] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.
- [31] L. Shen, Z. Lin, and Q. Huang, "Relay backpropagation for effective learning of deep convolutional neural networks," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 467–482.
- [32] J. Stanislaw, Z. Kenton, D. Arpit, N. Ballas, A. Fischer, Y. Bengio, and A. Storkey, "Finding flatter minima with sgd," in *ICLR Workshop*, 2018.
- [33] L. Trottier, P. Giguère, and B. Chaib-draa, "Multi-Task Learning by Deep Collaboration and Application in Facial Landmark Detection," *ArXiv e-prints*, Oct. 2017.