Please do not remove this page

An analysis of the accuracy of photo-based plant identification applications on 55 tree species

Schmidt, Ryan J; Casario, Brianna M; Zipse, Pamela C; et.al.

https://scholarship.libraries.rutgers.edu/esploro/outputs/acceptedManuscript/An-analysis-of-the-accuracy-of/991031655349704646/filesAndLinks?index=0

Schmidt, R. J., Casario, B. M., Zipse, P. C., & Grabosky, J. C. (2022). An analysis of the accuracy of photo-based plant identification applications on 55 tree species. In Arboriculture & urban forestry (Vol. 48). Rutgers University. https://doi.org/10.48044/jauf.2022.003
Document Version: Accepted Manuscript (AM)

Published Version: https://doi.org/10.48044/jauf.2022.003

This work is protected by copyright. You are free to use this resource, with proper attribution, for research and educational purposes. Other uses, such as reproduction or publication, may require the permission of the copyright holder.
Downloaded On 2024/04/28 19:00:34 -0400

An Analysis of the Accuracy of Photo-Based Plant Identification Applications on 55 Tree Species

Ryan J Schmidt[1,2], Brianna M Casario AS[1,2], Pamela C Zipse BS[1,2], Jason C Grabosky PhD[1,2,3]

Author Institutional Information

[1] Rutgers Urban Forestry Program, Rutgers, State University of New Jersey, New Brunswick, New Jersey

[2] Department of Ecology, Evolution and Natural Resources, Rutgers, State University of New Jersey, New Brunswick, New Jersey

[3] Corresponding Author

Corresponding Author Information

Jason Grabosky

14 College Farm Road

New Brunswick, NJ 08901

848-932-0050

grabosky@sebs.rutgers.edu

Abstract

Background With the creation of photo-based plant identification applications (apps), the ability to attain basic identifications of plants in the field is seemingly available to anyone who has access to a smartphone. The use of such apps as an educational tool for students and a major identification resource as for some community science projects calls into question the accuracy of the identifications they provide. We created a study to provide some context with our local tree species to provide an informed response to students asking for guidance in choosing a tool for their support in classes.

Methods Six mobile plant identification apps were tested on a set of 440 photographs representing the leaves and bark of 55 tree species common to the state of New Jersey.

Results Of the six apps tested, PictureThis was the most accurate followed by iNaturalist with PlantSnap failing to offer consistently accurate identifications. Overall, these apps are much more accurate in identifying leaf photos as compared to bark photos, and while these apps offer consistently accurate identifications to the genus-level, there seems to be little accuracy in successfully identifying photos to the species level.

Conclusions Therefore, while these apps cannot replace traditional field identification for these trees, they can be used with high confidence as a tool to assist inexperienced or unsure arborists, foresters or ecologists by helping to refine the pool of possible species for further identification.

Keywords
image recognition, tree identification, dendrology, botany, natural resource management

Introduction

With the creation of photo-based plant identification applications (apps), the ability to attain basic identifications of plants in the field is no longer limited to trained botanists or studied naturalists and is seemingly available to anyone who has access to a smartphone. This presents an incredible opportunity to engage young and emerging natural scientists, particularly in community science projects, where users can upload a picture of an unknown plant and receive a suggested identification from one of these mobile apps (Joly et al. 2014, Bilyk et al. 2020, Barre et al. 2017). While the accuracy of such cellular phone apps is not inherently imperative for casual botanical observations, the use of such apps as the sole or, at least, major identification resource for community science projects calls into question the accuracy of the identifications provided by these apps (Bonney et al. 2009,). We initiated a study to explore and evaluate a series of apps as a tool for educational training, as a supportive resource for early professionals in botanic fields and as a useful resource in volunteer training or resident engagement (Echeverria et al. 2021, Perdigones et al. 2021,Bilyk et al. 2020, Barre et al. 2017, Crall et al. 2011). We sought to provide some context with our local urban and rural tree species to provide an informed response to students asking for guidance in choosing a tool for their support in classes.

In urban tree inventories, the proper identification of trees is crucial in terms of understanding the implications, benefits and risks associated with the urban forest from a management perspective. Similarly, understanding the species composition within an area can lend insight into the ecological effects of trees on the community as a whole. These discussions on tree community structure, diversity, and resilience within an urban forest or landscape rely upon identifying the species in place.

While accurate identification of trees is fundamental to community assessment, the precision to which they need to be identified for sufficient understanding will likely be different depending on the goals or use of the identification information. For example, the identification of *Fraxinus* species to genus might be acceptable in order to determine which trees are susceptible to infection from Emerald Ash Borer (*Agrilus planipennis*) while identifying maples to species might be crucial to understanding a specific tree's susceptibility to storm damage, drawing distinctions between the sturdy *Acer saccharum* and the weak-wooded *Acer saccharinum*.

In terms of ecology, as each species has a specific set of preferred environmental conditions, understanding the species distribution within an area can help to attain a better working knowledge of the intricacies of the system being studied (Trowbridge and Bassuk 2004, Robichaud and Buell 1973),   In a natural setting, the linkage between site conditions and species distribution helps to illuminate trends in hydrology and soil types across a community and by applying these ideas to urban settings, understanding the disconnect between site conditions and species selection (Trowbridge and Bassuk 2004,   can be used to guide disease and pest management decisions as well as future planting stock selections (Laćan I. and McBride J.R. 2008 Scharenbroch et al. 2017).

A thorough knowledge of tree identification is needed to provide the plant-community inventory prior to making a site management plan or gaining an understanding of plant community-site relationships. There is growing evidence that volunteers can produce valid data streams in generating urban community inventories, particularly at the genus level (Bancks et al. 2018.) with the associated community stewardship benefits that come with citizen science engagement (Crown et al. 2018, Roman et al. 2017). To this end, community volunteers with varied levels of background training and—more generally—less experienced botanists and tree care professionals may use apps which offer help in identifying plants while in the field or at home from captured field images.

To use the typical app, the observer simply needs to take a close-up photograph of the tree (most frequently of the leaf, bark, flower, or fruit) and upload it to the app. Once uploaded, some apps prompt the user to specify the character being tested (again usually either the leaf, bark, flower, or fruit) and then the app will compare the user's photograph to photographs within its system (Joly et al. 2014, Bilyk et al. 2020, Barre et al. 2017). The output is a listing of one or more suggestions as to what the identity of the plant may be. The first listed suggestion is viewed as the primary identification for the plant and is henceforth referred to as the "Identification". Many apps provide additional suggestions for the identity of the plant (henceforth referred to as simply "Suggestions") in order to allow for some error in the primary identification. For a thorough review of the development and logic of plant identification apps, please refer to Wäldchen and Näder (2018).

Although these apps are often considered to be extremely helpful in species identification, there has been little done to compare the identification precision and accuracy of these apps as a whole, and we sought to inform our conversations with students, community volunteer groups, and beginning professionals. The lack of information beyond the details and claims produced by the developer reflects the difficulty in direct comparison in a technical sense. A challenge as detailed by Xing et al. (2020), the systems do not share data sets, system training approaches, common flora or focal plant organ, much less a comparable user interface (Keivani et al. 2020, Goëau et al. 2013, Wang et al. 2013, Cope et al. 2012, Kumar et al. 2012). Apps generally are developed in a machine-learning environment with improved function as additional data is accumulated, as an evolving "intelligence", as an algorithm using a probability-based neural network in some form. Such derived code can be pressed against open-sources image sets such as Flavia (Wu et al. 2007) and the Folio dataset (Munisami et al. 2015) which can then be automated into an image analysis as was developed by Keivani et al. (2020). Additional data sets have been used elsewhere, such as the Swedish Leaf dataset (Söderkvist 2001), or the LeafSnap image libraries used by Kumar et al. (2012). Generally speaking, the resultant code calibration yields incredibly high accuracy results, often exceeding 95% (Keivanai et al. 2020, Goëau et al. 2013, Kumar et al. 2012, Wang et al. 2013). Such accuracy cannot be assumed to predict the efficacy of the tools once beyond the code training environment, but accuracy claims would certainly flow from the initial training phase. Our study uses the tools beyond this training phase, specific to our limited purpose with non-curated field images. Our protocol to standardize and avoid extraneous non-target information was chosen to avoid deflation of accuracy due to the photo quality.

We set out to determine the accuracy of six of the most-downloaded apps (as per the Apple App Store® at the start of the project, 6 July 2020) in order to better understand what trends exist in the apps' cumulative abilities to identify different groups of trees to the genus and species level: iNaturalist™, Pl@ntNet™ (henceforth PlantNet), LeafSnap™, PlantSnap™, PictureThis™, and Plant Identify™ (Table 1). Our choice in selection by popularity stands in contrast to a similar study conducted by Xing et al. (2020) which selected apps based on function (foliar versus floral identification). As that study points out, performance within urban forests need to be checked since the species profiles are different between the locally natural and the designed plant community (Xing et al. 2020). Our evaluations were organized based on phylogenetic relatedness (i.e., trends within and between taxa) as well as morphological traits of the leaves and bark across all of the apps. We sought to understand the accuracy of each of the apps individually and determine how accurately the apps can identify trees from pictures of their leaves as opposed to their barks. We then considered their value as a teaching support for students or for early field professionals. The study differs from other work which usually considers a broader range of plant types (beyond trees) within a regional flora (Jones 2020,Kumar et al. 2012). We focused on in-field identification rather than identification from stock photos, such as used in Jones (2020) which considered several of the same apps, but on a wider range of plant types (i.e., multiple habits), and only two species within the genera of trees that we tested (i.e., *Quercus robur* and *Acer* *pseudoplatanus*). There have also been studies evaluating the identification ability of similar photo-based identification software as compared to the ability of botanists (of varying levels of experience) to identify the same photos. Bonnet et al. (2015) determined that while the apps did not come close to outperforming expert botanists, their identification skills were on par with somewhat-experienced botanists and even outperformed inexperienced botanists, indicating that these apps may have profound implications if they can be tactfully utilized by beginners in the field.

TABLE 1 HERE (Apps Considered in Study)

Materials and Methods

Study system

Our study system was the temperate seasonal climate of the State of New Jersey is located within the mid-Atlantic Region of the United States and can be described as spanning 5 physiographic regions from the Highlands and Ridge and Valley systems in the northwest through a Piedmont section and to the inner- and outter-coastal planes in the southeast (Robichaud and Buell 1973). The northern half of the state is dominated by glacial actions along the Appalachian Rib and a transition from an upland forest with mixed hardwood assemblages of Maple-Beech-Birch to a mixed oak-hickory forest. The southern half is dominated by non-glaciated sands, encompassing the New Jersey Pine Barrens and plant communities akin to the southeast oak-pine and south east bottomland systems (Collins and Anderson 1994, Tedrow 1986, Robichaud and Buell 1973). As a heavily urbanized state within the Northeast Megalopolis, there are many introduced tree species representative of design preferences and choices from over 250 years of development. Average minimum temperature hardiness zone is listed as ranging from a -23.3 to -20.6 C zone in the north to a -15 to -12.2 C zone in the southern coastal and more urbanized areas proximate to New York City, NY and Philadelphia, PA (USDA 2012).

Species Selection

To attain a general idea of the overall accuracy of the apps in terms of trees found in New Jersey forests and landscapes, a wide range of 55 species were selected for analysis (Table 2). Species were selected due to their prevalence within the state of New Jersey as common street or forest trees for both practicality and usefulness. The list included both native and introduced species. In terms of forest trees, the largest portion of the forests within New Jersey are categorized as Oak/Hickory forests with large percentages of Loblolly/Shortleaf, Oak/Pine, Northern Hardwood and Elm/Ash/Red Maple forests (Widmann 2005, Crocker et al. 2017). Therefore, several species of Oaks, Hickories, Pines, Maples and Birches were included to attempt to represent some of the more likely species that would be encountered in the forests around the state.

TABLE 2 HERE (Species Table)

Species such as the *Magnolia* spp., *Gleditsia triacanthos*, *Zelkova serrata*, *Platanus* spp., *Tilia* spp., and *Pyrus calleryana* were included due to their high prevalence as street and ornamental trees (Sanders et al. 2013). Due to the often very similar characteristics of the different subspecies and cultivars, no effort was made to distinguish them from one another and an identification to the species was all that was required (e.g., *Gleditsia triacanthos* subsp. *triacanthos* and *Gleditsia triacanthos* var. *inermis* were both treated simply as *Gleditsia triacanthos*). Cultivars and infraspecifics with extremely divergent leaf or bark characteristics (e.g., *Acer platanoides* 'Crimson King') were excluded from this study.

Additional species were chosen to increase both morphological and phylogenetic diversity amongst the testing specimens. For example, species such as *Salix babylonica*, *Ginkgo biloba*, *Taxodium distichum* and *Aesculus hippocastanum* were selected due to their leaf morphologies to expand the evaluation range of the study. Common invasive species such as *Ailanthus altissima* were included as they are often targeted for specific studies that seek to better understand the prevalence and distribution of invasive species within an area as well as for management efforts to control or eradicate them. Finally, the species *Castanea dentata* and *Nyssa sylvatica* were added after the beginning of the study due to the frequency that they were incorrectly suggested by the apps. Several of the planned test species were misidentified as these two taxa (10 times as *C. dentata* and 27 times as *N. sylvatica*). We thus included these less-common species to determine if they would be correctly identified when presented with images of the species in the field given their frequency as an incorrect suggestion for other species.

Photo Collection

For each of the species represented in the study, a minimum of four photos each of bark and leaves were taken from different individuals of the same species so that no two photos of a single character were taken from a single tree (a bark photo and a leaf photo from the same tree was however permissible). As the team collected images, photos from several individuals were collected and then aggregated into folders for the targeted species. Then, 4 leaf and 4 bark images were selected for each of the species being studied. When possible, leaves and bark without noticeable infection or infestation were selected (Cherry Leaf Spot, *Blumeriella jaapii* (Rehm) Arx, was not feasible to exclude in *Prunus serotina*).

Efforts were made when possible to attain photographs representing the phenotypic variation present in the species in terms of morphology and tree age. For example, bark photos of young, mature and old trees were included when possible, and for trees with multiple leaf shapes (e.g., *Sassafras albidum*) representatives of each leaf type were included. When possible, photos of each species from different locations were included in order to attempt to account for some of the ecotypic variation in the species (e.g., *Pinus rigida* from the Pitch Pine/Scrub Oak forests of North Jersey and the Pitch Pine forests of South Jersey). The majority of these photos were taken in Mahlon Dickerson Reservation in Morris County and on the Rutgers University-Cook/Douglass campus in New Brunswick as well as in Medford, Moorestown, and Pennsauken, New Jersey.

All of the photos used in this study were taken by authors of this paper, the vast majority of which were collected in the month of July 2020. Phenotypic variation between the photos of each species is therefore minimal due to the limited time of year they were collected. Photos were collected using the built-in cameras on either the Apple iPhone XS®, iPhone 11® as a 12-megapixel image, or a Samsung Galaxy S9® as a 12-megapixel image as well as a small number from a Nikon 3100 digital camera as a 13.5-megapixel image . Bark photos were taken so that the only character visible in the frame was the bark whenever possible (i.e., avoiding leaves, fruits, and epicormic sprouts). Some space was left to the sides of the tree so that the whole trunk section could be viewed. The 'zoom' feature was avoided when at all possible in order to assure that the photo would not be distorted. Leaf photos were taken so that there would be one leaf (or possibly a few if the leaves were smaller) centered and focused in the frame with the natural surroundings around it. Efforts were made to exclude fruit and bark from the photos to ensure that they were identifying from the leaf alone. Epicormic sprouts were avoided when possible as their form is often divergent from the typical canopy leaf.

Data Collection

Four bark photos and four leaf photos of each species were selected according to the above criteria and uploaded individually to each of the apps. For the sake of consistency, the photos were merely uploaded to the app and allowed to crop and focus on their own without any interference or the moving of frames. All photos were uploaded to a digital storage folder and then re-downloaded before uploading them to any of the apps so that there was no GPS data associated with the images. All apps were provided the same set of images and all photos were uploaded to the apps within the state of New Jersey. Once a photo was uploaded, each app typically offered one or more guesses (an identification was not always made by PictureThis and Plant Identify) as to the identity of the plant. These identifications and suggestions were given in the form of a species name with a generic name and specific epithet (e.g., *Acer rubrum*). For this study, only automated, or system-generated suggestions for plant identification were used. We did not consider the community aspects of some apps, wherein suggestions from experts or other users could have also been considered, negating an important supplemental aspect which is available in some apps (e.g., PlantNet, PlantSnap, and iNaturalist).

In order to determine the accuracy of these identifications and suggestions to both the genus- and species-levels, we coded the responses by breaking the app suggestion into the genus and then the specific epithet components to segregate correct genus level identifications. We then recorded separately if the app correctly identified the plant's genus *and* specific epithet. For clarity, and since completely different species can share the same specific epithet (e.g., '*americana'* in *Ulmus americana* and *Tilia americana*), the specific epithet identification/suggestion was not used, in isolation. The results were interpreted and recorded as follows:

- Genus Identification: If the tree was identified correctly to the genus in the first suggestion, it received a score of '1' for the Genus Identification. If it was not, it received a score of '0'.
- Species Identification: If the tree was identified correctly to the species in the first suggestion, it received a score of '1' for the Species Identification. If it was not, it received a score of '0'.
    - If the tree was identified to one of the hybrids of the correct species in the first suggestion (or identified as a parent of a tested hybrid), it received a score of '0.5' for Species Identification.
- Suggested Genus/Genera: If the tree was identified correctly to the genus in the first OR any other suggestion, it received a score of '1' for the Suggested Genus. If it was not, it received a score of '0'.
- Suggested Species: If the tree was identified correctly to the species in the first OR any other suggestion, it received a score of '1' for the Suggested Species. If it was not, it received a score of '0'.

    - If the tree was identified to a hybrid of the correct species in any suggestion (or identified as a parent of a tested hybrid), it received a score of '0.5' for Suggested Species.
    - If the tree was identified to *more than* one hybrid of the correct species in any suggestion (or identified as *both* parents of a tested hybrid), it received a score of '1' for Suggested Species.
- If the tree was misidentified in the first suggestion, the first proposed species was recorded.

Data was tabulated as the percentage of correct identification or suggestion across each species bark and leaf set, or across classification or app groupings. We arbitrarily defined evaluation categories of high, moderate and low confidence (95-100% correct, 80-94% correct, and <80% correct respectively). Data were developed and processed from the July 2020 photo collection through the following 50 days, so any inferences from the apps that were chosen are based on their program and algorithm development as of summer 2020. In order to ensure that the data collected would be consistent through multiple runs, all photos of four selected species (*Quercus alba*, *Betula lenta*, *Acer saccharinum*, and *Pinus rigida*) were run through all six apps for a second time several days after the first run, however before any updates were allowed to occur on any of the apps as this could have influenced the accuracy of the apps (Jones 2020). Then, a Chi Squared test was run in order to determine if there was a statistically significant difference between the outcomes of the multiple runs.

Finally, for interpretation of the results, species were categorized into groupings by bark characteristics as detailed in Wojtech (2011) to look for patterns in the app-response results: Peeling Horizontally, Lenticels Visible, Smooth Unbroken, Vertical Cracks or Seams in Otherwise Smooth Bark, Broken into Vertical Strips, Broken into Scales or Plates, or With Ridges and Furrows. For species with different bark types at different life stages or in different forms (e.g., the many bark types of *Acer rubrum*), the species was placed into each group (e.g., *Acer rubrum* being listed under Smooth

Unbroken, Vertical Cracks or Seams, and Vertical Strips). When a taxon was not explicitly mentioned within Wojtech (2011), species were categorized according to the text descriptions for each category.

Results

Chi-squared values of $\chi^2=0.1296$ and $\chi^2=0.0106$ were determined for identifications and suggestions, respectively. Their corresponding p values ($p=0.7188$ and $p=0.9179$, respectively) both fail to reject the null hypothesis that there is no difference in the accuracy of the apps' identifications of the same photographs on two different days with a significance level of 0.05.

TABLE 3 HERE (App Evaluation Table)

PlantSnap was able to correctly identify a comparable percentage of the tested leaf photos, however, the percentage of correct bark identifications was exceedingly low across all taxa. Due to the low levels of accuracy in the identification of North American trees by bark characters, the data collected from the PlantSnap app was excluded from consideration when looking for general trends across all apps as sorted by taxonomic order, family, genus or species (Tables 2,4).

Across all apps, leaf photos always outperformed bark photos by a large margin. In terms of bark images alone, none of the tested apps provided an overall accuracy of over 70% in identifications and none over 80% in overall suggestions. We observed a moderate confidence in Genus Identifications for leaf photos across our selected taxa in all but two cases: PlantSnap provided a low confidence and PictureThis provided a high confidence. Species Identifications for leaf images across all taxa where only moderately confident for PictureThis, and showed low confidence for all other apps tested. For Genus Suggestions and Species Suggestions, scores generally increased across all apps (some to a greater extent than others), excepting PictureThis which does not usually provide suggestions, beyond the initial identification. The only exception to this was for one leaf photo of *Taxodium distichum* which it misidentified *Taxodium mucronatum*. When uploaded to PictureThis, the app indicated that this tree was similar to *T. distichum* and that "it is not easy to distinguish them with just one photo", much like the suggestion descriptions on other apps. The iNaturalist app was observed to suggest the correct species 95.91% of the time for leaf photos, which is indicative of high confidence that one can narrow an observation to at least a correct species complex, if not a singular species. PictureThis failed to offer an identification of one image (bark of *Pseudotsuga menziesii*), while Plant Identify failed to make an identification in 47 of the uploaded images.

TABLE 4 HERE (Taxonomic Table)

Across all of the taxa studied, the apps were more accurate in identifying trees to the genus-level as opposed to the species-level (Tables 2,4). For each group except for the Betulaceae (including the genus *Betula*), the leaf photos had dramatically higher correct identification rates than the bark photos. While some taxa exhibited moderately confident, [80%-94%), species-level identifications of leaf photos (namely the Cupressaceae, Fabaceae, and Sapindaceae and the genera *Acer* and *Picea*), all species-level bark identifications were wholly unreliable.

The apps consistently offered correct leaf identifications to the genus-level for some genera (namely *Acer*, *Carya*, *Picea*, *Platanus*, *Quercus*, and *Tilia*) with an accuracy of 95% or above. However, the apps all failed to offer consistently accurate identifications for any of the *Magnolia* spp. for either bark or leaf photos (5.00% and 37.50%, respectively).

Many of the same points made above for the broader taxonomic divisions (Table 4) can also be seen exemplified at the species level (Table 2). Again, genus-level identifications are much more reliable than species-level identifications, and besides the members of the Betulaceae and several unrelated species (namely *Fagus grandifolia*, *Pinus sylvestris*, and *Platanus* x *hispanica*), bark remains mostly unreliable at any level.

Here it can be seen that there is a very high probability that the correct genus will be listed as either an identification or a suggestion for leaf photos (94.55% of species having a moderate-high confidence interval for genus-level leaf suggestions). In terms of identification of trees by bark, in spite of the much lower percent accuracy as compared to leaf identifications, there were some clear, trends that exist based on bark type. While most bark types exhibit a percent identification rate of less than 50%, there was a surprisingly high identification rate to genus for bark that is peeling horizontally (87.50%) and to a lesser extent bark with visible lenticels (69.44%). The high accuracy of *Betula* species is very likely linked to this observation.

Table 4 also illustrates the nuances between the identification rates of closely related taxa such as those of *Magnolia* spp. and *Liriodendron tulipifera*. While identification rate for the members of the Magnoliaceae in Table 4 can be seen to be very low (as would be expected due to the low percent accuracy for the *Magnolia* species), this table shows that *Liriodendron tulipifera* (also in the Magnoliaceae) had an impressive species-level identification rate of 100% for leaf photos. This helps to exemplify how species with more iconic characters may be more consistently identified correctly, even within typically underperforming taxa.

Discussion

We stress that this study was, by nature, limited in its scope (isolated to 55 species of trees commonly found in New Jersey urban and natural landscapes) and cannot be used as an accurate evaluation of these apps across all plant habits, taxa and morphologies. Therefore, it should be understood that the following observations are meant to guide users who are likely to encounter the same taxa in their activities. This survey also does not take into consideration the power of community and expert identifications available on some apps, (table 1): it only evaluates the suggestions given by the apps for immediate identification in the field. We acknowledge that the loss of a GPS coordinate may well influence output in some apps. The cosmopolitan species diversity of our regional urban plant community may negate the GPS value or it could influence the aptness of the tool and its success in a forest inventory when considering a choice. Indeed, Leafsnap was initially conceived and focuses on the tree species of the Northeastern Forest community (Kumar et al. 2012) but our study sample extended to species in southern New Jersey beyond that database. Furthermore, we chose the apps for this study based on their download frequency and availability of use. It is important to note that the apps are meant to engage larger aspects of the flora in total, and each app represents a different database which can range from thousands to hundreds of thousands of species as well as different algorithm learning trajectories for their own development for accuracy (table 1) These various apps host vastly different scales of species range and type, with iNaturalist spanning beyond the plant kingdom (including animals, fungi, and protists) as a community of experts and novices.

The Chi Square test suggested repeatability in the output for constancy of identifications and suggestions, but as observed in general, there are limits to what can be expected as a tool to aid in tree species identification. That said we fully expect that such outcomes would improve as any specific app evolves with increased data process. The fact that an experienced observer can locate and define multiple traits much faster than can be accomplished with a phone camera enforces the use of such tools in support in training and confirmation.

Across all apps, there was a general trend of higher percent accuracy in correctly identifying leaf photos as opposed to bark photographs. This is not surprising based since the process in developing such tools has focussed on image pattern recognition using shape, edge pattern, venation and similar characters consistent with foliar morphology (Keivani et al. 2020, Zhao C et al. 2015, Zhao Z et al. 2015, Goëau et al. 2013, Wang et al. 2014, Wang et al. 2013, Cope et al. 2012). Our result highlights the general difficulty of using bark characters alone for traditional tree identification due to the effects of convergent bark appearances across taxa as well as the effects of the environment on bark texture and qualities. For the identification of trees in forested areas (where twigs and leaves are not easily observed) and the identification of deciduous trees during the winter, the use of bark can become a very reliable characteristic deserving greater attention.

It would be reasonable to suggest that some ubiquitous species that have more data within an AI network, and those interesting species with iconic bark or leaf characters or aesthetically-charismatic leaf form would, in general, provide a higher confidence in either identifying or suggesting against a new image (e.g., the high genus-level identification rate for leaf photos of members of the Sapindaceae including the easily recognizable *Acer* and *Aesculus* leaves). For PlantSnap in particular, while the percent of correct leaf identifications was only slightly below the percentages for the other apps, the percent identifications for bark were exceedingly small with only 1.36% identification to genus and 0.00% identification to species.

The app with the highest percentage of correctly identified photographs was PictureThis with a combined leaf and bark correct-identification percentage of 81.36% to genus and 67.84% to species. This app also boasts a 97.27% identification rate to genus and an 83.86% identification rate to species for leaf photos as well as a 65.45% identification rate to genus and an 51.82% identification rate to species for bark photos. With such a high percent accuracy for identification of leaf photos to genus, we will likely suggest this app for our purposes with students if and when they feel they want to pay for such a tool as a confirmation to their own field

The PictureThis app always offered only one species identification for each photo upload in all but one taxon tested. That exception was with *Taxodium distichum* and *Taxodium mucronatum* which were listed as difficult to distinguish from photographs and misidentified one of 4 leaf images. PictureThis also failed to offer an identification for one photo: the bark of *Pseudotsuga menziesii*. While this might seem to be a drawback that the app might not make any identification at all, this failure to offer an identification when unsure indicates that when the app is not confident it will not make a potentially faulty identification. Ultimately, PictureThis, and arguably the iNaturalist app, offered identifications with a high confidence to genus that might be deployed in a number of practical approaches, particularly in early training or educational situations, or as an early support for emerging professionals.

For situations in which only a broad context of community is desired, identification to genus might be acceptable. For example, a study which seeks to determine the number of tree families or genera present in a patch of woods or a portion of a community might only require such identifications for useful data. Use of an app can help to attain large amounts of broad data in a short amount of time (and with inexperienced naturalists) which could then be refined by more experienced foresters as needed. This could be in the form of successively working through genera until all have been identified to a finer degree or to target desired genera for more specific detail.

These apps can also assist inexperienced or unsure arborists, foresters or ecologists who are not confident in their identifications by narrowing down their observations to the genus or species level. For example, the user could take a picture of the leaf of a palmately lobed tree to use the app to distinguish *Acer* from *Platanus* and *Liquidambar* with high confidence. The user could then utilize a more specific key or more refined section of a reference guide to distinguish between species within a genus. Such apps could be used by foresters or ecologists who simply want a second opinion on identifications to prevent potential consistent misidentifications. Apps could also be used as an educational tool in preparation for credentialling or licensure exams to practice leaf identifications to the genus.

PictureThis is, however, a paid app which may reduce its accessibility for those without the resources (or long-term need) to purchase the app. This therefore might make the standardized use of this app less probable, especially for students and volunteers. The investment might be worthwhile for beginning foresters or ecologists to help to validate their identifications and expose any biases they might have in their identifications.

As an alternative to a paid app, the second most accurate app, iNaturalist, offered many of the same values as PictureThis and includes some community-based assistance which can help to attain more confident identifications. iNaturalist had an observed 92.27% identification rate to genus and a 69.55% identification rate to species for leaf photos as well as a 48.18% identification rate to genus and an 31.82% identification rate to species for bark photos. With a percent leaf identification to genus of over 90.00%, iNaturalist can be used in a similar manner as stated for PictureThis, however, with only moderate confidence.

In contrast to the singular identification provided by Picture This, iNaturalist provided many suggestions as possible species. This can be useful for individuals with some knowledge of tree identification who can look through the list and reject some of the suggestions due to previous knowledge (e.g., rejecting trees with similar leaves that have widely different barks than the unknown specimen). This could result in a relatively short list of species to sort through and turn an almost unmanageable list of possibilities into one that can be used to quickly narrow the scope of a field guide, such as when guided as a "quest" (Kingsley and Grabner-Hagen 2015). iNaturalist offered higher level identifications if the software was confident in their identification such as being "pretty sure" for different families and genera. Of all of the photos which received a listing of "pretty sure" to a specific family, 90.03% of them were correct and identifications listed as "pretty sure" to a specific genus, 95.83% of the time across both bark and leaf pictures. Even when the identified broader taxon was not correct, there is a 99.17% chance that the correct genus will be listed in the suggestions and a 95.83% chance the correct species will be in the suggestions.

The iNaturalist app also utilizes community and expert verification on photos submitted through the app. Other apps such as Pl@nt Net, Plant Snap, also offer a community support function with their tool. While this may pose a challenge for large scale identification efforts such as comprehensive tree inventories, in smaller projects where time is not as much of a limiting factor and can help to ensure higher accuracy. A community support function can also be helpful to identify a species (or at least get a second opinion on a specimen) that is particularly difficult to identify or foreign to the naturalist or ecologist. As a teaching tool, the power of linking an interested person into a larger, more professionally adept community is an invaluable asset (Pollack et al. 2015). There is a potential for questing or gamification (Kingsley and Grabner-Hagen 2015) of early natural resource management or natural sciences students in the tactical use of these apps (Struwe et al. 2014).

In order to better understand the limitations of these apps and in turn how to best utilize them to attain the most confident data possible, we set out to explore the effects that different morphological features had on the ability of the apps to correctly identify a tree. Starting with the broadest morphological comparison, there seems to be a relatively small difference between the ability of the apps to successfully identify broadleaf species and needle/scale bearing species by leaf to genus (89.24% and 91.67% respectively) and to species (65.60% and 63.89% respectively). This is slightly surprising due to the apparent visual similarities between the leaves of needle-bearing trees. From a practical perspective, this could be a very important piece of information for community science projects and tree inventories as it is a common issue that many novices believe all needled-evergreen trees to belong to the genus *Pinus* (Bancks et al 2018). The use of these apps can help to ensure that needle-bearing trees can be more often identified correctly to at least the genus.

When just considering broadleaf species, there are several morphological characteristics that offer an important insight into the success of these apps. Across all runs, the apps seem to have a higher percent of correct identifications to the genus for trees with compound leaves than for simple leaves (96.00% as opposed to 87.36%). This is likely in large part due to the greater number of genera within the region containing a majority of simple leaves, as opposed to compound leaves.

In terms of the lobation of simple leaves, a similar trend seems to exist in regard to lobed leaves vs. unlobed leaves with the seemingly more numerous unlobed-genera having a lower percent correct identification than the lobed leaves. When, however, the type of lobation (palmately or pinnately) is distinguished, an interesting trend becomes apparent: when considering the identification of palmately lobed leaves there was a staggering 100% correct identification rate to genus. This is particularly important as we, the authors, find there to be a propensity for individuals new to tree identification misidentify *Platanus* species as *Acer* species and vice versa, unless there is a specific training emphasis in this area. This distinction was addressed by Roman et al. when completing a brief training session with beginner tree inventory volunteers, which resulted in a high level of accuracy (Roman et al. 2017). With such a high percentage of proper identifications for this leaf type, the use of these apps seems to offer the ability of even inexperienced naturalist to confidently distinguish genera of trees with palmately lobed leaves when a similar type of training is not feasible.

Taking a cue from the success of the apps with the bark of *Betula* species, it would be interesting to include only photos with bark containing visible lenticels (e.g., young *Prunus serotina*, *Pinus strobus*, and *Quercus rubra*) in order to determine if there is a correlation between bark with visible lenticels and a higher percent identification or if the *Betula* species are merely skewing the data. It is important to note that for deciduous species, a leafless condition or unreliable access to expanded leaves can occur in New Jersey from November to April, or 6 months of every year which can put more pressure on attempting to attain accurate identifications from bark (or bud) characters. Given the extremely low accuracy of these apps in identifying trees by bark images, however, such apps did not seem to offer an adequate solution to this problem at the time of our study. From a managerial perspective, this is an area in which targeted software development would greatly improve the apps' utility in the field.

The taxonomy of tree species has the potential to illuminate helpful trends in species characteristics that can help divide the list of possible species into more manageable groups. If a potential user is able to identify the taxonomic order, family or genus to which a particular specimen belongs, it can be very helpful to understand the reliability of the identification that the apps tend to provide. For example, if a tree with a nut is found, it can be predicted that a photograph of the leaves will correctly identify the tree to genus 87.81% of the time. If the tree can even be narrowed down to the walnut family or the beech family the confidence in correct identification to genus can increase even further to 100.00% and 96.25% for leaf photos, respectively. To take it even further, if a tree can be identified as an oak there is a 83.06% chance that the leaf can be used to correctly identify the taxonomic section to which it belongs (sections *Quercus* and *Lobatae* of the subgenus *Quercus* were tested). However, it is very likely that volunteer training could yield similar results without an app. (Bancks et al. 2018, Roman et al. 2017, Kosmala et al. 2016). While taxonomic section-level identification of the tree is often not specific enough to properly manage or understand the implications of the tree on the site (and conversely the site on the tree) it can help to sufficiently reduce the number of potential candidates and make further identification markedly easier. In addition to the Fagaceae and the Juglandaceae, trees in the Platanaceae, Sapindaceae, Malvaceae, and Cupressaceae all have a highly confident identification to genus (above 95.00%).

On the other end of this spectrum, it is important to note that certain taxonomic groups can be seen as chronic underperformers and therefore their identifications should not be inherently trusted. Species in the Betulaceae (and specifically the genus *Betula*) collectively have some of the lowest percent of identifications by leaf photos but conversely have one of the highest identification percentages by bark of 85.00%. The lowest percent accuracy determined through this study was in regard to the genus *Magnolia* which had only a 37.50% accuracy to genus with leaf photos and a meager 5.00% accuracy to genus with bark photos. While not inherently surprising given the difficulty even for trained foresters to distinguish *Magnolia* species without specific characters, it is clear that these apps do not seem to offer any reliability for this taxon in particular. This is likely due, in part, to the inability for any of the apps to utilize any other sensory characteristics in their identifications (e.g., the presence and quality of trichomes, smell of crushed leaves, and sound of snapping needles); all characters which are often relied on heavily in the training of professionals in the field.

The taxonomic groups listed in Table 2 were limited in order to attempt to ensure that the data would not be completely unrepresentative of the group. For instance, including percentages for a group such as the Lamiales for which our study only considered *Fraxinus* species would not be indicative of the apps' abilities to identify any species within the Lamiales but instead just indicate their ability to identify *Fraxinus* species: it is unknown whether the inclusion of species in the genera *Olea* and *Syringa* (also within the Lamiales) would have greatly changed the total percentages for the entire order. Similarly, all genera with only one tested species were excluded as the app's ability to identify one species is not necessarily indicative of its ability to identify another species within the same genus.

Some attention should also be paid to those species that were offered incorrectly as the primary identification very frequently throughout the study: *Carya glabra* (identified incorrectly 45 times), *Fraxinus americana* (39 times), *Betula pendula* (34 times)*, Liquidambar styraciflua* (32 times), and *Acer platanoides* (29 times) were all erroneously offered very frequently. While these misidentifications were mostly due to incorrect identifications of bark photos, it is important to understand which species are frequently suggested so that it is understood that even though some species might have extremely high correct identification percentages, not every identification can be trusted. For example, *Acer platanoides* has an impressive correct identification rate to species of 100.00% for leaf photos however 9 additional leaf photos (all of *Acer saccharum*) were incorrectly identified as *Acer platanoides*. The apps also frequently identified species that are not native to North America and are almost exclusively found in the planted landscape such as *Betula pendula* (34 times), *Carpinus betulus* (27 times), and *Quercus robur* (27 times) which can often be excluded quickly by form or site conditions if working in the natural landscape, especially those of European origins. Again, compared to earlier training studies such as Jones (2020) and the occurrences of *Q. robur* as suggestion, there is an artifact of training and a rationality to consider with choice of application which has to be balanced with the varied selections of urban landscapes. The unfortunate point to be made, however, is that we can make these observations from a vantage point of already possessing a positive identification before using the apps. The person needing or using the apps cannot be expected to know in such detail what to trust or avoid, otherwise they would not be likely to use the app in the first place (unless they were, for example, in a supervised training event with an expert to guide the process as a teaching tool).

Conclusion

For our purposes, the use of PictureThis would most likely offer the most accurate identifications for immediate responses to photo uploads from the field. This app could be considered if sufficient funds are available or the need for accuracy is of the utmost importance. If funds are limited, iNaturalist seems to be the closest to PictureThis in terms of identification ability and also offers a community-based feature within the app that can help to gain a second (and often expert) opinion on a troublesome identification if time is not a factor. This feature might also be very helpful from an educational or training support perspective by providing feedback on a user's identification. Of course over time and with different flora and context of use, other apps would possibly be preferred for other audiences.

These identification apps also seem to have areas of weakness that are not limited to an individual app such as the identification of unlobed leaves (79.69% vs. 98.13% for lobed leaves) and bark photos as a whole, in addition to relatively low identification rates for *Betula* leaves as well as the bark and leaves of *Magnolia* species. While currently problematic, this illuminates a very promising area for future, more targeted software development in order to better address these shared shortcomings.

In general, despite the perception that these apps can be used to correctly identify plants to the species level, it is clear that these apps can, as a whole, only be expected to provide consistent and accurate identifications of Northeastern trees to the genus level at best. While this level of identification may be very helpful in reducing the potential species pool for identification within a genus, it is clear that in their current form, they do not consistently possess the accuracy needed to replace traditional identification tools or experienced professionals.

Literature Cited

Barré P, Stöver BC, Müller KF, Steinhage, V. 2017. LeafNet: A computer vision system for automatic plantspecies identification. *Ecological Informatics.* 40:50–56. https://doi.org/10.1016/j.ecoinf.2017.05.005

Bancks N, North EA, Johnson GR. 2018. An analysis of agreement between volunteer- And researcher-collected urban tree inventory data. *Arboriculture and Urban Forestry*. 44(2): 73-86.

Bilyk ZI, Shapovalov YB, Shapovalov VB, Megalinska AP, Andruszkiewicz F, Dołhańczuk-Śródka A. 2020. Assessment of mobile phone applications feasibility on plant recognition: Comparison with Google Lens AR-app. *CEUR Workshop Proceedings 2731*. p. 61–78.

Bonnet P, Joly A, Goëau H, Champ J, Vignau C, Molino JF, Barthélémy D, Boujemaa N. 2015. Plant identification: man vs. machine: LifeCLEF 2014 plant identification challenge. *Multimedia Tools and Applications*. 75(3):1647–1665. https://doi.org/10.1007/s11042-015-2607-4

Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg KV, Shirk J. 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. BioScience. 54(11): 977-984.

Collins BR, Anderson KH. 1994. *Plant Communities of New Jersey: A Study in Landscape Diversity*. New Brunswick (NJ, USA): Rutgers University Press.

Cope JS, Corney D, Clark JY, Remagnino P, Wilkin P. 2012. Plant species identification using digital morphometrics: A review. *Expert Systems with Applications.* 39:7562–7573.

Crocker SJ, Barnett CJ, Butler BJ, Hatfield MA, Kurtz CM, Lister TW, Meneguzzo DM, Miles PD, Morin RS, Nelson MD, Piva RJ, Riemann R, Smith JE, Woodall CW, Zipse W. 2017. New Jersey Forests 2013. Newtown Square (PA, USA): U.S. Dept. of Agriculture, Forest Service, Northeastern Research Station. RB NRS-109. https://doi.org/10.2737/NRS-RB-109

Crall AW, Newman GJ, Stohlgren TJ, Holfelder KA, Graham J, Waller DM. 2011. Assessing citizen science data quality: an invasive species case study. *Conservation Letters*. 4(6):433–442.

Crown CA, Greer BZ, Gift DM, Watt FS. 2018. Every tree counts: reflections on NYC's third volunteer street tree inventory. *Arboriculture & Urban Forestry.* 44(2):49–58.

Echeverria A, Ariz I, Moreno J, Peralta J, Gonzalez E. 2021. Learning Plant Biodiversity in Nature: The Use of the Citizen–Science Platform iNaturalist as a Collaborative Tool in Secondary Education. *Sustainability.* 13(735):1-12. https://doi.org/10.3390/su13020735

Goëau H, Bonnet P, Joly A, Bakić V, Barbe J, Yahiaoui I, Selmi S, Carré J, Barthélémy D, Boujemaa N, Molino JF. 2013. Pl@ntNet mobile app. In: *Proceedings of the 21st ACM international conference on multimedia*. New York (NY, USA). p. 423–424.

Joly A, Goëau H, Bonnet P, Bakić V, Barbe J, Selmi, J, Yahaioui I, Carré J, Mouysset E, Molino JF, Boujemaa N, Barthélémy D. 2014. Interactive plant identification based on social image data. *Ecological Informatics.* 23(2014):22-34.

Jones H. 2020. Viewpoint: What plant is that? Tests of automated image recognition apps for plant identification on plants from the British flora. *AoB PLANTS*. 12(6):1-9. doi:10.1093/aobpla/plaa052

Keivani M, Mazloum J, Sedaghatfar E, Tavakoli MB. 2020. Automated analysis of leaf shape, texture, and color features for plant classification. *Trait Signal.* 37(1):17–28.

Kingsley TL, Grabner-Hagen MM. 2015. Gamification: Questing to integrate content knowledge, literacy, and 21st century learning. *Journal of Adolescent & Adult Literacy.* 59(1):51–61. doi: 10.1002/jaal.426

Kosmala M, Wiggins A, Swanson A, Simmons B. 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*. 14(10):551-560.

Kumar N, Belhumeur PN, Biswas A, Jacobs DW, Kress WJ, Lopez IC, Soares JV. 2012. Leafsnap: a computer vision system for automatic plant species identification. In: *Proceedings of the European conference on computer vision*. Springer (NY, USA). p. 502–516.

Laćan I, McBride J R. 2008. Pest vulnerability matrix (PVM): A graphic model for assessing the interaction between tree species diversity and urban forest susceptibility to insects and diseases. Urban For. Urban Greening, **7**(4): 291–300.

Munisami T, Ramsurn M, Kishnah S, Pudaruth S. 2015. Plant leaf recognition using shape features and colour histogram with k-nearest neighbour classifiers. *Procedia Computer Science.* 58:740-747.

Perdigones A, Ruiz-Mazarron F, Canas I, Garcia JL. 2021. Apps as a learning vehicle in urban naturing and sports fields. *South Florida Journal of Development, Miami.* v.2(1):432-439. DOI: 10.46932/sfjdv2n1-032

Pollock NB, Howe N, Irizarry NI, Lorusso N, Kruger A, Himmler K, Struwe L. 2015. Personal BioBlitz: a new way to encourage biodiversity discovery and knowledge in k-99 education and outreach. BioScience 65(12):1154-1164. doi: 10.1093/biosci/biv140

Robichaud B, Buell M F. 1973. Vegetation of New Jersey: A Study of Landscape Diversity. Rutgers Univ. Press New Brunswick, NJ. Pp 340.

Roman LA, Scharenbroch BC, Östberg JP, Mueller LS, Henning JG, Koeser AK, Sanders JR, Betz DR, Jordan RC. 2017. Data quality in citizen science urban tree inventories. *Urban Forestry & Urban Greening*. 22:124-135. doi:10.1016/j.ufug.2017.02.001

Sanders J, Grabosky J, Cowie P. 2013. Establishing maximum size expectations for urban trees with regard to designed space. *Arboriculture and Urban Forestry*. 39(2):68-73.

Scharenbroch BC, Carter D, Bialecki M, Fahey R, Scheberl L, Catania M, Roman L A, Bassuk N, Harper R W, Werner L, Siewert A, Miller S, Hutyra L, Raciti S. 2017. A rapid urban site index for assessing the quality of street tree planting sites, Urban Forestry & Urban Greening, 27:279-286. doi. 10.1016/j.ufug.2017.08.017

Söderkvist OJO. 2001. Computer vision classification of leaves from Swedish trees [Master’s Thesis]. Linkoping (Östergötland, Sweden): Linkoping University.

Struwe L, Poster LS, Howe N, Zambell CB, Sweeney PW. 2014. The hands-on, online learning project Flora of Rutgers Campus: campus floras as a student learning tool for plant systematics. *Plant Science Bulletin.* 60:159-169. doi:10.3732/psb.1400008

Tedrow J C F. 1986. Soils of New Jersey. Kreiger Pub Inc. Malabar, FL. Pp. 480 and map.

Trowbridge, P.J., and Bassuk, N.L. 2004. Trees in the urban landscape: Site assessment, design, and installation. Wiley, Hoboken, NJ. Pp207.

USDA Plant Hardiness Zone Map, 2012. Agricultural Research Service, U.S. Department of Agriculture. Accessed from https://planthardiness.ars.usda.gov/.

Wäldchen J, Näder P. 2018. Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering.* 25:507-543 https://doi.org/10.1007/s11831-016-9206-z

Wang B, Brown D, Gao Y, La Salle J. 2013. Mobile plant leaf identification using smart-phones. In: *Proceedings of the 2013 IEEE international conference on image processing*. New York (NY, USA): IEEE. p. 4417–4421.

Wang X, Liang J, Guo F. 2014. Feature extraction algorithm based on dual-scale decomposition and local binary descriptors for plant leaf recognition. *Digital Signal Processing*. 34:101–107.

Wu SG, Bao FS, Xu EY, Wang YX, Chang YF, Shiang CL. 2007. A Leaf Recognition Algorithm for Plant classification Using Probabilistic Neural Network, *IEEE 7th International Symposium on Signal Processing and Information Technology*. Cairo (Cairo, Egypt): IEEE.

Widmann RH. 2005. Forests of the Garden State. Newtown Square (PA, USA): U.S. Dept. of Agriculture, Forest Service, Northeastern Research Station. RB NRS-163. https://doi.org/10.5962/bhl.title.70046

Wojtech M. 2011. *Bark: A Field Guide to Trees of the Northeast*. Lebanon (NH, USA): University Press of New England.

Xing D, Yang J, Jin J, Luo X. 2020. Potential of plant identification apps in urban forestry studies in China: comparison of recognition accuracy and user experience of five apps. *Journal of Forestry Research.* https://doi.org/10.1007/s11676-020-01234-3

Zhao ZQ, Maa LH, Cheung Y, Wu X, Tang Y, Chen CLP. 2015. ApLeaf: An efficient android-based plant leaf identification system. *Neurocomputing*. 151(2015):1112–1119.

Zhao C, Chan SSF, Cham WK, Chu LM. 2015. Plant identification using leaf shapes—A pattern counting approach. *Pattern Recognition.* 48(2015):3203–3215.

PlantSnap (4.00.1) [Mobile application software]. Telluride (CO, USA): PlantSnap, Inc. 10 July 2020. https://apps.apple.com/us/app/plantsnap-plant-identifier/id1451054346

PictureThis (2.0.3) [Mobile application software]. Hangzhou (Zhejiang, People's Republic of China): Glority LLC. 16 July 2020 https://apps.apple.com/us/app/picturethis-plant-identifier/id1252497129.

Plant Identification (1.3.2) [Mobile application software]. Milpitas (CA, USA): Touchberry, Inc. 27 May 2020. https://apps.apple.com/us/app/plant-identification/id1488376022.

PlantNet (3.0.5) [Mobile application software]. (France): plantnet-project.org. 12 March 2020. https://apps.apple.com/us/app/plantnet/id600547573.

LeafSnap (1.6) [Mobile application software]. Hanoi (Red River Delta, Vietnam): Appixi. 15 February 2020. https://apps.apple.com/us/app/leafsnap-plant-identification/id1487972880.

iNaturalist (2.8.7) [Mobile application software]. San Francisco (CA, USA): iNaturalist, LLC. 10 January 2020. https://apps.apple.com/us/app/inaturalist/id421397028.

Acknowledgements
Funding for this work was provided by:

The John and Eleanor Kuser faculty endowment for Urban and Community Forestry

The USDA McIntire-Stennis program: Project NJ17318 Building an Urban Context for Modelling Efforts to Describe and Inform Urban Tree Growth for Services and Management

USDA Forest Service and NJ Dept of Environmental Protection Forestry program: Project Urban Forestry Support 827994

Tables

Apps Considered in Study

App Name	Cost of Use (USD)*	Suggestions Consistently Offered?	Community Identifications	Plant Species database from website claim	Developer
PictureThis	29.99/year	-	-	10,000 +	Glority LLC
iNaturalist	N/A	+	+	Not stated	iNaturalist, LLC
Plant Identify++	1.49/week, 3.99/month, 29.99/year	+	-	Not stated	Touchberry, Inc.
PlantNet (Pl@ntNet)	N/A	+	+	26,722 - world flora 8,490 - USA flora	plantnet-project.org
LeafSnap	N/A	+	-	All 185 tree species in the NE Forest	Appixi
PlantSnap	2.99/month, 19.99/year	+	+	650,000	PlantSnap, Inc.

Table 1: A listing of the plant-identification apps tested during this study. *Cost of Use values are based on the cost at the time of the study, in summer 2020.

Species Evaluation

Family	Genus	Species	Genus Identification		Species Identification		Genus Suggestions		Species Suggestions	
			Bark	Leaf	Bark	Leaf	Bark	Leaf	Bark	Leaf
Altingiaceae	*Liquidambar*	*styraciflua*	20%	100%	20%	100%	65%	100%	65%	100%
Anacardiaceae	*Rhus*	*typhina*	25%	95%	20%	65%	45%	100%	35%	95%
Betulaceae	*Betula*	*allegheniensis*	85%	25%	35%	15%	100%	70%	45%	65%
Betulaceae	*Betula*	*lenta*	75%	45%	50%	45%	95%	90%	55%	50%
Betulaceae	*Betula*	*nigra*	90%	85%	65%	50%	95%	95%	100%	95%
Betulaceae	*Betula*	*populifolia*	95%	100%	35%	40%	95%	100%	70%	100%
Betulaceae	*Carpinus*	*caroliniana*	80%	80%	50%	30%	95%	100%	65%	85%
Cornaceae	*Cornus*	*florida*	50%	95%	50%	60%	75%	100%	75%	100%
Cupressaceae	*Juniperus*	*virginiana*	20%	100%	20%	80%	80%	100%	80%	90%
Cupressaceae	*Taxodium*	*distichum*	50%	95%	50%	90%	95%	100%	95%	100%
Fabaceae	*Cercis*	*canadensis*	5%	95%	5%	90%	5%	100%	5%	100%
Fabaceae	*Gleditsia*	*triacanthos*	45%	85%	45%	85%	80%	100%	80%	100%
Fabaceae	*Robinia*	*pseudoacacia*	50%	100%	50%	100%	70%	100%	70%	100%
Fagaceae	*Castanea*	*dentata*	25%	100%	25%	90%	50%	100%	50%	100%
Fagaceae	*Fagus*	*grandifolia*	90%	100%	70%	95%	100%	100%	90%	100%
Fagaceae	*Quercus*	*alba*	70%	100%	55%	70%	95%	100%	85%	100%
Fagaceae	*Quercus*	*bicolor*	10%	100%	0%	70%	55%	100%	20%	95%
Fagaceae	*Quercus*	*coccinea*	45%	100%	0%	0%	90%	100%	25%	70%
Fagaceae	*Quercus*	*montana*	75%	70%	50%	40%	90%	100%	50%	55%
Fagaceae	*Quercus*	*palustris*	60%	100%	20%	85%	85%	100%	60%	90%
Fagaceae	*Quercus*	*rubra*	55%	100%	50%	90%	95%	100%	80%	95%
Ginkgoaceae	*Ginkgo*	*biloba*	5%	100%	5%	100%	35%	100%	35%	100%
Juglandaceae	*Carya*	*cordiformis*	60%	100%	0%	10%	70%	100%	15%	45%
Juglandaceae	*Carya*	*ovata*	40%	100%	40%	10%	75%	100%	50%	50%
Juglandaceae	*Juglans*	*nigra*	40%	100%	40%	100%	90%	100%	75%	100%
Lauraceae	*Sassafras*	*albidum*	15%	100%	15%	100%	40%	100%	40%	100%
Magnoliaceae	*Liriodendron*	*tulipifera*	30%	100%	30%	100%	75%	100%	75%	100%
Magnoliaceae	*Magnolia*	*soulangeana*	0%	50%	0%	13%	25%	75%	10%	28%
Magnoliaceae	*Magnolia*	*stellata*	10%	25%	0%	10%	30%	55%	0%	25%
Malvaceae	*Tilia*	*cordata*	25%	100%	10%	90%	65%	100%	55%	100%
Malvaceae	*Tilia*	*tomentosa*	10%	100%	0%	10%	60%	100%	5%	40%
Moraceae	*Maclura*	*pomifera*	15%	55%	15%	55%	35%	90%	35%	90%
Nyssaceae	*Nyssa*	*sylvatica*	5%	60%	5%	60%	35%	85%	35%	85%
Oleaceae	*Fraxinus*	*americana*	80%	90%	75%	40%	85%	100%	80%	80%
Oleaceae	*Fraxinus*	*pennsylvanica*	45%	90%	15%	30%	75%	100%	25%	70%
Pinaceae	*Picea*	*abies*	60%	90%	60%	75%	90%	95%	70%	75%
Pinaceae	*Picea*	*pungens*	60%	100%	0%	100%	90%	100%	35%	100%
Pinaceae	*Pinus*	*rigida*	85%	90%	15%	10%	90%	90%	75%	70%
Pinaceae	*Pinus*	*strobus*	50%	100%	45%	95%	75%	100%	75%	100%
Pinaceae	*Pinus*	*sylvestris*	100%	90%	0%	15%	100%	90%	60%	60%
Pinaceae	*Pseudotsuga*	*menziesii*	0%	60%	0%	60%	20%	90%	20%	90%
Pinaceae	*Tsuga*	*canadensis*	20%	100%	20%	50%	50%	100%	50%	100%
Platanaceae	*Platanus*	*occidentalis*	75%	100%	55%	100%	85%	100%	70%	100%
Platanaceae	*Platanus*	*x hispanica*	95%	100%	38%	30%	100%	100%	50%	40%
Rosaceae	*Prunus*	*serotina*	80%	90%	55%	90%	90%	100%	75%	100%
Rosaceae	*Pyrus*	*calleryana*	5%	100%	0%	50%	25%	100%	10%	95%
Salicaceae	*Salix*	*babylonica*	55%	100%	35%	90%	65%	100%	55%	100%

Sapindaceae	*Acer*	*platanoides*	40%	100%	35%	100%	75%	100%	70%	100%
Sapindaceae	*Acer*	*rubrum*	35%	100%	20%	100%	75%	100%	50%	100%
Sapindaceae	*Acer*	*saccharinum*	40%	100%	35%	100%	70%	100%	60%	100%
Sapindaceae	*Acer*	*saccharum*	35%	100%	35%	55%	80%	100%	50%	100%
Sapindaceae	*Aesculus*	*hippocastanum*	15%	100%	15%	100%	40%	100%	40%	100%
Simaroubaceae	*Ailanthus*	*altissima*	20%	100%	20%	100%	20%	100%	20%	100%
Ulmaceae	*Ulmus*	*americana*	10%	90%	10%	75%	60%	95%	60%	95%
Ulmaceae	*Zelkova*	*serrata*	45%	80%	45%	80%	50%	95%	50%	95%
Average			44.09%	89.64%	28.23%	65.32%	69.09%	96.64%	52.36%	85.77%

Table 2: A listing of species used in the evaluation of phone apps for plant identification. Percentages of correct responses across 5 apps are shown for bark and for leaf images, for both identification and for suggestion at genus and species levels.

App Evaluation

	Genus Identification			Species Identification		
App	Combined	Bark	Leaf	Combined	Bark	Leaf
PictureThis	81.36%	65.45%	97.27%	67.84%	51.82%	83.86%
iNaturalist	70.23%	48.18%	92.27%	50.68%	31.82%	69.55%
Plant Identify	63.86%	40.00%	87.73%	44.09%	25.00%	63.18%
PlantNet	60.00%	34.55%	85.45%	36.36%	17.27%	55.45%
LeafSnap	59.09%	32.27%	85.91%	35.11%	14.77%	55.45%
PlantSnap	36.59%	1.36%	71.82%	20.45%	0.00%	40.91%
PictureThis*	81.55%	65.75%	97.27%	67.54%	52.05%	82.95%
Plant Identify*	71.50%	49.16%	90.19%	49.36%	30.73%	64.95%
	Genus Suggestions			Species Suggestions		
	Combined	Bark	Leaf	Combined	Bark	Leaf
PictureThis	81.36%	65.45%	97.27%	67.84%	52.27%	84.32%
iNaturalist	89.55%	79.55%	99.55%	83.41%	70.91%	95.91%
Plant Identify	66.59%	41.36%	91.82%	50.23%	27.27%	73.18%
PlantNet	88.41%	79.55%	97.27%	70.68%	54.55%	86.82%
LeafSnap	88.86%	79.55%	98.18%	72.73%	56.82%	88.64%
PlantSnap	46.36%	2.27%	90.45%	39.55%	0.45%	78.64%
PictureThis*	81.55%	65.75%	97.27%	68.00%	52.51%	83.41%
Plant Identify*	74.55%	50.84%	94.39%	56.23%	33.52%	75.23%

Table 3: A listing of the percentage correct response for six plant identification apps across 55 commonly observed tree species in New Jersey forests and landscapes. Percentages are means of 4 images for each species for bark and for leaf. Combined percentage values are means of 8 observations for leaf and bark combined. PictureThis* and Plant Identify* percentages have been adjusted from their respective percentages by lowering the photo count by the number of photos which the apps failed to make any identification (1 and 47 photos, respectively).

Taxonomic Evaluation

Taxonomic Orders						
Order (Tested Species Count)	Genus Identification			Species Identification		
	Combined	Bark	Leaf	Combined	Bark	Leaf
Fabales (3)	63.33%	33.33%	93.33%	62.50%	33.33%	91.67%
Fagales (16)	75.00%	62.19%	87.81%	44.53%	36.56%	52.50%
Magnoliales (3)	35.83%	13.33%	58.33%	25.42%	10.00%	40.83%
Pinales (9)	70.56%	49.44%	91.67%	43.61%	23.33%	63.89%
Rosales (4)	62.50%	35.00%	90.00%	50.63%	27.50%	73.75%
Taxonomic Families						
Family (number of species tested within family)	Genus Identification			Species Identification		
	Combined	Bark	Leaf	Combined	Bark	Leaf
Betulaceae (5)	76.00%	85.00%	67.00%	41.50%	47.00%	36.00%
Cupressaceae (2)	66.25%	35.00%	97.50%	60.00%	35.00%	85.00%
Fabaceae (3)	63.33%	33.33%	93.33%	62.50%	33.33%	91.67%
Fagaceae (8)	75.00%	53.75%	96.25%	50.63%	33.75%	67.50%
Juglandaceae (3)	73.33%	46.67%	100.00%	33.33%	26.67%	40.00%
Magnoliaceae (3)	35.83%	13.33%	58.33%	25.42%	10.00%	40.83%
Pinaceae (7)	71.79%	53.57%	90.00%	38.93%	20.00%	57.86%
Rosaceae (2)	68.75%	42.50%	95.00%	48.75%	27.50%	70.00%
Sapindaceae (5)	66.50%	33.00%	100.00%	59.50%	28.00%	91.00%
Genera						
Genus (number of species represented within genus)	Genus Identification			Species Identification		
	Combined	Bark	Leaf	Combined	Bark	Leaf
Acer (4)	68.75%	37.50%	100.00%	60.00%	31.25%	88.75%
Betula (4)	75.00%	86.25%	63.75%	41.88%	46.25%	37.50%
Carya (2)	75.00%	50.00%	100.00%	15.00%	20.00%	10.00%
Fraxinus (2)	76.25%	62.50%	90.00%	40.00%	45.00%	35.00%
Magnolia (2)	21.25%	5.00%	37.50%	5.63%	0.00%	11.25%
Picea (2)	77.50%	60.00%	95.00%	58.75%	30.00%	87.50%
Pinus (3)	85.83%	78.33%	93.33%	30.00%	20.00%	40.00%
Platanus (2)	92.50%	85.00%	100.00%	55.63%	46.25%	65.00%
Quercus (6)	73.75%	52.50%	95.00%	44.17%	29.17%	59.17%
Tilia (2)	58.75%	17.50%	100.00%	27.50%	5.00%	50.00%

Table 4: A listing of the of correct response percentage across five plant identification apps as organized by classification level for 55 tree species in New Jersey forests and landscapes. Only Orders

and Families with two or more genera represented in the data were included in this table. Similarly, only Genera with two or more tested species were included.