Predicting Hardwood Porosity Domains: Toward Cascading Computer-Vision Wood Identification Models

Frank C. Owens,[a,*] Prabu Ravindran,[b,c] Adriana Costa,[a] Rubin Shmulsky,[a] and Alex C. Wiedenhoeft [a,b,c,d,e]

Prior work on computer-vision wood identification (CVWID) for North American hardwoods yielded two independent deep learning models – a 22-class model for diffuse-porous woods and a 17-class model for ring-porous woods – but did not address semi-ring-porous woods nor provide a CVWID solution for an unknown specimen without a human first determining which model to deploy. As untrained human operators would lack the anatomical proficiency to differentiate among porosity domains, it is necessary to develop a consolidated model that can identify diffuse-, ring-, and semi-ring-porous woods. Previous research suggests that prediction accuracy might decrease as class number grows. A potential strategy to reduce the number of classes a CVWID system must consider at a time is to hierarchically deploy a cascade of models. In pursuit of a unified model that can cover North American hardwoods of all porosity types, this study compared the accuracies of a consolidated 39-class (ring- + diffuse-porous) model and a consolidated 42-class (ring- + diffuse- + semi-ring-porous) model with a two-tiered, cascading model scheme whereby images are first differentiated into three porosity domain classes and then again into only those taxonomic classes with that porosity. The results showed that the cascading model scheme can mitigate the accuracy reductions incurred by the 42-class model and nearly eliminate the occurrence of cross-domain misidentifications.

DOI: 10.15376/biores.19.4.9741-9772

Keywords: Wood identification; XyloTron; Computer vision; Machine learning; Deep learning; Porosity domain; Cascading models

Contact information: a: Department of Sustainable Bioproducts, Mississippi State University, Starkville, MS, USA; b: Department of Botany, University of Wisconsin, Madison, WI, USA; c: Center for Wood Anatomy Research, USDA Forest Service, Forest Products Laboratory, Madison, WI, USA; d: Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN, USA; e: Departamento de Ciências Biológicas (Botânica), Universidade Estadual Paulista – Botucatu, São Paulo, Brasil.
** Corresponding author: fco7@msstate.edu*

INTRODUCTION

Wood identification is an important tool for combatting the global and pervasive problem of illegal logging and timber trade (Johnson and Laestadius 2011; Dormontt *et al*. 2015; Koch *et al*. 2015; Lowe *et al*. 2016; UNODC 2016), investigating supply chain integrity (Wiedenhoeft *et al*. 2019), and verifying compliance with import/export declaration requirements. Conventional wood identification relies on humans trained in wood anatomy to differentiate among taxa by observing anatomical features under a hand lens and/or microscope (Wheeler and Bass 1998; Gasson 2011). Learning to manually identify wood anatomically requires extensive training, and there is a severe shortage of reliable conventional wood identification capacity in North America (Wiedenhoeft *et al*.

2019) and worldwide. For this reason, scientists are utilizing advances in technology to develop non-conventional wood identification methods based on computer-vision, DNA, mass spectroscopy/spectrometry, and others to reduce dependence on human expertise in wood anatomy (Johnson and Laestadius 2011; Dormontt *et al*. 2015; Lowe *et al*. 2016). Proliferation of these new, non-conventional tools promises to mitigate the shortage of wood identification capacity and better empower industrial compliance with laws and regulations while also enabling law enforcement around the world to detect and prosecute illicit activity.

In conventional human-based identification, a wood anatomist classifies a specimen based on observations of anatomical features (Wheeler and Bass 1998; Gasson 2011). Until an anatomist can recognize a taxon on sight, s/he typically relies on a wood identification key to arrive at a potential identification and then compares the unknown specimen to a known reference specimen and confirms an identification. A key is a decision tree comprising a series of (often dichotomous) choices arranged in a hierarchy. At each level of the hierarchy, the anatomist must decide which of two (or more) descriptions best characterize the anatomical features they observe in the specimen. The option they select determines the next set of choices they must consider further down in the hierarchy, and so on, until they arrive at a terminal description and the name of the taxon (usually a genus, species, or subgeneric category). A well-designed wood identification key works because it focuses attention on only a limited set of anatomical features at a time and prevents wildly deviant conclusions – presuming each user decision is correct – by limiting the remaining options based on previous choices. It also reduces the number of possible solutions as the anatomist descends from one level to the next. In computer science terms, using a wood identification key is the process of choosing a path from the root node (the first level in the key) to a leaf node (the terminal identification) by means of a cascade of solitary decisions at each node on this path.

In many wood identification keys that include temperate hardwood species, a set of alternatives regarding the size and distribution of pores across a growth ring (often referred to as porosity) appears early in the decision tree. These typically include diffuse-porous, ring-porous, and semi-ring-porous. Diffuse-porous woods are characterized by growth rings in which the earlywood pores are not conspicuously larger than the latewood pores (Fig. 1a, Panshin and de Zeeuw 1980). Conversely, ring-porous woods typically exhibit growth rings with a zone of larger earlywood pores abruptly transitioning in size to latewood with conspicuously smaller pores (Fig. 1b, Panshin and de Zeeuw 1980).

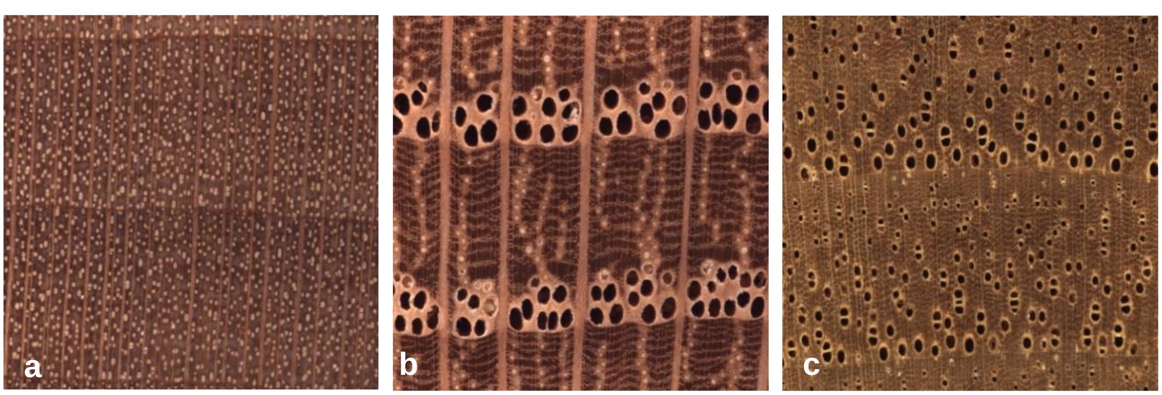

Fig. 1. Examples, from left to right, of the three porosity domains: diffuse-porous (a: *Acer saccharum*), ring-porous (b: *Quercus falcata*), and semi-ring-porous (c: *Diospyros virginiana*) woods

Semi-ring-porous (or sometimes semi-diffuse-porous) is a category intermediate between the former and the latter in which growth rings exhibit a comparatively steady decrease in pore diameter from the earlywood to the latewood (Fig. 1c, Panshin and de Zeeuw 1980).

The concept of porosity domain is well established in the wood anatomy and identification literature. The terms ring-porous and diffuse-porous have been used since at least the 19th Century (Roth 1895). The addition of the category semi-ring-porous in later years (Lodewick 1928; Brown *et al.* 1949) suggests that porosity exists along a spectrum. Wood belonging to the genera *Carya* and *Populus*, for example, exhibit pore sizes and distributions that commonly intergrade between ring- and semi-ring-porous, and semi-ring-porous and diffuse-porous, respectively (Hoadley 1990). Despite the potential for confusion at the border of discrete categories assigned to a more-or-less continuous characteristic, these porosity categories, or domains, have proven useful for separating woods anatomically for more than a century.

A non-conventional wood identification method that has shown great promise is computer-vision (Khalid *et al*. 2008; Esteban *et al*. 2009; Nasirzadeh *et al*. 2010; Wang *et al*. 2010; Hermanson and Wiedenhoeft 2011; Ravindran *et al*. 2018, 2022a; Hwang and Sugiyama 2021). Computer-vision wood identification (CVWID) involves the capture and analysis of digital images of wood specimens by trained classification models leading to an identification (as reviewed in Hwang and Sugiyama 2021). Recently, deep learning, a powerful and flexible approach to training classification models, has been employed to sort images into their appropriate, pre-defined classes (typically corresponding to species, genera and/or anatomically distinct subgeneric categories, *e.g.*, Ravindran *et al*. 2018, 2022a; Liu *et al*. 2024). In addition to high predictive accuracy, the relatively low cost and portability of CVWID systems such as the XyloTron (Ravindran *et al*. 2020) and XyloPhone (Wiedenhoeft 2020) make them readily deployable in the field.

In previous work on North American hardwoods, Ravindran *et al.* (2022a,b) designed separate XyloTron models to differentiate 22 classes of diffuse-porous and 17 classes of ring-porous woods. Deploying each model separately would require the operator to make a porosity domain classification (*e.g.*, diffuse- or ring-porous) prior to selecting the appropriate model. As differentiating among porosity domains requires at least some understanding of wood anatomy, maintaining separate models falls short of the promise of full automation. To realize a unified CVWID model for North American hardwoods, it is necessary to combine ring-porous and diffuse-porous woods into the same model and add semi-ring-porous woods to cover the entire porosity domain spectrum.

Consolidating models from disparate porosity domains has the potential to impact model performance. Previous research suggests that predictive accuracy could decrease as class number grows (Bilal *et al*. 2018; Shigei 2019; Ravindran *et al*. 2022a, b), so it is not unreasonable to expect an increase in misclassifications when combining a 22-class label space with a 17-class label space to yield a 39-class model. In addition, even highly accurate CVWID models have been shown to occasionally produce anatomically unexplainable predictions of the kind that no human would likely make (Ravindran *et al.* 2021, 2022a,b), defined as a Type 3 misclassification in Ravindran *et al*. (2022a). In a consolidated model of North American hardwoods, credibility-reducing cross-domain misclassifications between woods of disparate porosity domains become a possibility.

While they both arrive at a common end – namely, the assignment of a taxonomic label to a specimen of wood – the process of CVWID is different from conventional human-based wood identification. The image-based CVWID used in Ravindran *et al.*

(2022a,b) differs from human-based identification in two important ways. First, it is not clear how the features the model detects in the digital image correspond to anatomical features that would be used in a key. Second, computer-vision image classification is based on a single decision step as opposed to the explicit, cascading sequence of decisions found in a wood identification key. In short, though the features employed by the model are implicitly hierarchical, it is difficult to rigorously explain why a decision was made by a trained classification model.

These differences contribute to a few limitations. CVWID, though rapid and highly accurate, still has the potential to make Type 3 misclassifications, those that trained humans would almost never make such as cross-domain misidentifications (*e.g.*, confusing a hardwood for a softwood or a diffuse-porous wood for a ring-porous wood). Also, as the number of classes increases, the model has no explicit mechanism to break down the task of identification into smaller steps to limit the number of classes it must discriminate. Breaking down a CVWID model with many classes into a cascade or tree of models with fewer classes at each level might help reduce the occurrence of Type 3 misclassifications and better enable a CVWID system to handle larger numbers of classes without a drop in accuracy, especially in the medium-sized dataset regime typical in CVWID.

Until a CVWID model can employ explicit and enumerable semantic information allowing it to identify the pixels in an image that correspond to particular anatomical features (such as vessels, rays, fibers, *etc.*), it is not possible to duplicate the decision-making process of a wood identification key. It is possible to approximate that process by developing new classifiers based on wood anatomical character-based label spaces rather than taxonomic label spaces – for example, woods with marginal parenchyma *vs.* woods without marginal parenchyma.

In pursuit of a unified model that can cover all commercial North American hardwoods, this study had two objectives. The first was to determine how accurately a convolutional neural network (CNN)-based model, trained using the same pipeline from Ravindran *et al.* (2022a, 2022b), can predict the 39 classes of (22) diffuse- and (17) ring-porous woods from those publications plus three additional classes of semi-ring-porous woods. The second was to determine if gains in accuracy might be achieved (or losses in accuracy might be mitigated) by creating a two-level decision tree wherein images are first classified by a root classifier into one of three porosity domain classes (diffuse-, ring- or semi-ring-porous) and then again by a first level model covering only those woods with that porosity. The results of this investigation should help determine viable options for structuring future CVWID models aimed at covering large numbers of classes using a richer set of anatomical characters akin to a traditional wood identification key.

EXPERIMENTAL

Materials

Specimens and images

The datasets for this study comprised 1) the images of diffuse-porous specimens used in Ravindran *et al*. (2022a), 2) the images of ring-porous specimens used in Ravindran *et al*. (2022b), and 3) a new set of images from specimens of three semi-ring-porous species (*Diospyros virginiana*, *Juglans cinerea*, and *Juglans nigra*). The images for the training datasets were captured from specimens, chosen to represent characteristic wood anatomical variability for the classes, sourced exclusively from a) the MADw and SJRw collections at the USDA Forest Products Laboratory and b) the Tw collection at the Royal Museum of Africa. The images for the testing datasets were captured from specimens sourced exclusively from the David A. Kribs (PACw) and teaching collections at Mississippi State University. Specimens in all xylaria were at moisture contents consistent with ambient, indoor, conditioned conditions (typically assumed to range from ~5% to 9% moisture content depending on season and location).

To prepare each wood specimen for imaging, the transverse surface was polished in coarse-to-fine progression on a benchtop disc sander at grits of 80, 180, 240, 400, 600, 800, and 1500, but note that testing-dataset images produced at much coarser grit levels or with a knife as in the field (Ravindran *et al*. 2023) are still largely identifiable, as are digitally perturbed testing dataset images (Owens *et al*. 2024) for a Peruvian woods CVWID model (Ravindran *et al*. 2021). The polished surfaces were imaged using the XyloTron platform at a linear resolution of ∼3.1 microns per pixel. Multiple, non-overlapping 2048 × 2048-pixel images (representing 6.35mm-by-6.35mm of tissue) were acquired from each specimen. Commonly, images are of heartwood, though some sapwood images are doubtless present for each class. Images with a sapwood-heartwood transition were culled.

Label spaces

Descriptions of the six label spaces used in this study are shown in Table 1. The first three were used for the “domain-based” models: the 22-class label space for the diffuse-porous woods from Ravindran *et al.* (2022a), the 17-class label space for the ring-porous woods from Ravindran *et al.* (2022b), and a new 3-class label space for semi-ring-porous woods.

Table 1. Descriptions of the Six Label Spaces Used in This Study

	Label Spaces	**Description**
Domain-based Models	22DP	The original 22-class label space including a class for each diffuse-porous wood from Ravindran *et al.* 2022a (left column of Table 2)
	17RP	The original 17-class label space including a class for each ring-porous wood from Ravindran *et al.* 2022b (center column of Table 2)
	3SRP	A new 3-class label space including a class for each semi-ring-porous wood from the current study (right column of Table 2)
Consolidated Models	39DP-RP	A new 39-class label space including the 22 diffuse-porous classes from Ravindran *et al*. 2022a plus the 17 ring-porous classes from Ravindran *et al.* 2022b (left and center columns of Table 2)
	42DP-RP-SRP	A new 42-class label space including the 22 diffuse-porous classes from Ravindran *et al*. 2022a, the 17 ring-porous classes from Ravindran *et al.* 2022b, and the 3 semi-ring classes from the current study (all three columns of Table 2)
Porosity Model	3POR	A new 3-class label space representing the 3 porosity domains: 1) an aggregate diffuse-porous class including all the woods with diffuse-porous structure (entire left column of Table 2), 2) an aggregate ring-porous class including all the woods with a ring-porous structure (entire center column of Table 2), and 3) an aggregate semi-ring-porous class including all the woods with a semi-ring-porous structure (entire right column of Table 2)

The classes comprising these three label spaces are detailed in Table 2. The next two label spaces are used for the “consolidated” models: a 39-class label space (39DP-RP) comprising all the classes from 22DP and 17RP, and a 42-class label space (42-DP-RP-SRP) comprising all the classes from 22DP, 17RP and 3SRP. The final 3-class label space models the three porosity domains: an aggregate diffuse-porous (DP) class including all the woods with diffuse-porous structure (entire left column of Table 2); an aggregate ring-porous (RP) class including all the woods with a ring-porous structure (entire center column of Table 2); and an aggregate semi-ring-porous (SRP) class including all the woods with a semi-ring-porous structure (entire right column of Table 2). Details of label, specimen, image, and taxa counts are shown in Tables 3 and 4 by label space. When referring to a CVWID class, class names are written without italics, while italicization is used when referring to the same woods as botanical entities (*e.g.*, the class Diospyros vs. the genus *Diospyros*). Finer details of the taxa included in the training and testing datasets for 39DP-RP and 42DP-RP-SRP are provided in Tables S1, S2 in the Appendix.

Table 2. Porosity Domain Membership of Woods Used in Training and Testing

Diffuse-porous (DP) (Ravindran *et al.* 2022a)	**Ring-porous (RP)** (Ravindran *et al.* 2022b)	**Semi-ring-porous (SRP)**
AcerH (hard maples) AcerS (soft maples) Aesculus Alnus Arbutus Betula Carpinus Fagus Frangula Fruitwood (comprising:) • *Crataegus* spp. • *Malus* spp. • *Prunus* spp. (except *P. serotina*) • *Pyrus* spp. • *Sorbus* spp. Liquidambar Liriodendron Magnolia Nyssa Ostrya Oxydendrum Platanus Populus Prunus (*P. serotina* only) Rhamnus Salix Tilia	Asimina Carya Castanea Catalpa Celtis Cladrastis Fraxinus Gleditsia Gymnocladus Maclura Morus QuercusR (red oaks) QuercusW (white oaks) Robinia Sassafras UlmusH (hard elms) UlmusS (soft elms)	Diospyros JuglansC (*J. cinerea*) JuglansN (*J. nigra*)

Table 3. Training Dataset Details by Label Space

Detail	22DP (counts)	17RP (counts)	3SRP (counts)	39DP-RP (counts)	42DP-RP-SRP (counts)	3POR (counts)
Images	5184	4045	734	9229	9963	2716
Specimens	504	452	54	956	1010	1001
Taxa	98	64	3	162	165	163
Classes	22	17	3	39	42	3
These wood specimens were sourced from the USDA Forest Products Laboratory's MADw and SJRw collections and the Royal Museum of Africa's Tw collection. Images of the specimens from these xylaria were used **only** for training and **not** for testing.						

Table 4. Test Dataset Details by Label Space

Detail	22DP (counts)	17RP (counts)	3SRP (counts)	39DP-RP (counts)	42DP-RP-SRP (counts)	3POR (counts)
Images	1209	936	131	2145	2276	2276
Specimens	284	198	28	482	510	510
Taxa	69	40	3	109	112	112
Classes	22	17	3	39	42	3
These wood specimens were sourced from Mississippi State University's PACw and teaching collections. Images of the specimens from these xylaria were used **only** for testing and **not** for training.						

Methods

Machine learning models

CNNs (LeCun *et al.* 1989) with ImageNet (Russakovsky *et al.* 2015) pretrained ResNet34 (He *et al.* 2015) backbones and custom classification heads were trained for each of the six label spaces. A two-stage transfer learning strategy (freezing the backbone and training only the randomly initialized custom head followed by full network fine-tuning) was employed to train the models along with a data augmentation strategy that included horizontal/vertical flips, small rotations, and cutout (Devries and Taylor 2017). For both the training stages, random patches of 2048 × 768 pixels were downsampled to 512 × 192 pixels and fed into the network in mini batches of size 16. The Adam optimizer (Kingma and Ba 2015) with cosine annealing (Smith 2018) of the learning rate and momentum was used for updating the model weights during the training process. Further details about the architecture and the two-stage (Howard and Gugger 2020) transfer learning (Pan and Yang 2010) training methodology can be found in Ravindran *et al*. (2019) and Arévalo *et al*. (2021). Scientific Python tools (Pedregosa *et al.* 2011) and the PyTorch deep learning framework (Paszke *et al.* 2019) were used for model definition, training, and evaluation.

The predictive performance of the trained field models (the trained models obtained by using the entire training data) was evaluated using the top-1 and top-2 specimen level accuracies on the mutually exclusive (from completely different collections) testing dataset. The majority of the class predictions for the (up to 5) images contributed by a specimen was taken as the top-1 specimen level prediction. For top-2 accuracy analysis, if the true class of a specimen was one of the top-2 predicted classes in an equally weighted voting of (up to 5) image-level top-2 predictions, then the specimen was considered to be correctly classified. The field model performance was evaluated using the proxy field testing approach (in which specimens for training and testing were obtained from different xylaria) introduced in Ravindran *et al.* (2021). The importance of mutually exclusive training and evaluation datasets is elaborated in Ravindran and Wiedenhoeft (2022).

Models with a ResNet50 backbone were also trained and evaluated, and these results are presented in Table S3 and Figs. S1, S2 in the Appendix. Additionally, results of five-fold cross-validation analyses (i.e. internal validation) for both ResNet34 and ResNet50 based model architectures are presented in Table S4 and Figs. S3 – S6 in the Appendix. The field models were evaluated in the following manner under the following assumptions.

Domain-based model scheme

To test the individual accuracies of the domain-based diffuse- (22DP), ring- (17RP) and semi-ring-porous (3SRP) models, each model was run on the test dataset that corresponded to its porosity domain. In the case of actual field deployment, a specimen would first be examined by a human operator who would make a porosity domain determination and then select the corresponding domain-based XyloTron model: 22DP, 17RP, or 3SRP (Fig. 2). These accuracies were evaluated under a best-case scenario assuming that the human operator made the initial porosity domain classification without error.

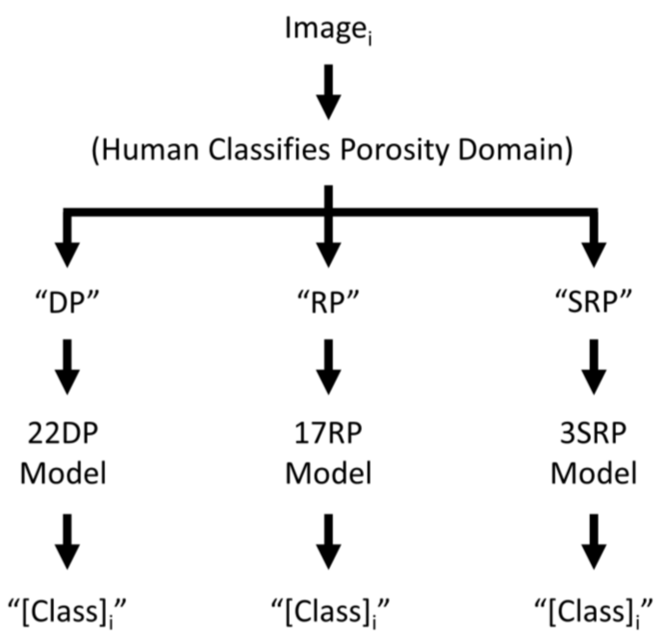

Fig. 2. In the domain-based model scheme, a human would first make a (correct) porosity domain determination (diffuse-porous, DP; ring-porous, RP; or semi-ring-porous, SRP) for the specimen and then select the corresponding domain-based XyloTron model.

Consolidated model schemes

To test the accuracies of the consolidated models, the 39DP-RP model was run on the test dataset that included all diffuse- and ring-porous woods, and the 42-DP-RP-SRP model was run on the test dataset that included all diffuse-, ring-, and semi-ring-porous woods. In the case of actual field deployment, the human operator would deploy one model without first making a porosity domain determination (Fig. 3). In the case of the 39DP-RP model, we assume that there are no semi-ring-porous woods in this particular region of deployment.

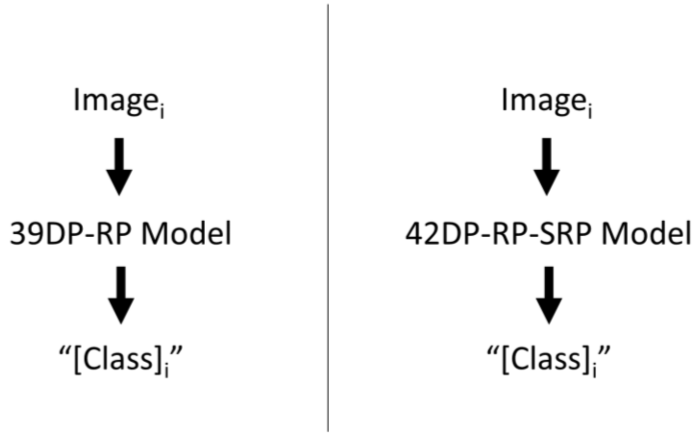

Fig. 3. In the consolidated model schemes, a human does not need to first make a (correct) porosity domain determination. One XyloTron model would be deployed. In the case of the 39DP-RP model (left), we assume that there are no semi-ring-porous woods in the region of deployment. (Abbreviations: diffuse-porous, DP; ring-porous, RP; semi-ring-porous, SRP)

Cascading model scheme

To test the accuracy of a cascading model scheme, the 3POR XyloTron model was deployed first to classify the porosity domain of the image. Based on the output, the image was submitted to the XyloTron model that corresponded to the identified porosity domain. In the case of actual field deployment, the XyloTron software would implement the two models sequentially without need for the human operator to make a porosity domain determination (Fig. 4) or need to manually submit the image to the corresponding model as we have done here.

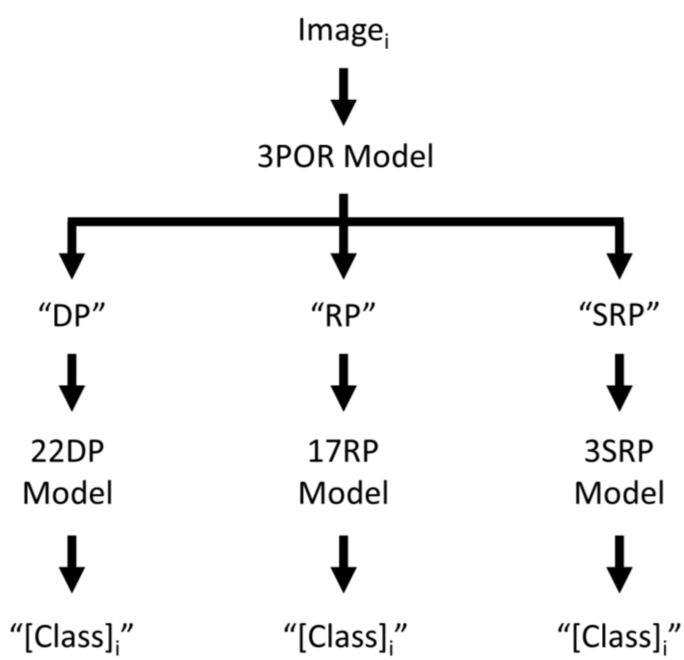

Fig. 4. In the cascading model scheme, a human does not need to first make a porosity domain determination. Instead, the 3POR XyloTron model is deployed first to classify the porosity of the specimen (diffuse-porous, DP; ring-porous, RP; or semi-ring-porous, SRP). Based on the output, the XyloTron model that corresponded to that porosity domain type would then be deployed.

Model top-1 accuracies are reported below, and top-2 accuracies can be found in the Appendix.

RESULTS AND DISCUSSION

Domain-based Model Scheme Accuracy

Confusion matrices for the domain-based diffuse- (22DP) and ring-porous (17RP) models can be found in Ravindran *et al.* 2022a and 2022b, respectively, along with misidentification analyses. Their respective top-1 predictive accuracies were 80.6% and 91.4%.

The top-1 predictive accuracy for the domain-based semi-ring-porous model (3SRP) was 100.0%, and it was tested on 3 specimens of *Diospyros virginiana*, 12 specimens of *Juglans cinerea*, and 13 specimens of *Juglans nigra*.

Assuming the XyloTron operator would be able to separate all the specimens into their correct porosity domains and apply the appropriate domain-based model to each specimen, the overall accuracy for the domain-based model scheme would be 85.9%.

Consolidated Model Scheme Accuracies

39-class consolidated model

The top-1 predictive accuracy for the consolidated 39-class model (39DP-RP) was 84.2%. Broken down by porosity domain, the predictive accuracy for diffuse-porous woods was 80.1% and for ring-porous woods was 90.0% (Fig. 5, Tables 5, 6).

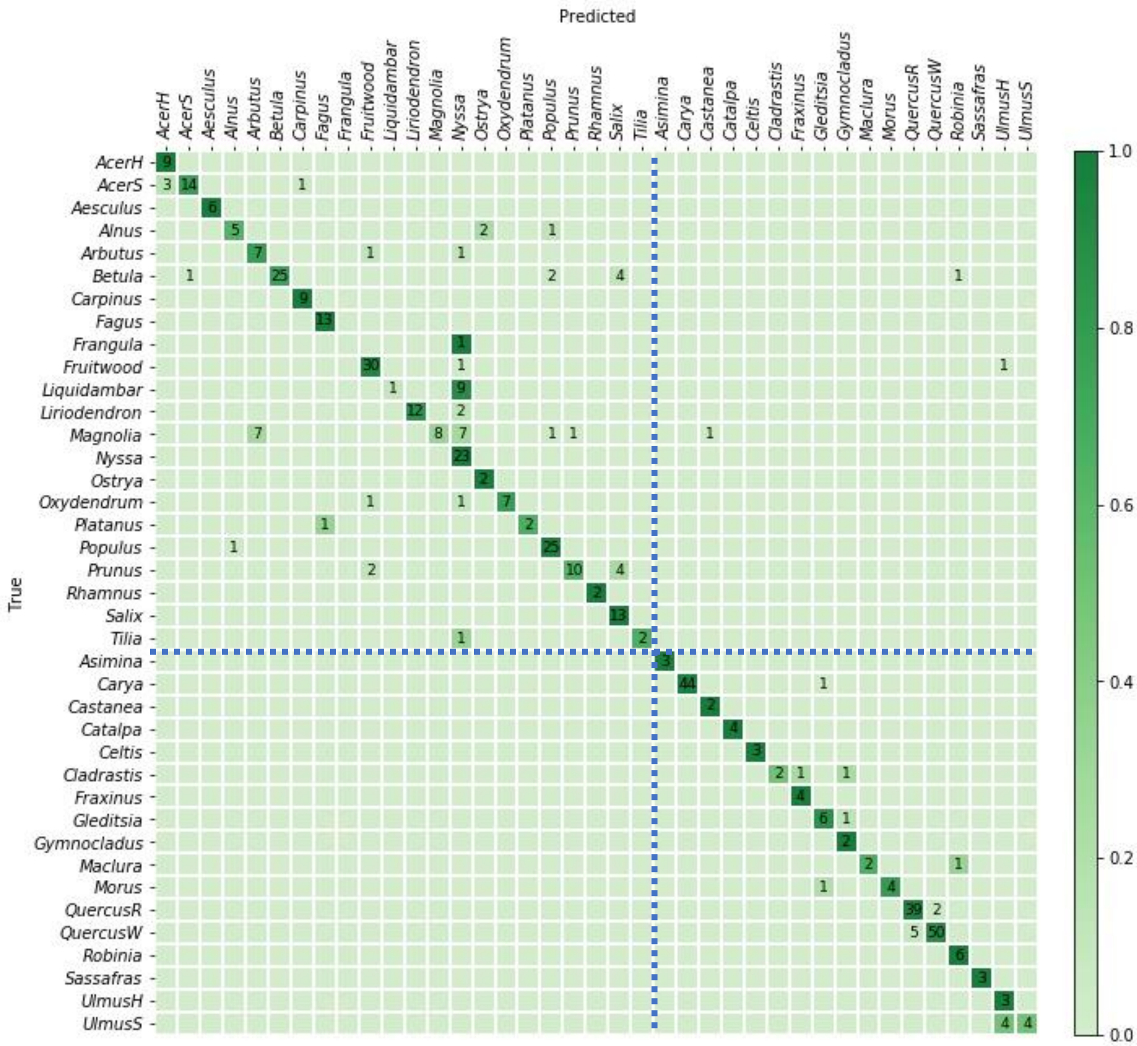

Fig. 5. Confusion matrix of ResNet34 field model on test specimens for the consolidated 39DP-RP model. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right region are cross-domain misclassifications for porosity (diffuse-porous woods mistaken for ring-porous woods). N = 482

The confusion matrix for the consolidated 39DP-RP model appears in Fig. 5 and is separated into four regions by blue dotted lines, which delineate the boundaries between the two porosity domains on the vertical and horizontal axes. The diffuse-porous classes run from AcerH to Tilia while the ring-porous classes run from Asimina to UlmusS. Counts in the diagonal represent correct predictions. Off-diagonal counts indicate misidentifications. The upper right region of the matrix is populated by three cross-domain prediction errors representing three different diffuse-porous specimens misclassified as three different ring-porous woods. In contrast, the lower left region contains no cross-domain errors showing that none of the ring-porous woods were misclassified as diffuse-porous woods. Off-diagonal counts in the upper left and lower right regions indicate erroneous predictions within a porosity domain.

Table 5. Prediction Errors and Accuracies by Input Class for 39DP-RP (Diffuse-porous portion of the dataset)

Input Class Label	Specimen Total Count	Count Missed	% Missed	Mistaken Classes (Count)	Count Correct	% Accurate
AcerH	9	0	0.0	--	9	100.0
AcerS	18	4	22.2	AcerH (3) Carpinus (1)	14	77.8
Aesculus	6	0	0.0	--	6	100.0
Alnus	8	3	37.5	Ostrya (2) Populus (1)	5	62.5
Arbutus	9	2	22.2	Fruitwood (1) Nyssa (1)	7	77.8
Betula	33	8	24.2	Salix (4) Populus (2) AcerS (1) * Robinia (1)	25	75.8
Carpinus	9	0	0.0	--	9	100.0
Fagus	13	0	0.0	--	13	100.0
Frangula	1	1	100.0	Nyssa (1)	0	0.0
Fruitwood	32	2	6.2	Nyssa (1) * UlmusH (1)	30	93.8
Liquidambar	10	9	90	Nyssa (9)	1	10.0
Liriodendron	14	2	14.3	Nyssa (2)	12	85.7
Magnolia	25	17	68.0	Arbutus (7) Nyssa (7) Populus (1) Prunus (1) * Castanea (1)	8	32.0
Nyssa	23	0	0.0	--	23	100.0
Ostrya	2	0	0.0	--	2	100.0
Oxydendrum	9	2	22.2	Fruitwood (1) Nyssa (1)	7	77.8
Platanus	3	1	33.3	Fagus (1)	2	66.7
Populus	26	1	3.8	Alnus (1)	25	96.2
Prunus	16	6	37.5	Salix (4) Fruitwood (2)	10	62.5
Rhamnus	2	0	0.0	--	2	100.0
Salix	13	0	0.0	--	13	100.0
Tilia	3	1	33.3	Nyssa (1)	2	66.7

Class labels marked with an asterisk (*) indicate cross-domain misclassifications for porosity. Class labels highlighted in black indicate classes for which the domain-based 22DP model (from Ravidran *et al*. 2022a) classified all specimens correctly but for which the consolidated 39DP-RP model misclassified at least one specimen. Class labels in gray indicate the opposite: the domain-based 22DP model misclassified at least one specimen while the consolidated 39DP-RP model classified all specimens correctly.

Table 6. Prediction Errors and Accuracies by Input Class for 39DP-RP (Ring-porous portion of the dataset)

Input Class Label	Specimen Total Count	Count Missed	% Missed	Mistaken Classes (Count)	Count Correct	% Accurate
Asimina	3	0	0.0	--	3	100.0
Carya	45	1	2.2	Gleditsia (1)	44	97.8
Castanea	2	0	0.0	--	2	100.0
Catalpa	4	0	0.0	--	4	100.0
Celtis	3	0	0.0	--	3	100.0
Cladrastis	4	2	50.0	Fraxinus (1) Gymnocladus (1)	2	50.0
Fraxinus	4	0	0.0	--	4	100.0
Gleditsia	7	1	14.3	Gymnocladus (1)	6	85.7
Gymnocladus	2	0	0.0	--	2	100.0
Maclura	3	1	33.3	Robinia (1)	2	66.7
Morus	5	1	20.0	Gleditsia (1)	4	80.0
QuercusR	41	2	4.9	QuercusW (2)	39	95.1
QuercusW	55	5	9.1	QuercusR (5)	50	90.9
Robinia	6	0	0.0	--	6	100.0
Sassafras	3	0	0.0	--	3	100.0
UlmusH	3	0	0.0	--	3	100.0
UlmusS	8	4	50.0	UlmusH (4)	4	50.0

Class labels marked with an asterisk (*) indicate cross-domain misclassifications for porosity. Class labels highlighted in black indicate classes for which the domain-based 17RP model (from Ravidran *et al*. 2022b) classified all specimens correctly but for which the consolidated 39DP-RP model misclassified at least one specimen. Class labels in gray indicate the opposite: the domain-based 17RP model misclassified at least one specimen while the consolidated 39DP-RP model classified all specimens correctly.

A breakdown of prediction errors and accuracies by class appears in Tables 5 and 6. Class labels marked with an asterisk (*) indicate cross-domain misclassifications for porosity. Class labels highlighted in black indicate classes for which the domain-based 22DP or 17RP model (from Ravidran *et al*. 2022a,b) classified all specimens correctly but for which the consolidated 39DP-RP model misclassified at least one specimen. Class labels in gray indicate the opposite: the domain-based 22DP or 17RP model misclassified at least one specimen while the consolidated 39DP-RP model classified all specimens correctly.

42-class consolidated model

The top-1 predictive accuracy for the consolidated 42-class consolidated model (42DP-RP-SRP) was 77.1%. Broken down by porosity domain, the predictive accuracy for diffuse-porous woods is 69.7% and for ring-porous woods is 83.9%, and for semi-ring-porous woods is 100% (Fig. 6, Tables 7, 8).

The confusion matrix for the consolidated 42DP-RP-SRP model appears in Fig. 6 and is separated into nine regions by blue dotted lines, which delineate the boundaries between the three porosity domains on the vertical and horizontal axes. The diffuse-porous woods run from AcerH to Tilia, the semi-ring-porous woods from Diospyros to JuglansN, and the ring-porous woods from Asimina to UlmusS. Counts in the diagonal represent correct predictions. Off-diagonal counts indicate misclassifications. The upper right region of the matrix is populated by seven cross-domain prediction errors representing five different diffuse-porous woods misclassified as three different ring-porous woods. In contrast, the lower left region contains no cross-domain errors, showing that none of the ring-porous woods were misclassified as diffuse-porous woods. Off-diagonal counts in the upper left and lower right regions indicate erroneous predictions within the same porosity domain. The semi-ring-porous woods had no misclassifications.

A breakdown of prediction errors and accuracies by class appears in Tables 7 and 8. Class labels marked with an asterisk (*) indicate cross-domain misclassifications for porosity. Class labels highlighted in black indicate classes for which the domain-based 22DP or 17RP model (from Ravidran *et al.* 2022a,b) identified all specimens without any misclassifications but for which the consolidated 42DP-RP-SRP model misclassified at least one specimen. Class labels in gray indicate the opposite: the domain-based 22DP or 17RP model misclassified at least one specimen while the consolidated 42DP-RP-SRP model correctly classified all specimens.

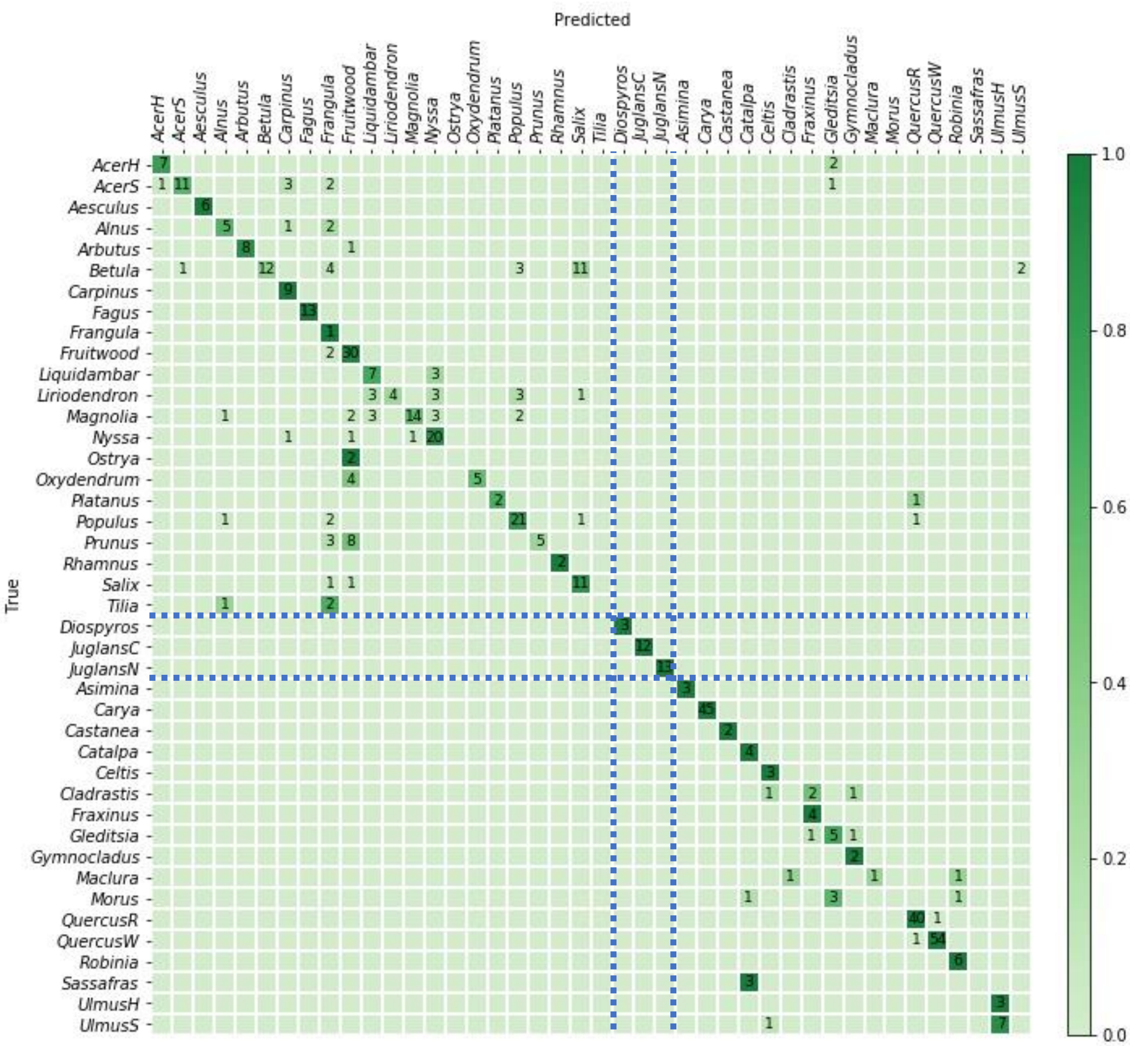

Fig. 6. Confusion matrix of ResNet34 field model on test specimens for label space 42DP-RP-SRP. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes, from Diospyros to JuglansN are semi-ring-porous classes, and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right region are the only cross-domain misclassifications for porosity domain (diffuse-porous mistaken for ring-porous). N= 510

Table 7. Prediction Errors and Accuracies by Input Class for 42DP-RP-SRP (Diffuse-porous portion of the dataset)

Input Class Label	Specimen Total Count	Count Missed	% Missed	Mistaken Classes (Count)	Count Correct	% Accurate
AcerH	9	2	22.2	* Gleditsia (2)	7	77.8
AcerS	18	7	38.9	Carpinus (3) Frangula (2) AcerH (1) * Gleditsia (1)	11	61.1
Aesculus	6	0	0.0	--	6	100.0
Alnus	8	3	37.5	Frangula (2) Carpinus (1)	5	62.5
Arbutus	9	1	11.1	Fruitwood (1)	8	88.9
Betula	33	21	63.6	Salix (11) Frangula (4) Populus (3) * UlmusS (2) AcerS (1)	12	36.4
Carpinus	9	0	0.0	--	9	100.0
Fagus	13	0	0.0	--	13	100.0
Frangula	1	0	0.0	--	1	100.0
Fruitwood	32	2	6.3	Frangula (2)	30	93.8
Liquidambar	10	3	30.0	Nyssa (3)	7	70.0
Liriodendron	14	10	71.4	Liquidambar (3) Nyssa (3) Populus (3) Salix (1)	4	28.6
Magnolia	25	11	44.0	Liquidambar (3) Nyssa (3) Fruitwood (2) Populus (2) Alnus (1)	14	56.0
Nyssa	23	3	13.0	Carpinus (1) Fruitwood (1) Magnolia (1)	20	87.0
Ostrya	2	2	100.0	Fruitwood (2)	0	0.0
Oxydendrum	9	4	44.4	Fruitwood (4)	5	55.6
Platanus	3	1	33.3	* QuercusR (1)	2	66.7
Populus	26	5	19.2	Frangula (2) Alnus (1) Salix (1) * QuercusR (1)	21	80.8
Prunus	16	11	68.8	Fruitwood (8) Frangula (3)	5	31.3
Rhamnus	2	0	0.0	--	2	100.0
Salix	13	2	15.4	Fruitwood (1) Frangula (1)	11	84.6
Tilia	3	3	100.0	Frangula (2) Alnus (1)	0	0.0

Class labels marked with an asterisk (*) indicate cross-domain misclassifications for porosity. Class labels highlighted in black indicate classes for which the domain-based 22DP model (from Ravidran *et al*. 2022a) classified all specimens correctly but for which the consolidated 42DP-RP-SRP model misclassified at least one specimen. Class labels in gray indicate the opposite: the domain-based 22DP model misclassified at least one specimen while the consolidated 42DP-RP-SRP model classified all specimens correctly.

Table 8. Prediction Errors and Accuracies by Input Class for 42DP-RP-SRP (Ring-porous and Semi-ring-porous portions of the dataset)

Porosity	Input Class	Specimen Total Count	Count Missed	% Missed	Mistaken Classes (Count)	Count Correct	% Accurate
Ring	**Asimina**	3	0	0.0	--	3	100.0
	Carya	45	0	0.0	--	45	100.0
	Castanea	2	0	0.0	--	2	100.0
	Catalpa	4	0	0.0	--	4	100.0
	Celtis	3	0	0.0	--	3	100.0
	Cladrastis	4	4	100.0	Fraxinus (2) Celtis (1) Gymnocladus (1)	0	0.0
	Fraxinus	4	0	0.0	--	4	100.0
	Gleditsia	7	2	28.6	Fraxinus (1) Gymnocladus (1)	5	71.4
	Gymnocladus	2	0	0.0	--	2	100.0
	Maclura	3	2	66.7	Cladrastis (1) Robinia (1)	1	33.3
	Morus	5	5	100.0	Gleditsia (3) Catalpa (1) Robinia (1)	0	0.0
	QuercusR	41	1	2.4	QuercusW (1)	40	97.6
	QuercusW	55	1	1.8	QuercusR (1)	54	98.2
	Robinia	6	0	0.0	--	6	100.0
	Sassafras	3	3	100.0	Catalpa (3)	0	0.0
	UlmusH	3	0	0.0	--	3	100.0
	UlmusS	8	8	100.0	UlmusH (7) Celtis (1)	0	0.0
Semi-ring	**Diospyros**	3	0	0.0	--	3	100.0
	JuglansC	12	0	0.0	--	12	100.0
	JuglansW	13	0	0.0	--	13	100.0

Class labels marked with an asterisk (*) indicate cross-domain misclassifications for porosity. Class labels highlighted in black indicate classes for which the domain-based 17RP model (from Ravidran *et al*. 2022a) classified all specimens correctly but for which the consolidated 42DP-RP-SRP model misclassified at least one specimen. Class labels in gray indicate the opposite: the domain-based 17RP model misclassified at least one specimen while the consolidated 42DP-RP-SRP model classified all specimens correctly.

Cascading Model Scheme Accuracy

The top-1 predictive accuracy for the 3-class porosity model (3POR) was 99.8%. Broken down by porosity domain, the ring-porous and semi-ring-porous accuracies were both 100.0%, while the diffuse-porous accuracy was 99.6% (Fig. 7).

The confusion matrix for 3POR appears in Fig. 7. One specimen of *Populus* was misclassified as semi-ring-porous, which, as noted in the Introduction, is a sensible error (a Type 1 misclassification, per Ravindran *et al*. 2022a) for this taxon.

If the output of the 3POR model were used to select which domain-based model to deploy next, the overall accuracy of the cascading models would be the product of the

3POR model accuracy × the accuracy of respective domain-based model (22DP, 17RP, or 3SRP). The top-1 predictive accuracy for the cascading scheme in total would be 85.7%. Broken down by porosity domain, the diffuse-porous accuracy would be 80.3% (= 99.6% × 80.6%), the ring-porous accuracy 91.4% (=100.0% × 91.4%), and the semi-ring-porous accuracy 100.0% (=100.0% × 100.0%, Table 9).

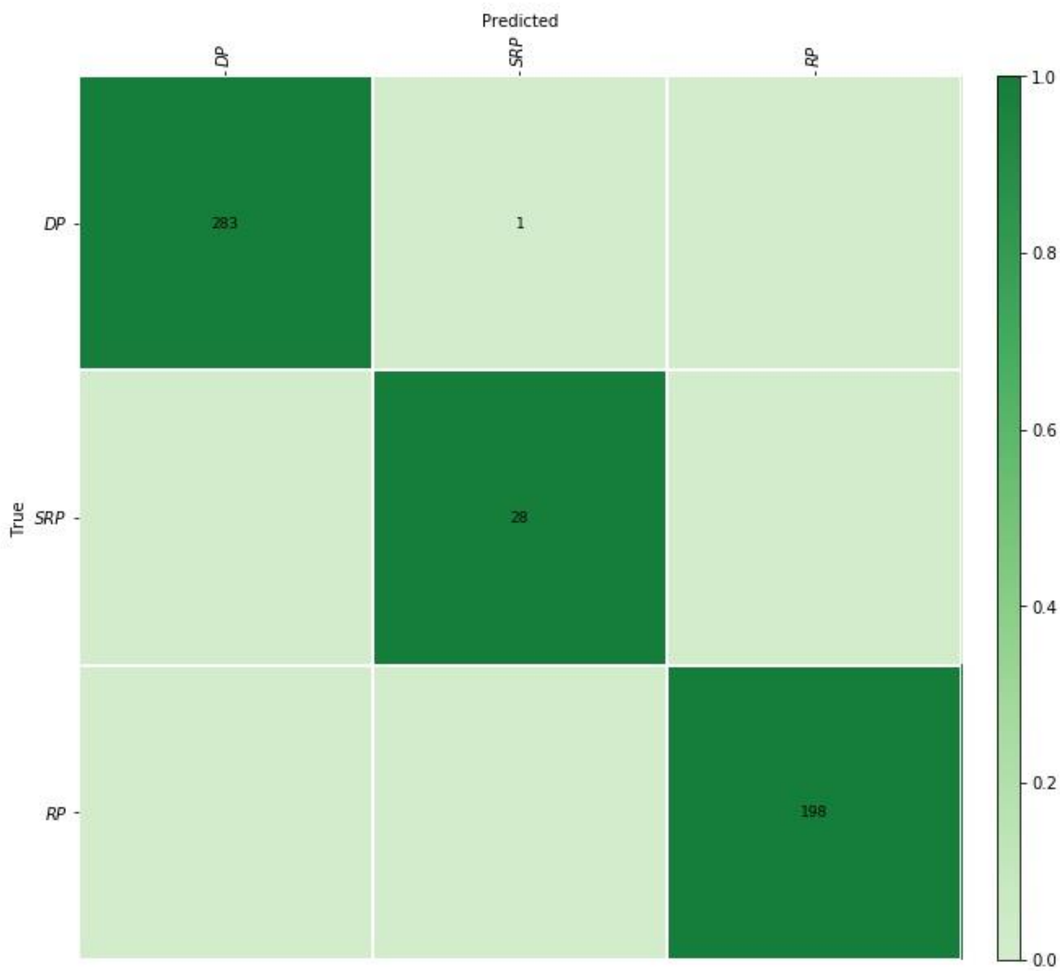

Fig. 7. Confusion matrix of ResNet34 field model on test specimens for label space 3POR. Counts in the diagonal cells indicate correct predictions. The lone off diagonal misclassification was a specimen of *Populus* (diffuse-porous) mistaken for a semi-ring-porous wood. N= 510.

Accuracy Comparison Among Schemes

The results of all models/schemes are summarized and compared in Table 9.

Table 9. Summary Counts and Predictive Accuracies for Each Model/Scheme

Porosity Domain	Test Specimen Count	Domain-based Model Accuracies	39DP-RP Model Accuracy	42DP-RP-SRP Model Accuracy	3POR × Domain-based Cascading Model Accuracies	
Diffuse-porous	284	80.6%[*1]	80.1%	69.7%	99.6% × 80.6% =	80.3%
Ring-porous	198	91.4%[*2]	90.0%	83.9%	100.0% × 91.4% =	91.4%
Semi-ring-porous	28	100.0%	---	100.0%	100.0% × 100.0% =	100.0%
Overall	510	85.9%	84.2%	77.1%	99.8% × 85.9% =	85.7%
*1 from Ravindran *et al.* 2022a; *2 from Ravindran *et al.* 2022b.						

Table S3 and Figs. S1, S2 provides the accuracies of the ResNet50 field models and associated confusion matrices. The cross-validation accuracies and confusion matrices for ResNet34 and ResNet50 models are presented in Table S4 and Figs. S3 – S6.

The results of this investigation allow for multiple inferences about the impact of increasing class number and the viability of employing cascading model schemes.

Performance of the Consolidated Models

While it was reasonable to anticipate a sizable drop in predictive accuracy when combining the 22-class DP and 17-class RP label spaces, the consolidated 39-class model (39DP-RP) performed better than expected. When compared to the aggregate performance (85.9%) of the domain-based 22DP and 17RP models, the overall predictive accuracy (84.2%) of 39DP-RP fell by less than 2.0% (Table 9). Breaking that performance down by porosity domain, the table shows a decrease of less than 1.0% each for diffuse-porous and ring-porous domains. These data demonstrate that an increase in class number does not necessarily result in sizable reductions in predictive accuracy. In this study, that might be the case in part because the additional classes were fundamentally anatomically disparate. It is plausible that roughly doubling the number of classes in a model with anatomically similar classes might result in more substantial drops in top-1 accuracy, but until such studies are conducted, it remains an open question.

When consideration is given to the number of classes identified without error, 39DP-RP performed worse than the domain-based models. While 22DP and 17RP collectively identified 20 out of 39 classes perfectly (Ravindran *et al*. 2022a), 39DP-RP-SRP lagged at only 17 (Tables 5 and 6). These data show that although the predictive accuracy of the 39-class model decreased very little, it was confusing specimens among more classes, some of which were Type 3 misclassifications across porosity domains.

When the three classes of semi-ring-porous woods were added to the 39-class model to form a consolidated 42-class model, greater decreases in predictive accuracy were observed. When compared to the aggregate performance of the domain-based 22DP, 17RP, and 3SRP models (85.9%), the overall predictive accuracy of 42DP-RP fell by 8.8% (Table 9). Breaking that performance down by porosity domain, the table shows decreases of 10.9% and 7.5%, respectively, for diffuse-porous and ring-porous domains. The predictive accuracy for the semi-ring-porous woods remained unchanged at 100.0%. These data demonstrate that an increase in class number sometimes results in a sizable reduction in predictive accuracy, but interestingly the semi-ring-porous woods, which in broad terms of porosity domains are between ring-porous and diffuse-porous on the porosity continuum, did not provide either source or sink misclassifications (per Ravindran *et al*. 2022a). Rather, with semi-ring-porous classes present in the model, cross-domain errors between diffuse-porous and ring-porous classes increased, with several diffuse-porous woods misclassified as ring-porous woods, rather than being misclassified as semi-ring-porous woods. The anatomical similarity between the latewood of some ring-porous woods and diffuse-porous woods might account for this, but it suggests that the consolidated 42-class model is not placing sufficient weight on the presence of large earlywood vessels characteristic of ring-porous woods.

When consideration is given to the number of classes identified without error, 42DP-RP-SRP performed worse than the domain-based models. While 22DP, 17RP and 3SRP collectively identified 20 out of 42 classes perfectly, 42DP-RP-SRP lagged at only 17. These data show that the sizable decrease in predictive accuracy of the 42-class consolidated model was accompanied by confusion among more classes.

Viability of the Cascading Model Scheme

As long as the predictive accuracy of the root model in a cascade (in this case the 3POR model) is perfect – or close to perfect – cascading models seem to be a viable way to mitigate the loss of accuracy incurred by the consolidated models, especially 42DP-RP-SRP. As the overall accuracy of the cascading scheme is the mathematical product of predictive accuracies of the root and subsequent level models, less-than-nearly-perfect performance at the root could ruin the prospects of mitigation at deeper nodes. This would apply even more so to cascading model schemes deeper than two levels.

Another benefit of the cascading model scheme for classification of North American hardwoods is the reduction of cross-domain misidentifications for porosity. If the 3POR model were perfect and 22DP, 17RP and 3SRP exclusively comprise diffuse-porous, ring-porous, and semi-ring-porous classes, respectively, there would be no opportunity for cross-domain misidentification to occur. Of course, the misidentification of one *Populus* specimen in the 3POR model (Fig. 7) guarantees a cross-domain misclassification at the first level due to those same conditions. This is emphasizing the importance of perfect-to-near-perfect predictive accuracy of the root model. Within the context of this work the goal is to deploy already-developed models, thus the emphasis on a high-performing root level classifier. If one were developing an entirely new cascade of trained classifiers, classes that were misclassified by root- and first-level models could be included in all daughter nodes where they appeared. Such an approach is also used in the better-informed wood anatomical keys – taxa that a variably ring-porous or semi-ring-porous, for example, are included in both sections of the key (*e.g.* Arévalo and Wiedenhoeft 2022).

Implications for Future CVWID Model Design

While models with fewer classes might suffice for wood products manufactured from temperate logs harvested in the U.S.A., imported wood products necessitate the use of CVWID models that can identify a greater number of woods potentially from other continents including tropical woods, many of which lack growth rings altogether, thus rendering the 3POR model functionally irrelevant in the long term. As the need for expanding the taxonomic coverage of CVWID models grows, developers must discern ways to mitigate both predictive accuracy reduction and cross-domain misidentifications in larger models, and cascading model schemes might prove useful, but the extent to which they are scalable is unknown. Cascading model schemes also create a training problem in that when a new class is added, models at every tier of the decision tree might have to be retrained. To reduce that burden, developing highly accurate root and first-level classifiers targeting crisp categorical traits (*e.g.* vessels present *vs.* vessels absent, or even wood vs. not-wood as the root classifier) trained on a wide range of woods from around the world might provide the initial binning necessary to train smaller classifiers at deeper nodes in the cascade (even if some of those classifications are incorrect, as noted above). Such targeted and crisp anatomical feature classifiers, rather than taxonomic classifiers, can be highly influential in pooling and chaining trained classifiers from different studies along with application-specific lower tier (deep and shallow) classifiers (see He *et al*. 2024 for a conceptual commentary on using feature-based CVWID for taxonomic identification).

CONCLUSIONS

1. When compared to the aggregate performance of the domain-based 22DP and 17RP models, the overall predictive accuracy of the 39DP-RP model fell by less than 2.0%. Breaking that performance down by porosity domain, the results showed a decrease of less than 1.0% each for diffuse-porous woods and ring-porous woods. These data demonstrate that an increase in class number does not necessarily result in a sizable reduction in predictive accuracy; however, the consolidated 39-class model confused more classes than the domain-based models and committed several Type 3 cross-domain misclassifications for porosity.
2. When compared to the aggregate performance of the domain-based 22DP, 17RP, and 3SRP models, the overall predictive accuracy of the 42DP-RP model fell by 8.8%. Breaking that performance down by porosity domain, the table shows decreases of 10.9% and 7.5%, respectively, for diffuse-porous and ring-porous woods while the predictive accuracy of the semi-ring-porous woods remained unchanged at 100.0%. These data demonstrate that an increase in class number can result in a sizable reduction in predictive accuracy; moreover, the consolidated 42-class model confused more classes than the domain-based models and committed more than double the cross-domain misclassifications for porosity than that of the 39-class model.
3. The overall accuracy of the cascading model scheme exceeded that of both the consolidated 39DP-RP and 42DP-RP-SRP models and came in only slightly lower than that of the domain-based model scheme (<1.0%). The cascading model scheme also reduced the number of cross-domain misidentifications for porosity domain to one.
4. As long as the predictive accuracy of the 3POR model (or the any root model in a cascade) is perfect – or close to perfect – model cascading seems a viable way to mitigate the loss of accuracy incurred by the consolidated models, especially that of the 42DP-RP-SRP model, though the scalability of cascading models is as yet unknown.

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of U.S. Department of Agriculture (USDA), Research, Education, and Economics, Agriculture Research Service, Administrative and Financial Management, Financial Management and Accounting Division, Grants and Agreements Management Branch, under Agreement No. 58-0204-9-164, specifically for support of FO and RS. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the USDA. This material is also based upon work supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, McIntire Stennis project under accession number 7004014. This publication is a contribution of the Forest and Wildlife Research Center, Mississippi State University.

The authors wish to gratefully acknowledge the specimen preparation and imaging efforts of Adam Wade, Nicholas Bargren, Karl Kleinschmidt, Caitlin Gilly, Richard Soares, Brunela Rodrigues, and Flavio Ruffinatto.

The software apps for image dataset collection and trained model deployment along with the weights of the trained model will be made available at https://github.com/fpl-xylotron. A minimal data set can be obtained by contacting the corresponding author, but the full data set used in the study is protected for up to five years by a CRADA between FPL, UW-Madison and FSC.

All authors (Prabu Ravindran, PR; Frank Owens, FO; Adriana Costa, AC; Rubin Shmulsky, RS; Alex Wiedenhoeft, AW) contributed actionable feedback that improved the presentation of the paper. FO and RS provided access to and supervised data acquisition from the PACw test specimens. AC prepared a portion of the MADw and SJRw specimens. FO, AC, and AW established the wood anatomical scope of the study. PR implemented the machine learning pipelines for the study. PR and AW conducted data analysis and synthesis. PR, AW, AC and FO wrote the paper. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES CITED

Arévalo, R., Pulido, E. N. R., Solórzano, J. F. G., Soares, R., Ruffinatto, F., Ravindran, P., and Wiedenhoeft, A. C. (2021). "Image based identification of Colombian timbers using the XyloTron: a proof of concept international partnership," *Colombia Forestal*. 24(1), 5-16. DOI: 10.14483/2256201X.16700

Arévalo, R., and Wiedenhoeft, A. C. (2022). *Identificación de las Maderas de Centroamérica, México y el Caribe — Identification of Central American, Mexican, and Caribbean Woods.* General Technical Report FPL-GTR-293. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory. 376 p. DOI: 10.2737/FPL-GTR-293

Bilal, A., Jourabloo, A., Ye, M., Liu, X., and Ren, L. (2018). "Do convolutional neural networks learn class hierarchy?," *IEEE Transactions on Visualization and Computer Graphics* 24(1), 152-162. DOI: 10.1109/TVCG.2017.2744683

Brown, H. P., Panshin, A. J., and Forsaith, C. C. (1949). *Textbook of Wood Technology. Volume I. Structure, Identification, Uses, and Properties of the Commercial Woods of the United States*, Mc-Graw-Hill, New York.

DeVries, T., and Taylor, G. W. (2017). "Improved regularization of convolutional neural networks with cutout," (arXiv:1708.04552 [cs.CV]), Accessed 14 Feb 2024

Dormontt, E. E., Boner, M., Braun, B., Breulmann, G., Degen, B., Espinoza, E., Gardner, S., Guillery, P., Hermanson, J. C., and Koch, G. (2015). "Forensic timber identification: It's time to integrate disciplines to combat illegal logging," *Biological Conservation* 191, 790-798. DOI: 10.1016/j.biocon.2015.06.038

Esteban, L. G., Fernandez, F. G., Palacios, P. D., Romero, R. M., and Cano, N. N. (2009). "Artificial neural networks in wood identification: The case of two *Juniperus* species from The Canary Islands," *IAWA Journal* 30(1), 87-94. DOI:10.1163/22941932-90000206

Gasson, P. (2011). "How precise can wood identification can be? Wood anatomy's role in support of the legal timber trade especially CITES," *IAWA Journal* 32(2), 137-154. DOI: 10.1163/22941932-90000049

He, K., Zhang, X., Ren, S., and Sun, J. (2015). "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," *IEEE International Conference on Computer Vision*. DOI: 10.1109/ICCV.2015.123

He, X., Pelt, D., Gao, J., Gravendeel, B., Zhu, P., Chen, S., Qiu, J., and Lens, F. (2024). "Machine learning-based wood anatomy identification: Towards anatomical feature recognition," *IAWA Journal*. DOI: 10.1163/22941932-bja10157

Hermanson, J. C., and Wiedenhoeft, A. C. (2011). "A brief review of machine vision in the context of automated wood identification systems," *IAWA Journal* 32(2), 233-250. DOI: 10.1163/22941932-90000054

Hoadley, R. B. (1990). *Identifying Wood- Accurate Results with Simple Tools*, Taunton Press, Newtown, CT, USA.

Howard, J., and Gugger, S. (2020). "Fastai: A layered API for deep learning," *Information* 11, article 108. DOI: 10.3390/info11020108

Hwang, S. W., and Sugiyama, J. (2021). "Computer vision-based wood identification and its expansion and contribution potentials in wood science: A review," *Plant Methods* 17, 47. DOI: 10.1186/s13007-021-00746-1

Johnson, A., and Laestadius, L. (2011). "New laws, new needs: The role of wood science in global policy efforts to reduce illegal logging and associated trade," *IAWA Journal* 32(2), 125-136. DOI: 10.1163/22941932-90000048

Khalid, M., Lee, E. L. Y., Yusof, R., and Nadaraj, M. (2008). "Design of an intelligent wood species recognition system," *International Journal of Simulation: Systems, Science & Technology* 9(3).

Kingma, D. P., and Ba, J. (2015). "Adam: A method for stochastic optimization," (arXiv:1412.6980 [cs.LG]), Accessed 14 Feb 2024

Koch, G., Haag, V., Heinz, I., Richter, H., and Schmitt, U. (2015). "Control of international traded timber — The role of macroscopic and microscopic wood identification against illegal logging," *Journal of Forensic Research* 6, article 317. DOI: 10.4172/2157-7145.1000317

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). "Backpropagation applied to handwritten zip code recognition," *Neural Computation* 1(4), 541-551. DOI: 10.1162/neco.1989.1.4.541

Liu, S., Zheng, C., Wang, J., Lu, Y., Yao, J., Zou, Z., Yin, Y., and He, T. (2024). "How to discriminate wood of CITES-listed tree species from their look-alikes: Using an attention mechanism with the ResNet model on an enhanced macroscopic image dataset," *Frontiers in Plant Science* 15, article 1368885. DOI: 10.3389/fpls.2024.1368885

Lodewick, J. E. (1928). "Seasonal activity of the cambium in some northeastern trees," *Bulletin of the New York State College of Forestry* (No. 23-24).

Lowe, A. J., Dormontt, E., Bowie, M., Degen, B., Gardner, S., Thomas, D., Clarke, C., Rimbawanto, A., Wiedenhoeft, A. C., Yin, Y., and Sasaki, N. (2016). "Opportunities for improved transparency in the timber trade through scientific verification," *BioScience* 66 (11), 990-998. DOI: 10.1093/biosci/biw129

Nasirzadeh, M., Khazael, A. A., and bin Khalid, M. (2010). "Woods recognition system based on local binary pattern," in: *Proceedings of the Second International Conference on Computational Intelligence, Communication Systems and Networks, CICSyN 2010*, Liverpool, UK, pp. 308-313.

Owens, F. C., Ravindran, P., C., Costa, A., Chavesta, M., Montenegro, R., Shmulsky, R., and Wiedenhoeft, A. C. (2024) "Robustness of a macroscopic computer-vision wood identification model to digital perturbations of test images," *IAWA Journal*. Published online July 11, 2024.

Pan, S. J., and Yang, Q. (2010). "A survey on transfer learning," *IEEE Transactions on Knowledge Data Engineering* 22(10), 1345-1359. DOI: 10.1109/TKDE.2009.191

Panshin, A. J. and De Zeeuw, C. (1980). *Textbook of Wood Technology*, Fourth Edition, McGraw-Hill, New York.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). "Pytorch: An imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems* 2019, 8026-8037.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research* 12(85), 2825-2830. DOI: 10.48550/arXiv.1201.0490

Ravindran, P., Costa, A., Soares, R., and Wiedenhoeft, A. C. (2018). "Classification of CITES-listed and other neotropical Meliaceae wood images using convolutional neural networks," *Plant Methods* 14(1), article 25. DOI: 10.1186/s13007-018-0292-9

Ravindran, P., Ebanyenle, E., Ebeheakey, A. A., Abban, K. B., Lambog, O., Soares, R., Costa, A., and Wiedenhoeft, A. C. (2019). "Image based identification of Ghanaian timbers using the XyloTron: opportunities, risks, and challenges," in: *Proceedings 2019 Workshop on Machine Learning for the Developing World*, Vancouver, BC.

Ravindran, P., Thompson, B. J., Soares, R. K., and Wiedenhoeft, A. C. (2020). "The XyloTron: flexible, open-source, image-based macroscopic field identification of wood products," *Frontiers in Plant Science* 11, article 1015. DOI: 10.3389/fpls.2020.01015

Ravindran, P., Owens, F. C., Wade, A. C., Vega, P., Montenegro, R., Shmulsky, R., and Wiedenhoeft, A. C. (2021). "Field-deployable computer vision wood identification of Peruvian timbers," *Frontiers in Plant Science* 12, article 647515. DOI: 10.3389/fpls.2021.647515

Ravindran, P., and Wiedenhoeft, A.C. (2022). "Caveat emptor: On the need for baseline quality standards in computer vision wood identification," *Forests* 13, article 632. DOI: 10.3390/f13040632

Ravindran, P., Owens, F.C., Costa, A., Rodrigues, B.P., Chavesta, M., Montenegro, R., Shmulsky, R., and A.C. Wiedenhoeft. (2023). "Evaluation of test specimen surface preparation on computer vision wood identification," *Wood and Fiber Science* 55(2): 176-202. DOI: 10.22382/wfs-2023-15

Ravindran, P., Owens, F. C., Wade, A. C., Shmulsky, R., and Wiedenhoeft, A. C. (2022a). "Towards sustainable North American wood product value chains, Part I: Computer vision identification of diffuse-porous hardwoods," *Frontiers in Plant Science* 12, article 758455. DOI: 10.3389/fpls.2021.758455

Ravindran, P., Wade, A. C., Owens, F. C., Shmulsky, R., and Wiedenhoeft, A. C.

(2022b). “Towards sustainable North American wood product value chains, part 2: computer vision identification of ring-porous hardwoods,” *Canadian Journal of Forest Research* 52, 1014-1027. DOI: 10.1139/cjfr-2022-0077

Roth, F. (1895). “Timber: An elementary discussion of the characteristics and properties of wood,” Bulletin No. 10, Department of Agriculture, Forestry Division.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. (2015). “Imagenet large scale visual recognition challenge,” *International Journal of Computer Vision* 115(3), 211-252. DOI: 10.1007/s11263-015-0816-y

Shigei, N., Mandai, K., Sugimoto, S., Takaesu, R., and Ishizuka, Y. (2019). “Land-use classification using convolutional neural network with bagging and reduced categories,” in *Proceedings of the International MultiConference of Engineers and Computer Scientists 2019*, Hong Kong.

Smith, L. (2018). *A Disciplined Approach to Neural Network Hyper-parameters: Part 1 – Learning Rate, Batch Size, Momentum, and Weight Decay* (US Naval Research Laboratory Technical Report 5510-026), US Naval Research Laboratory, Washington, DC.

United Nations Office on Drugs and Crime (UNODC) (2016). *Best Practice Guide for Forensic Timber Identification*, New York.

Wang, B. H., Wang, H. J., and Qi, H. N. (2010). “Wood recognition based on grey-level co-occurrence matrix,” in: *Proceedings of the 2010 International Conference on Computer Application and System Modeling*, Taiyuan, China, pp. 269-272.

Wheeler, E. A., and Baas, P. (1998). “Wood identification – A review,” *IAWA Journal* 19(3), 241-264. DOI: 10.1163/22941932-90001528

Wiedenhoeft, A. C. (2020). “The XyloPhone: Toward democratizing access to high-quality macroscopic imaging for wood and other substrates,” *IAWA Journal* 41, 699-719. DOI: 10.1163/22941932-bja10043

Wiedenhoeft, A. C., Simeone, J., Smith, A., Parker-Forney, M., Soares, and R., Fishman, A. (2019). “Fraud and misrepresentation in retail forest products exceeds U.S. forensic wood science capacity,” *PloS One* 14, 1-13. DOI: 10.1371/journal.pone.0219917

Article submitted: July 12, 2024; Peer review completed: September 14, 2024; Revised version received: October 7, 2024; Accepted: October 16, 2024; Published: October 30, 2024.
DOI: 10.15376/biores.19.4.9741-9772

APPENDIX

Table S1. Taxa Used in Model Training

Acer macrophyllum
Acer negundo
Acer rubrum
Acer saccharinum
Acer saccharum
Aesculus californica
Aesculus glabra
Aesculus hippocastanum
Aesculus octandra
Alnus rhombifolia
Alnus rubra
Alnus rugosa
Alnus serrulata
Alnus tenuifolia
Arbutus menziesii
Arbutus texana
Asimina triloba
Betula alleghaniensis
Betula lenta
Betula nigra
Betula occidentalis
Betula papyrifera
Betula populifolia
Carpinus caroliniana
Carya aquatica
Carya cordiformis
Carya glabra
Carya illinoinensis
Carya laciniosa
Carya myristiciformis
Carya ovata
Carya texana
Carya tomentosa
Castanea dentata
Castanea pumila
Catalpa bignonioides
Catalpa speciosa
Celtis laevigata
Celtis occidentalis
Celtis reticulata
Cladrastis kentukea
Cladrastis lutea
Crataegus aestivalis
Crataegus assurgens
Crataegus compacti
Crataegus cordata
Crataegus cuneiformis
Crataegus douglasii
Crataegus macracantha
Crataegus mollis
Crataegus nitida
Crataegus rivularis
Crataegus rotundifolia
Crataegus spathulata
Crataegus succulenta
Crataegus tomentosa
Diospyros virginiana
Fagus grandifolia
Fraxinus americana
Fraxinus nigra
Fraxinus oregona
Fraxinus pennsylvanica
Fraxinus quadrangulata
Gleditsia aquatica
Gleditsia triacanthos
Gymnocladus dioica
Juglans cinerea
Juglans nigra
Liquidambar styraciflua
Liriodendron tulipifera
Maclura pomifera
Magnolia acuminata
Magnolia fraseri
Magnolia grandiflora
Magnolia macrophylla
Magnolia tripetala
Magnolia virginiana
Malus baccata
Malus coronaria
Malus domestica
Malus pumila
Malus rivularis
Malus sp.
Morus alba
Morus rubra
Nyssa aquatica
Nyssa biflora
Nyssa ogeche
Nyssa sylvatica
Ostrya virginiana
Oxydendrum arboreum
Platanus occidentalis
Populus angustifolia
Populus balsamifera
Populus deltoides
Populus fremontii
Populus grandidentata
Populus heterophylla
Populus tremuloides
Malus angustifolia
Populus trichocarpa
Prunus americana
Prunus angustifolia
Prunus avium
Prunus caroliniana
Prunus emarginata
Prunus myrtifolia
Prunus nigra
Prunus serotina
Pyrus ioensis
Quercus alba
Quercus arkansana
Quercus bicolor
Quercus coccinea
Quercus ellipsoidalis
Quercus falcata
Quercus georgiana
Quercus ilicifolia
Quercus incana
Quercus laevis
Quercus laurifolia
Quercus lyrata
Quercus macrocarpa
Quercus marilandica
Quercus michauxii
Quercus montana
Quercus myrtifolia
Quercus nigra
Quercus palustris
Quercus phellos
Quercus rubra
Quercus shumardii
Quercus stellata
Quercus texana
Quercus velutina
Rhamnus californica
Rhamnus cathartica
Rhamnus caroliniana
Rhamnus crocea
Rhamnus frangula
Rhamnus lanceolata
Rhamnus purshiana
Rhamnus tomentella
Robinia neo-mexicana
Robinia pseudoacacia
Salix laevigata
Salix lasiandra
Salix nigra
Salix nuttellii
Salix scouleriana
Sassafras albidum
Sorbus americana
Sorbus aucuparia
Sorbus decora
Tilia americana
Tilia caroliniana
Tilia floridana
Tilia heterophylla
Tilia pubescens
Ulmus alata
Ulmus americana
Ulmus crassifolia
Ulmus rubra
Ulmus serotina
Ulmus thomasii

Table S2. Taxa Used in Model Testing

Acer macrophyllum
Acer negundo
Acer rubrum
Acer saccharinum
Acer saccharum
Aesculus californica
Aesculus glabra
Aesculus octandra
Alnus incana
Alnus rhombifolia
Alnus rubra
Arbutus menziesii
Arbutus xalapensis
Asimina triloba
Betula lenta
Betula nigra
Betula occidentalis
Betula papyrifera
Betula populifolia
Carpinus caroliniana
Carya aquatica
Carya cordiformis
Carya glabra
Carya illinoinensis
Carya laciniosa
Carya ovata
Carya tomentosa
Castanea dentata
Catalpa speciosa
Celtis occidentalis
Celtis sp.
Cladrastis lutea
Crataegus aestivalis
Crataegus calpodendron
Crataegus douglasii
Crataegus mollis
Crataegus rivularis
Crataegus spathulata
Diospyros virginiana
Fagus grandifolia
Frangula purshiana
Fraxinus nigra
Fraxinus pennsylvanica
Fraxinus quadrangulata
Gleditsia triacanthos
Gymnocladus dioica
Juglans cinerea
Juglans nigra
Liquidambar styraciflua
Liriodendron tulipifera
Maclura pomifera
Maclura sp.
Magnolia acuminata
Magnolia fraseri
Magnolia grandiflora
Magnolia macrophylla
Magnolia tripetala
Magnolia virginiana
Malus angustifolia
Malus coronaria
Malus fusca
Malus pumila
Morus alba
Morus rubra
Nyssa aquatica
Nyssa ogeche
Nyssa sylvatica
Nyssa sylvatica var. *biflora*
Ostrya virginiana
Oxydendrum arboreum
Platanus occidentalis
Populus angustifolia
Populus balsamifera
Populus deltoides
Populus fremontii
Populus grandidentata
Populus heterophylla
Populus tremuloides
Populus trichocarpa
Prunus americana
Prunus angustifolia
Prunus caroliniana
Prunus emarginata
Prunus myrtifolia
Prunus serotina
Quercus alba
Quercus bicolor
Quercus coccinea
Quercus falcata
Quercus laurifolia
Quercus lyrata
Quercus macrocarpa
Quercus marilandica
Quercus montana
Quercus nigra
Quercus phellos
Quercus shumardii
Quercus velutina
Rhamnus crocea
Robinia pseudoacacia
Salix laevigata
Salix lasiandra
Salix nigra
Salix scouleriana
Sassafras albidum
Sassafras sp.
Sorbus americana
Sorbus decora
Tilia americana
Tilia americana var. *heterophylla*
Ulmus americana
Ulmus thomasii

Prediction Metrics for ResNet50 Field Model

These tables and figures show the accuracies and confusion matrices of the ResNet50 field model for the 39- and 42-class models (39DP-RP and 42DP-RP-SRP, respectively). The model performance metrics were created using the predictions of the ResNet50 field model (the model that was trained with all the training datasets) on the testing dataset obtained from the specimens in the PACw and teaching wood collections at Mississippi State University.

The accuracies of the ResNet50 models are in Table S3. Confusion matrices appear in Figs. S1 and S2.

Table S3. Top-1 and Top-2 Specimen Level Accuracies of ResNet50 Field Model

Label space	Top-1 Accuracy (%)	Top-2 Accuracy (%)
39DP-RP	75.9	87.3
42DP-RP-SRP	83.7	94.3

Top-1 and top-2 specimen level accuracies of ResNet50 field model. The accuracies were computed using the testing specimens sourced from Mississippi State University's PACw and teaching collections and were not used for model training.

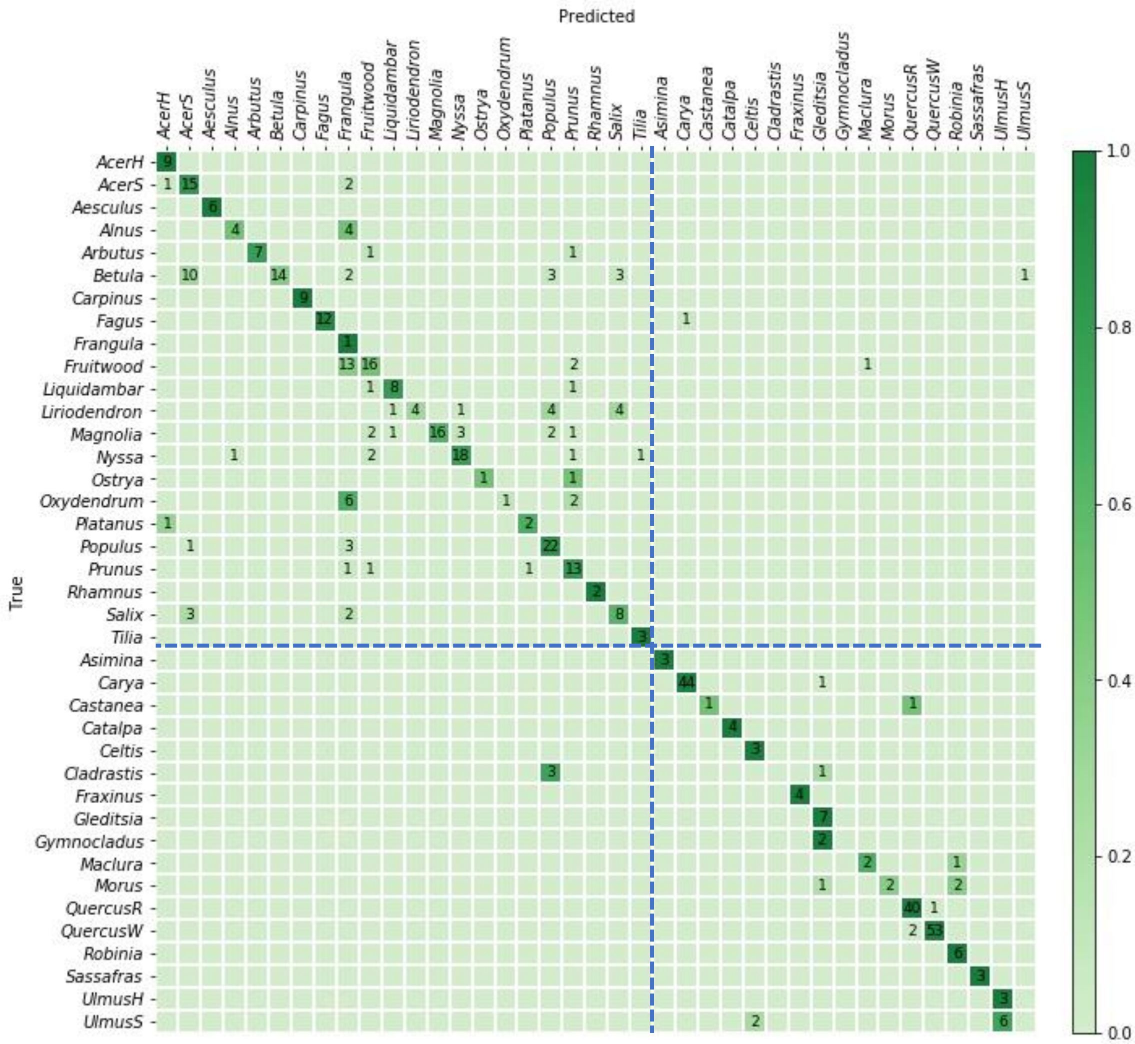

Fig. S1. Confusion matrix of ResNet50 field model on test specimens for the consolidated 39DP-RP model. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right and bottom left regions are cross-domain misclassifications for porosity. N = 482

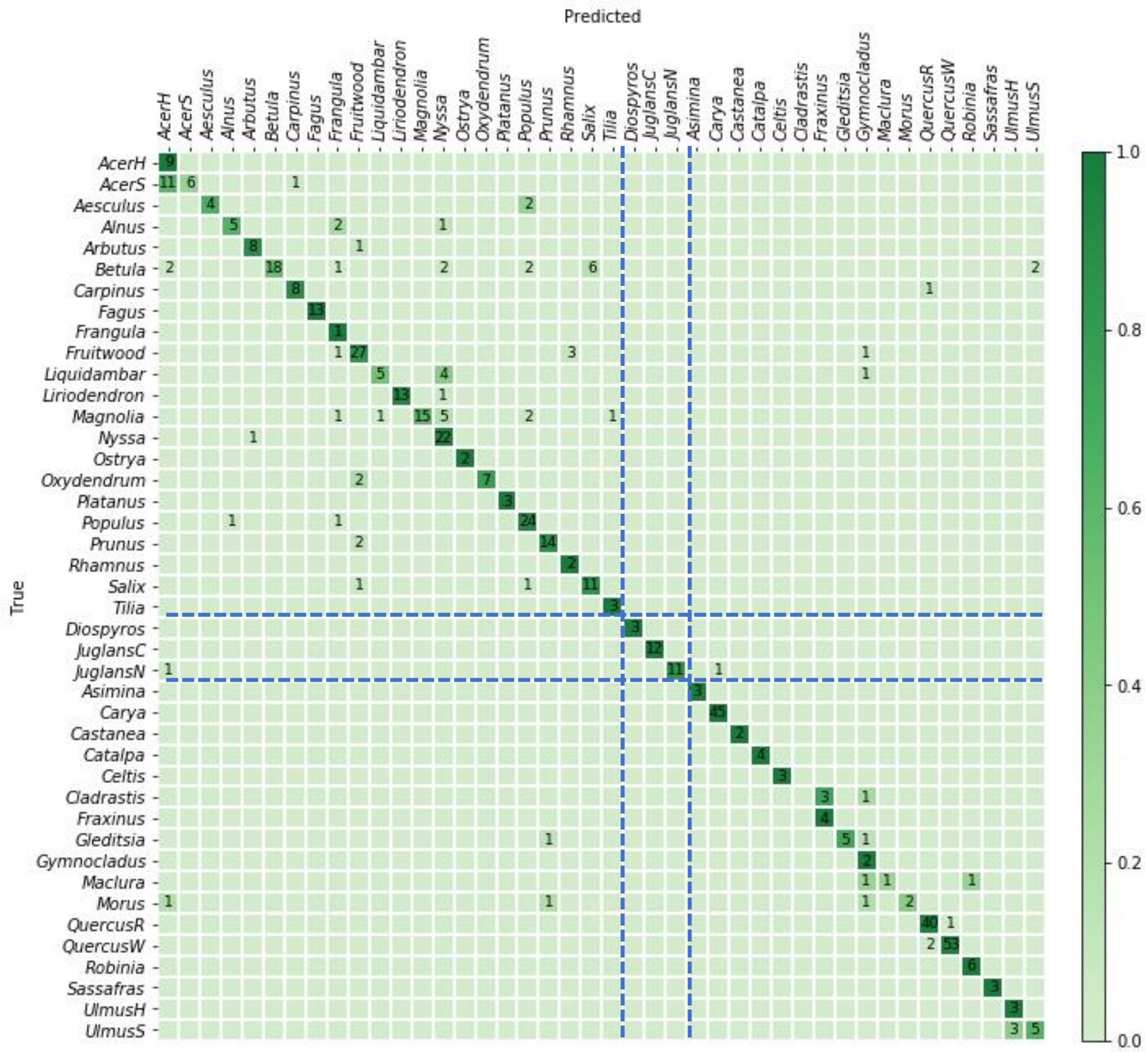

Fig. S2. Confusion matrix of ResNet50 field model on test specimens for label space 42DP-RP-SRP. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes, from Diospyros to JuglansN are semi-ring-porous classes, and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right and bottom left regions are cross-domain misclassifications for ring-porous and diffuse-porous porosity domains. There were no semi-ring-porous misclassifications. N= 510

5-fold Cross Validation Prediction Metrics

These tables and figures show the five-fold cross-validation accuracies and confusion matrices of the ResNet34 and ResNet50 backbones for the 39- and 42-class models (39DP-RP and 42DP-RP-SRP, respectively). The confusion matrices and accuracies are obtained by aggregating model predictions on the validation splits over the five folds.

The specimen level five-fold cross-validation accuracies of the ResNet34 and ResNet50 are shown in Table S4. Confusion matrices for the Top-1 results ResNet34 and ResNet50 appear in Figures S3, S4 and S5, S6, respectfully.

Table S4. Top-1 and Top-2 Specimen Level Accuracies for the Cross-validation Analysis

Label space	ResNet34 Accuracies (%)		ResNet50 Accuracies (%)	
	Top-1	Top-2	Top-1	Top-2
39DP-RP	93.9	97.8	92.2	95.6
42DP-RP-SRP	95.1	98.5	93.0	96.2

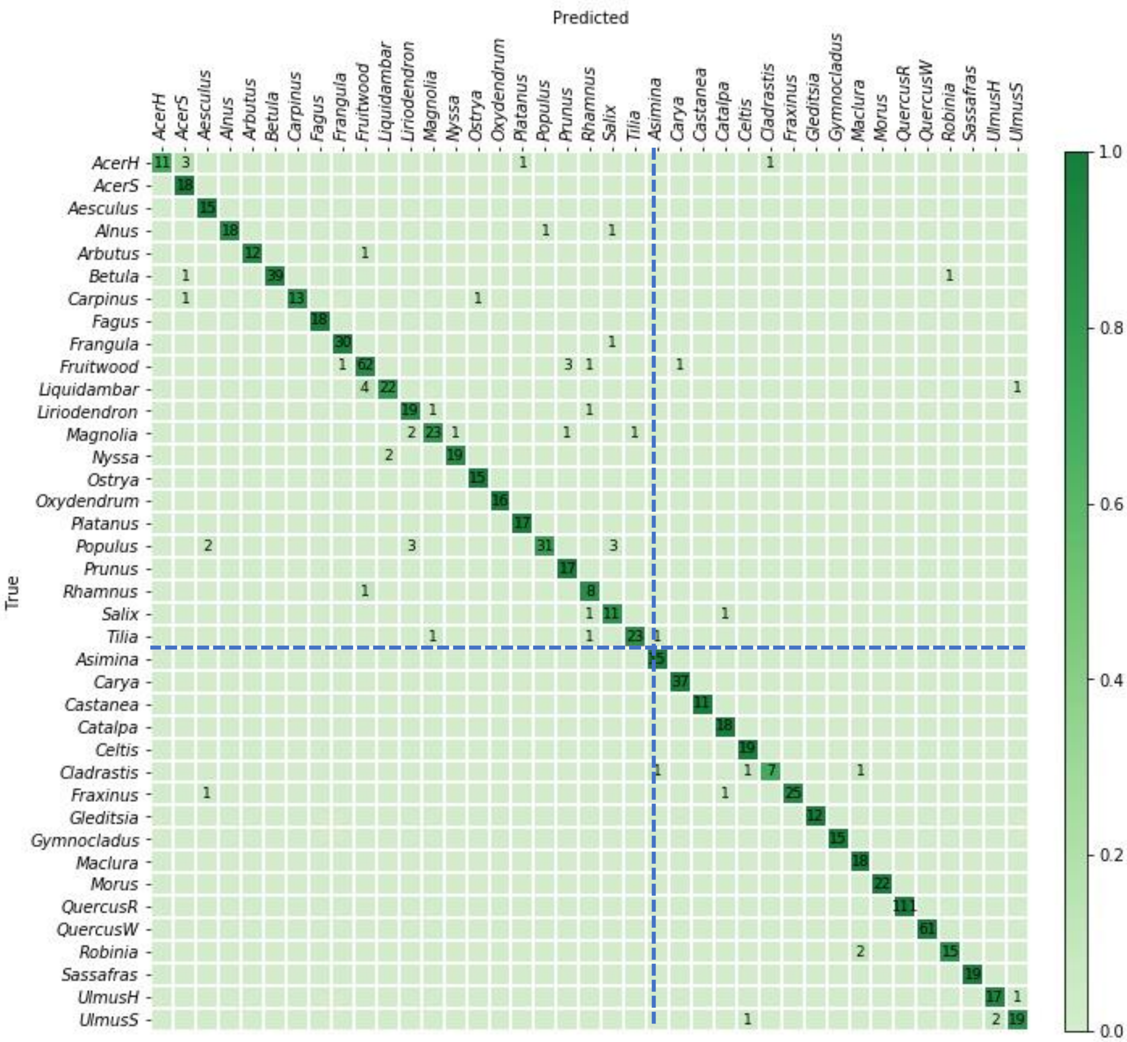

Fig. S3. Confusion matrix for specimen-level five-fold cross validation analysis of ResNet34 model for 39DP-RP model. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right and bottom left regions are cross-domain misclassifications for porosity. N = 956.

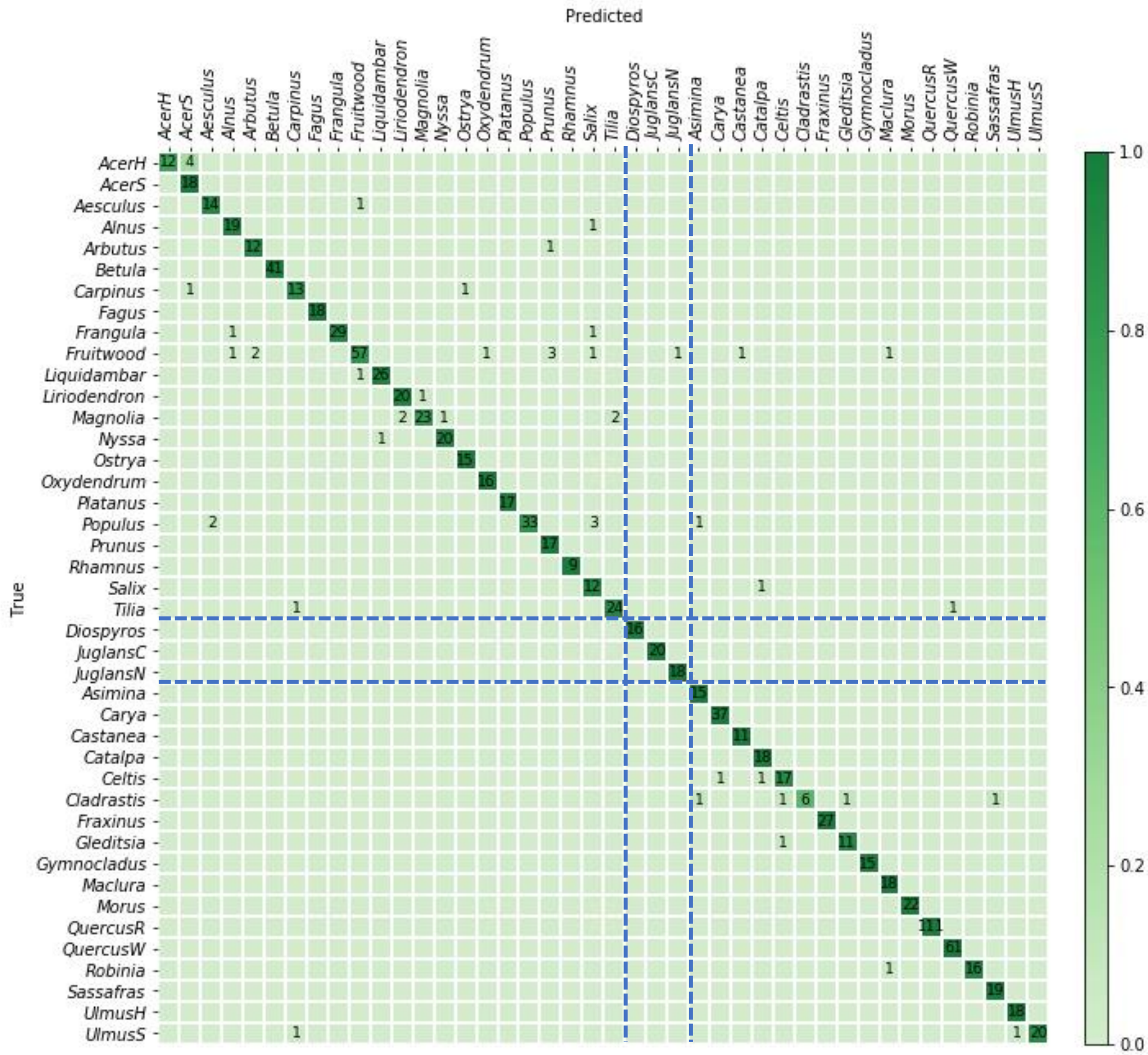

Fig. S4. Confusion matrix for specimen-level five-fold cross validation analysis of ResNet34 model for 42DP-RP-SRP. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes, from Diospyros to JuglansN are semi-ring-porous classes, and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right and bottom left regions are cross-domain misclassifications for ring-porous and diffuse-porous porosity domains. One specimen of fruitwood was misclassified as a semi-ring-porous wood. N = 1010.

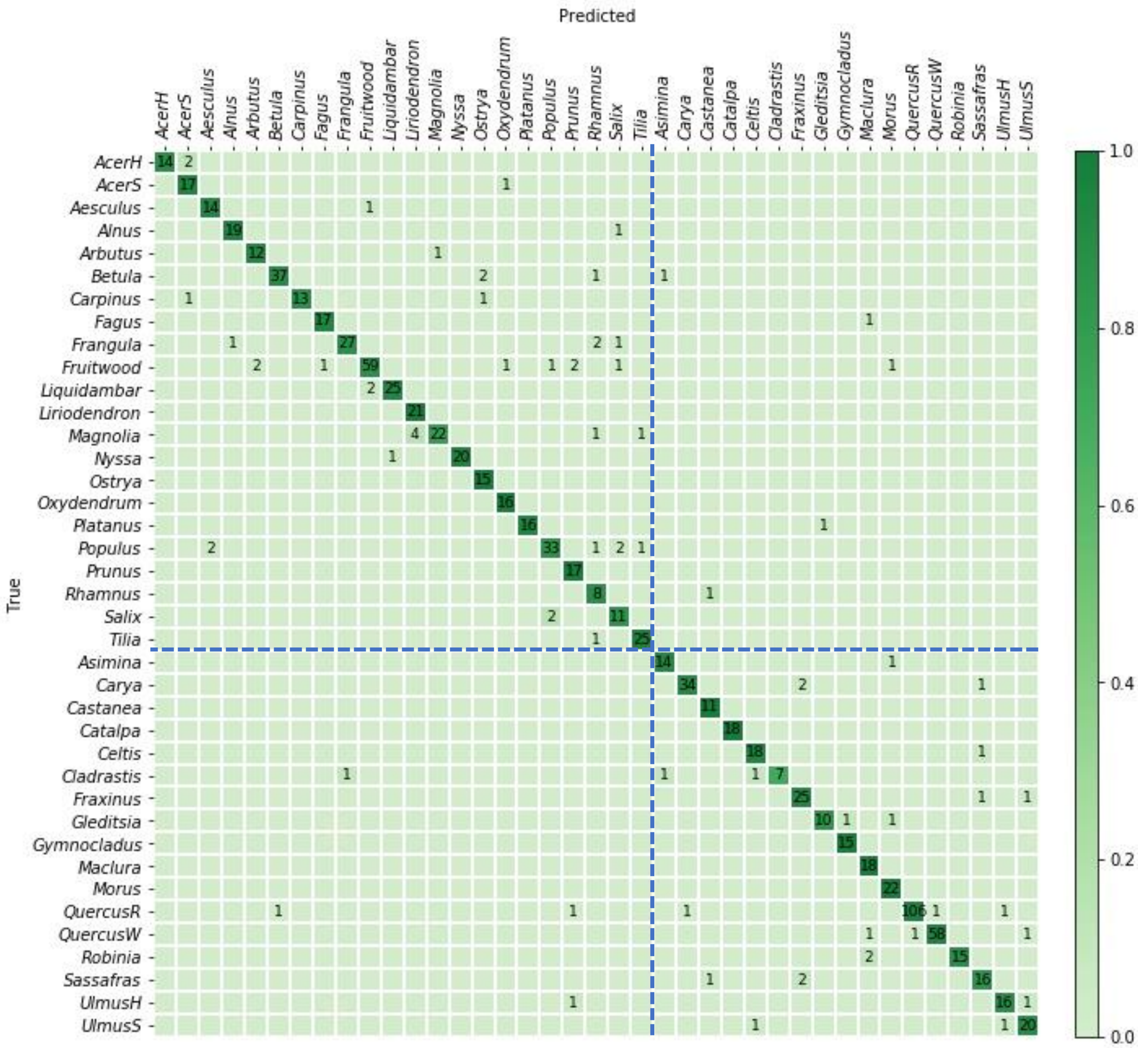

Fig. S5. Confusion matrix for specimen-level five-fold cross validation analysis of ResNet50 model for 39DP-RP model. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right and bottom left regions are cross-domain misclassifications for porosity. N = 956.

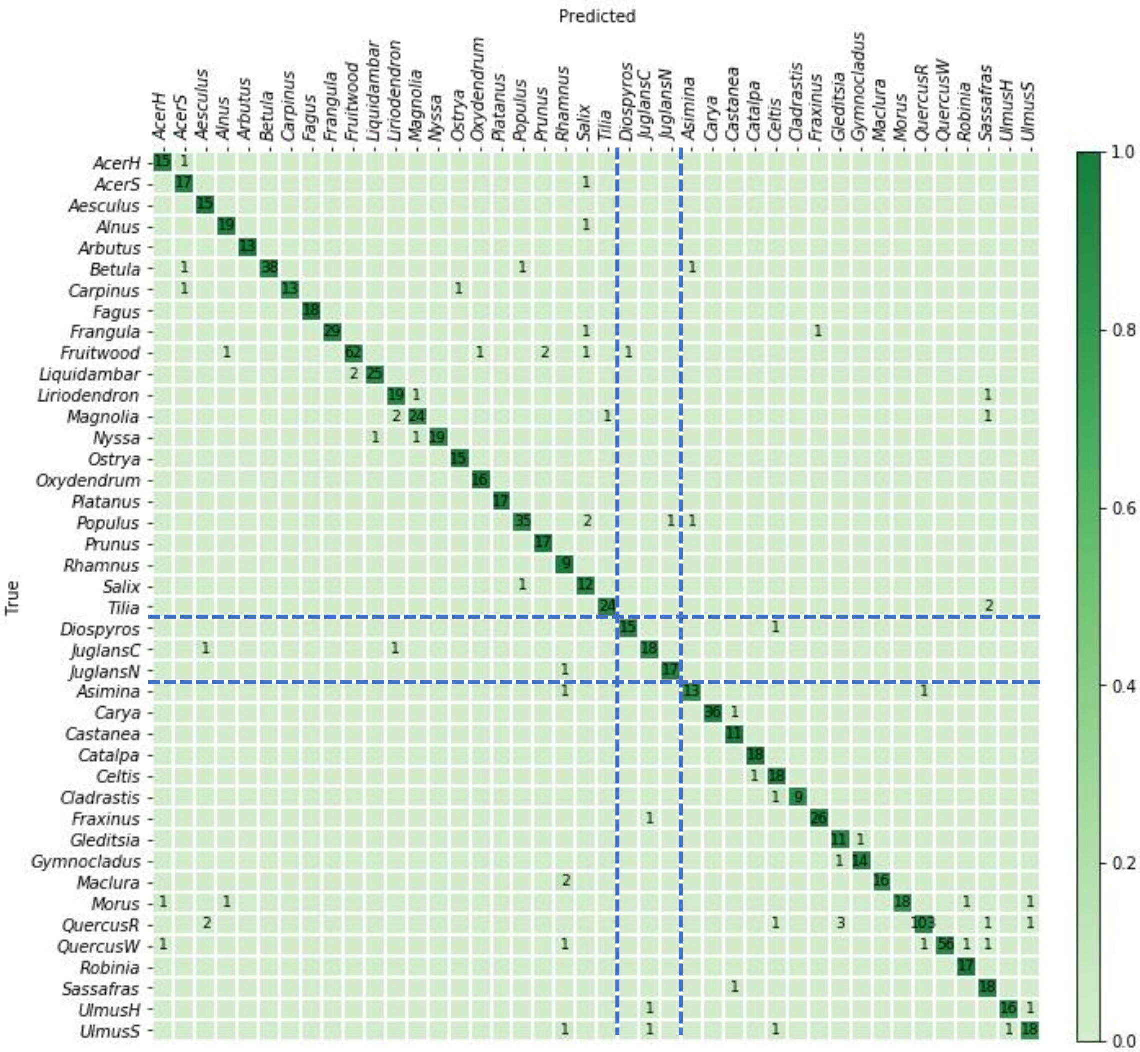

Fig. S6. Confusion matrix for specimen-level five-fold cross validation analysis of ResNet50 model for 42DP-RP-SRP. Counts in the diagonal cells indicate correct predictions. Off-diagonal counts indicate misclassifications. The blue dotted lines delineate porosity domains. From AcerH to Tilia are diffuse-porous classes, from Diospyros to JuglansN are semi-ring-porous classes, and from Asimina to UlmusS are ring-porous classes. Counts appearing in the top right and bottom left regions are cross-domain misclassifications for ring-porous and diffuse-porous porosity domains. Semi-ring-porous misclassifications are in the non-diagonal nearly central spaces between the dashed blue lines. N = 1010.