

Validation Study on the Practical Accuracy of Wood Species Identification *via* Deep Learning from Visible Microscopic Images

Te Ma,^a Fumiya Kimura,^a Satoru Tsuchikawa,^a Miho Kojima,^b and Tetsuya Inagaki^{a,*}

This study aimed to validate the accuracy of identifying Japanese hardwood species from microscopic cross-sectional images using convolutional neural networks (CNN). The overarching goal is to create a versatile model that can handle microscopic cross-sectional images of wood. To gauge the practical accuracy, a comprehensive database of microscopic images of Japanese hardwood species was provided by the Forest Research and Management Organization. These images, captured from various positions on wood blocks, different trees, and diverse production areas, resulted in substantial intra-species image variation. To assess the effect of data distribution on accuracy, two datasets, D1 and D2, representing a segregated and a non-segregated dataset, respectively—from 1,000 images (20 images from each of the 50 species) were compiled. For D1, distinct images were allocated to the training, validation, and testing sets. However, in D2, the same images were used for both training and testing. Furthermore, the influence of the evaluation methodology on the identification accuracy was investigated by comparing two approaches: patch evaluation and E2 image evaluation. The accuracy of the model for uniformly sized images was approximately 90%, whereas that for variably sized images it was approximately 70%.

DOI: 10.15376/biores.19.3.4838-4851

Keywords: Wood species identification; Microscopic cross-sectional images; Convolutional neural networks (CNN); Practical accuracy; Interactive platform; Web-based identification

Contact information: a: Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan; b: Forestry and Forest Products Research Institute, Matsunosato, Tsukuba 305-8687, Japan; *Corresponding author: inatetsu@agr.nagoya-u.ac.jp

INTRODUCTION

The accurate identification of wood species is important for efficient resource utilization and archaeological research. However, discerning species using microscopic cross-sectional images requires prior knowledge or experience related to the sizes and positions of vessel or tracheid in wood cells. In the context of addressing illegal logging practices and conducting comprehensive wood property analyses, there is a pressing contemporary requirement for the advancement and implementation of a machine learning-based systems dedicated to wood species identification. Deep learning has emerged as a dominant research trend worldwide and is applicable to classification, segmentation, and detection with high accuracy. It is characterized by a structure that mimics the neurons and synapses in the human brain. Each neuron receives information as an input and transmits the calculated data *via* a synapse (LeCun *et al.* 1998). Four dominant machine learning methods are employed in the deep learning domain: Convolutional neural networks (CNNs) for image recognition (Drakopoulos *et al.* 2021), recurrent neural

networks used for time-series data analysis (Sherstinsky 2020), autoencoders used for dimensionality reduction, and generative adversarial networks used for image generation (Goodfellow *et al.* 2014).

Among the methodologies utilized, CNNs have been used for the recognition of wood defects and the identification of wood species. Oktaria *et al.* (2019) reported an accuracy of 97% in the identification of 30 wood species from macroscopic images using a CNN *via* ResNet transfer learning. Likewise, an accuracy of $97.31 \pm 1.85\%$ was obtained *via* ResNet transfer learning when classifying 11 wood species of macroscopic images by Geus *et al.* (2021). They subsequently identified 281 wood species using DenseNet with an accuracy of 98.75% (Geus *et al.* 2020). Transfer learning is an effective approach for handling limited datasets (Sun *et al.* 2021). In their study, 25 species (120 for each species) were successfully identified with 99.6% accuracy. He *et al.* (2021) trained and evaluated nine CNN architectures using two macroscopic wood image datasets. Their proposed network achieved a 100% test rate on a dataset comprising eight wood species and 918 images after two rounds of training. In another dataset with 41 species and 11,984 images, it attained a 98.81% test recognition rate after three training cycles. Moulin *et al.* (2022) developed a custom deep CNN model to differentiate between images of Brazilian native and introduced wood species. The custom model achieved excellent accuracy (>0.90) and, in some cases, even surpassed human identification with an F1-score of 0.99. Lens *et al.* (2020) asserted that computers can differentiate between wood species by focusing on the corners and edges of tissues, such as vessel elements. Kwon *et al.* (2017) and Lopes *et al.* (2020) demonstrated through their respective studies that CNNs, utilizing transfer learning techniques, exhibit remarkable accuracy in the identification of wood species across datasets encompassing both hardwood and softwood specimens. Notably, one study achieved a classification accuracy of 97.32% when employing a CNN to classify the microscopic images of Brazilian wood specimens, encompassing 112 species (Hafemann *et al.* 2014). Kirbaş and Çifci (2022) delved into classifying wood species using the WOOD-AUTH dataset and assessing the effectiveness of different deep learning architectures including ResNet-50, Inception V3, Xception, and VGG19 with transfer learning. The dataset comprised macroscopic images of 12 wood species across the cross, radial, and tangential sections. The results indicated the superior performance of Xception, achieving a classification accuracy of 95.88%, surpassing that of the other models. Hwang and Sugiyama (2021) presented a thorough assessment of this subject and provided an essential foundation for the development of a framework for automatic wood identification. This study also highlighted the potential for expanding the use of computer vision in wood science, offering an insightful discourse on the future trajectory of the field. Similarly, noteworthy contributions have been made to the utilization of computer vision techniques for the identification of ring-porous hardwood species (Ravindran *et al.* 2022). That research highlights the significance of technological advancements in promoting sustainability within North American wood product value chains, while presenting a novel approach for the precise identification of specific types of wood. Fundamentally, these two studies made significant and essential contributions to the advancement of knowledge regarding wood species identification, with particular focus on the application of computer vision methodologies. These invaluable insights and directions guide the methodological and theoretical basis of this current research.

While numerous previous studies have reported high accuracy in wood species identification, many have not thoroughly addressed the specific factors that influence accuracy and robustness, such as image dimensionality and the distribution of data between

the training and testing sets. To develop a model for the identification of wood species based on microscopic cross-sectional images, it is essential to discern pertinent features, such as the position and size of vessels and tracheids in wood cells. However, these approaches ideally require consistent factors, including image resolution and calibration, when using microscopes with different magnification ratios.

The objective of this study was to assess the discriminatory capabilities of a CNN for classifying different wood species, irrespective of specific image characteristics. To achieve this objective, this study utilized images with various pixel sizes, enabling a comparison of model accuracy against scenarios in which only images of identical pixel sizes were analyzed. In addition, the impact of dataset partitioning (*i.e.*, training and test sets) on model accuracy was examined. Furthermore, this study presented a novel model capable of generating reasonably accurate estimations, regardless of the type of microscopic photograph employed. To facilitate practical implementation and promote collaborative efforts, the secondary goal was to establish a publicly accessible website that facilitated hardwood identification using customized image data.

EXPERIMENTAL

Image Processing

Microscopic cross-sectional images of 50 Japanese hardwood species (totaling 1,000 images, with 20 images per species) were sourced from the Japanese Wood Identification Database of the Forest Research and Management Organization. All images were captured using a D100 camera (Nikon, Tokyo, Japan) or a DP72 camera (Olympus). The 50 species included in the dataset, spanning 39 genera and 29 families, are summarized in Table 1. This means that the dispersion of anatomical features was quite high. When the database contained more than 20 images of any species, surplus images were harnessed to evaluate the robustness of the model. The images varied in terms of resolution (spanning $3,840 \times 3,072$; $3,200 \times 2,560$; $3,008 \times 2,000$; and $1,360 \times 1,024$). Although certain properties such as resolution are not preserved as digital values, one certainty is that if the TWTw No. differs, and it indicates a difference in individual specimens. The camera used varied across the images. This variation served as an appropriate test for the adaptability of the model to images of different sizes. In the process of data extraction from the images, the distance per pixel and field of view (FOV) on the X- and Y-axes varied according to the image resolution. The following parameters were observed for each resolution category:

For images with a resolution of $3,840 \times 3,072$ pixels, the distance per pixel was 0.9 μm , and the FOV was 3.5 mm and 2.8 mm on the X- and Y-axes, respectively.

For images with a resolution of $3,200 \times 2,560$ pixels, the distance per pixel was 0.9 μm , and the FOV was 2.9 mm and 2.3 mm on the X- and Y-axes, respectively.

For images with a resolution of $3,008 \times 2,000$ pixels, the distance per pixel was 1.3 μm , and the FOV was 3.9 mm and 2.6 mm on the X- and Y-axes, respectively.

For images with a resolution of $1,360 \times 1,024$ pixels, the distance per pixel was 3.2 μm , and the FOV was 4.4 mm and 3.3 mm on the X and Y axes, respectively.

Table 1. The 50 Species Included in the Dataset

SN* ¹	Sorted into 29 families	Scientific Name of Wood Species	NSI* ²	Sorted into 29 families	SN* ¹	Scientific Name of Wood Species	NSI* ²
1	Sapindaceae	<i>Acer sieboldianum</i>	2	Fagaceae	26	<i>Quercus stenophylla</i>	3
2	Actinidiaceae	<i>Actinidia arguta</i>	5		27	<i>Quercus serrata</i>	8
3	Anacardiaceae	<i>Rhus chinensis</i>	7		28	<i>Quercus crispula</i>	12
4		<i>Rhus trichocarpa</i>	3	Salicaceae	29	<i>Idesia polycarpa</i>	6
5	Aquifoliaceae	<i>Ilex macropoda</i>	0	Lauraceae	30	<i>Cinnamomum japonicum</i>	8
6	Araliaceae	<i>Aralia elata</i>	1		31	<i>Lindera erythrocarpa</i>	3
7	Betulaceae	<i>Betula grossa</i>	4	Moraceae	32	<i>Lindera umbellata</i>	2
8		<i>Carpinus laxiflora</i>	7		33	<i>Ficus erecta</i>	4
9		<i>Carpinus japonica</i>	4		34	<i>Morus australis</i>	12
10	Adoxaceae	<i>Viburnum furcatum</i>	0	Oleaceae	35	<i>Ligustrum japonicum</i>	3
11		<i>Viburnum dilatatum</i>	5	Rosaceae	36	<i>Pourthiaea villosa</i>	3
12	Celastraceae	<i>Euonymus oxyphyllus</i>	7		37	<i>Prunus grayana</i>	11
13	Clethraceae	<i>Clethra barbinervis</i>	10		38	<i>Prunus jamasakura</i>	19
14	Garryaceae	<i>Aucuba japonica</i>	0	Rutaceae	39	<i>Zanthoxylum piperitum</i>	0
15	Cornaceae	<i>Cornus kousa</i>	6	Hydrangeaceae	40	<i>Deutzia crenata</i>	13
16		<i>Cornus macrophylla</i>	6		41	<i>Hydrangea paniculata</i>	8
17		<i>Cornus controversa</i>	15		42	<i>Schizophragma hydrangeoides</i>	4
18	Daphniphyllaceae	<i>Daphniphyllum teijsmannii</i>	1	Stachyuraceae	43	<i>Stachyurus praecox</i>	19
19	Ericaceae	<i>Lyonia ovalifolia</i>	10	Styracaceae	45	<i>Styrax japonicus</i>	13
20		<i>Pieris japonica</i>	7	Staphyleaceae	44	<i>Euscaphis japonica</i>	10
21		<i>Rhododendron kaempferi</i>	9	Theaceae	46	<i>Camellia japonica</i>	11
22	Euphorbiaceae	<i>Mallotus japonicus</i>	10	Pentaphragmaceae	47	<i>Eurya japonica</i>	17
23	Eupteleaceae	<i>Euptelea polyandra</i>	5	Lamiaceae	48	<i>Callicarpa japonica</i>	11
24	Fagaceae	<i>Castanea crenata</i>	8		49	<i>Callicarpa mollis</i>	3
25		<i>Quercus acuta</i>	4		50	<i>Clerodendrum trichotomum</i>	0

*¹ Species number, *² Number of surplus images

Because these images were captured under a variety of microscope settings, they presented different resolutions and dimensionalities. Although challenging, these variable conditions helped test the robustness of this model in handling diverse and realistically inconsistent data. Because each image featured a scale bar in the bottom-right corner, the corresponding pixels were cropped out. The images were then processed, as shown in Fig. 1, and the previously mentioned image dimensions were reduced to $3,840 \times 2,591$, $3,200 \times 2,153$, $3,008 \times 2,000$, and $1,360 \times 972$, respectively. Subsequently, square areas corresponding to $2,501 \times 2,501$, $2,153 \times 2,153$, $2,000 \times 2,000$, and 972×972 pixels were randomly extracted from each image category and resized to 640×640 pixels. It is important to note that the exact scale lengths between pixels vary across images. Next, 64×64 nonoverlapping grid patches were extracted from these images, accumulating a total of 100 patches per image. These patches were divided into training, validation, and testing datasets. The RGB color of each patch was converted to grayscale to minimize the impact of differences in microscope devices.

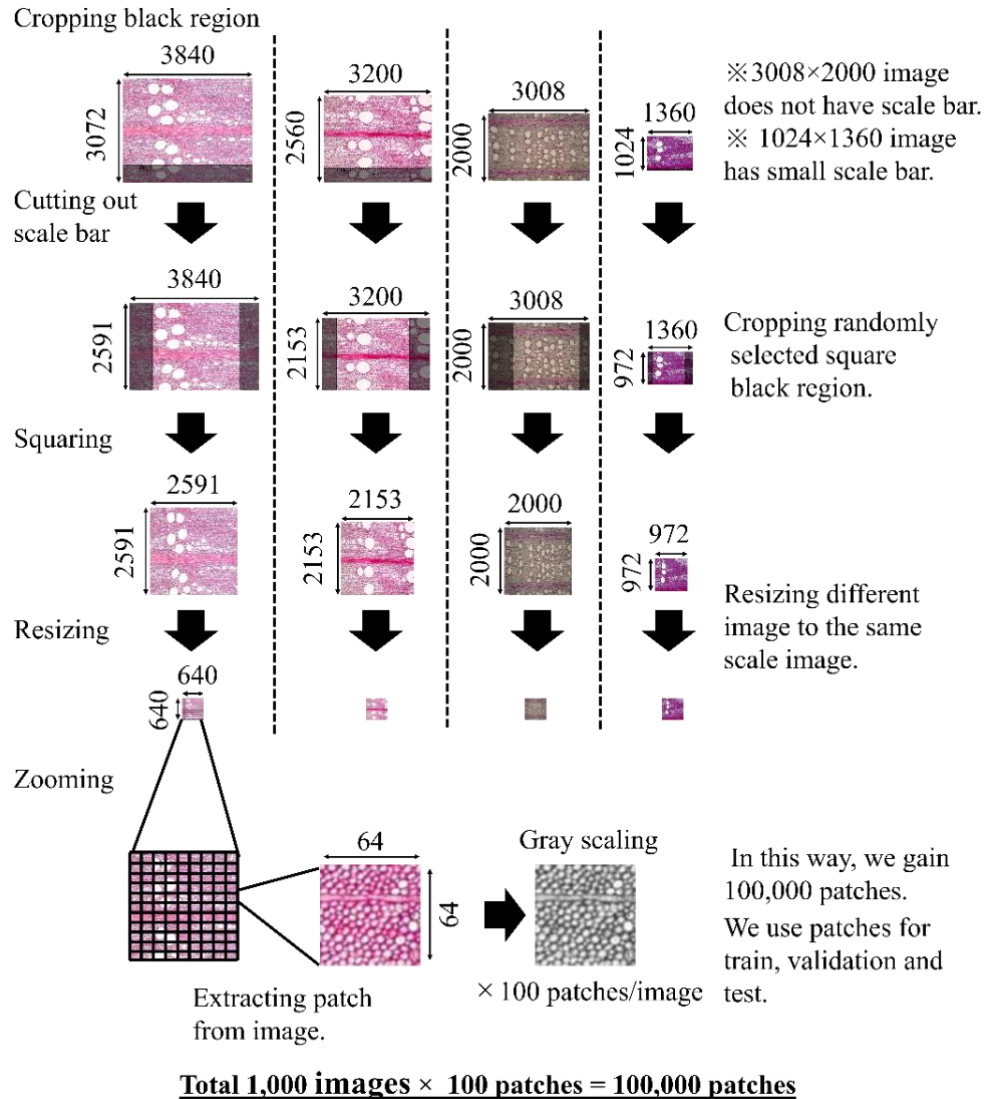


Fig. 1. Diagram of the image preprocessing

To ascertain the effect of image size on wood identification accuracy, an additional dataset was composed of images from 10 species (denoted in orange in Table 1) in their original dimensions ($3,840 \times 3,720$ pixels), processed in line with the main dataset. The present study utilized this diverse dataset to assess the practical identification performance of various CNN models. A fundamental principle of machine learning is the exclusion of identical images from the training, validation, and testing sets. To elucidate the impact of the data distribution, data allocation was evaluated using two prepared datasets, as shown in Fig. 2. Two datasets, D1 and D2—from 1,000 images (20 images per each of the 50 species) were assembled. For D1, entirely distinct images were used for training, validation, and testing. Conversely, D2 used the same images for both training and testing. In addition, two evaluation methodologies were compared: patch evaluation (E1) and image evaluation (E2). For the former, species predictions were made for each patch, and accuracy was calculated accordingly. In the latter scenario, following the species prediction for each patch, the cumulative accuracy was computed by considering the summation of all patches within the image.

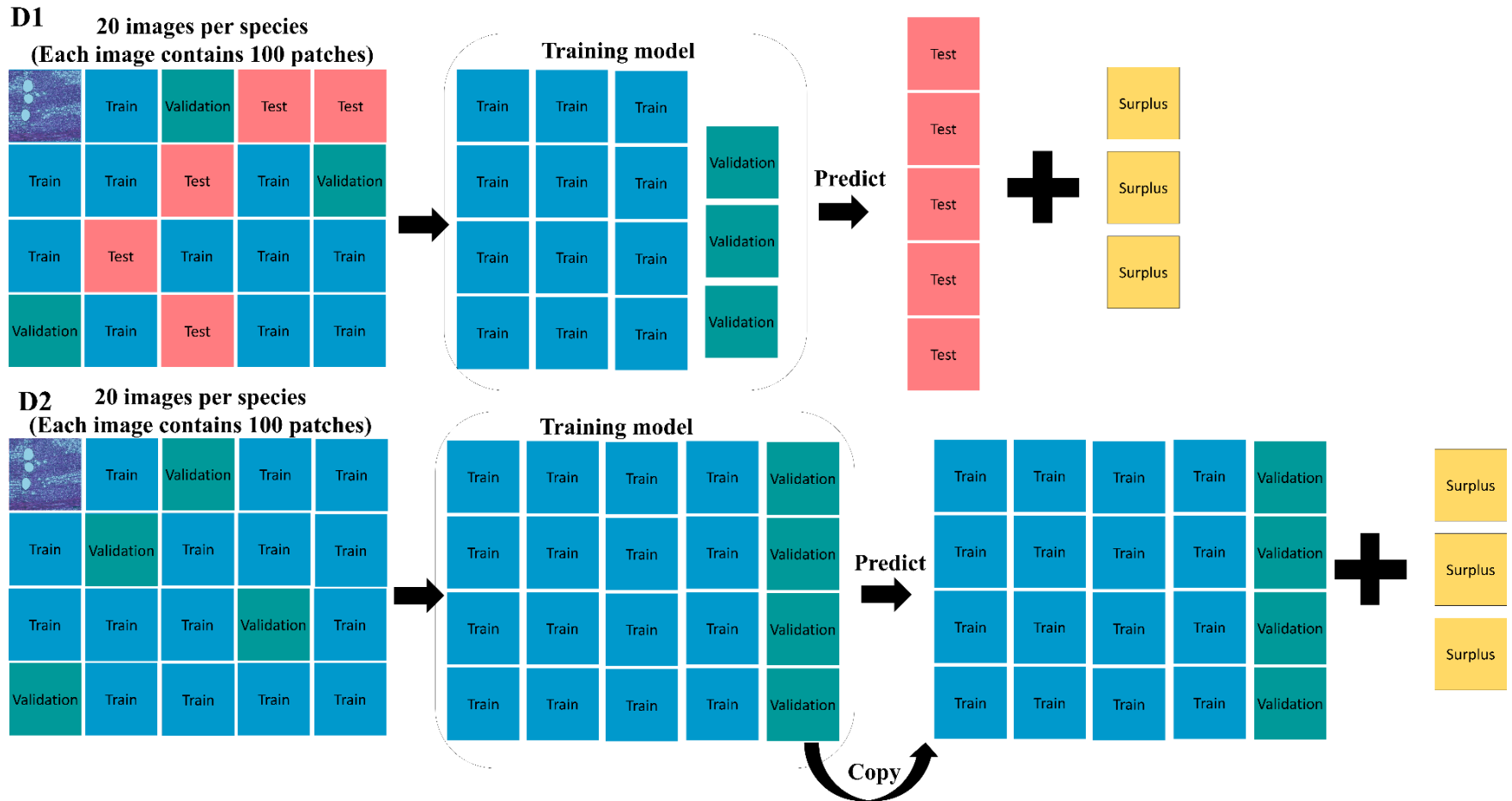


Fig. 2. Data allocation of D1 and D2

The potential for fine-tuning was also considered in the exploration of accuracy enhancement. For D1, 20 images from each species were apportioned from the training (12 images = 1,200 patches), validation (three images = 300 patches), and testing (five images = 500 patches) datasets without overlap. The impact of variations in location and camera settings across images on identification accuracy was assessed. For D2, 20 images from each species were allocated to the training (16 images = 1,600 patches) and validation (four images = 400 patches) datasets. After the development of the CNN models, these 20 images were reused for testing. The findings from this study suggest superior accuracy in wood identification from D2 compared to D1. However, model robustness should be higher in D1 than in D2. To evaluate model robustness, surplus data corresponding to species with more than 20 images in the original database were prepared. The quantity of surplus data for each species is shown in Table 1. To evaluate the accuracy of the model, it is crucial to use a test set in which the number of samples across each category is equal. However, achieving such a balanced representation is challenging in the context of real-world scientific data. Considering these practical limitations, this study was not restricted to a perfectly balanced test set. Instead, all remaining samples were used as surplus data to evaluate the robustness of the model. This approach enabled assessment of how well the model performed on a broader scale and reflected its true capacity for generalization and adaptability. It is crucial to emphasize that different TWTw No. signifies distinct individuals from which samples were extracted. A comprehensive summary of TWTw No. for the database used in this study is available at <https://inatetsu2nd-woodspecrecog-01-home-f1zh5g.streamlit.app/>.

Network Architecture

In this study, a CNN structure was employed (Hafemann *et al.* 2014). The rationale for this choice stems from its remarkable accuracy in the specific task of wood species identification. Such a high level of precision in this context is astonishing, making their approach particularly compelling. This structure incorporates a 64×64 input layer, two convolutional layer sets with 5×5 sliding kernels and a stride of one pixel, a pooling layer with 3×3 kernels and a stride of two pixels, two locally connected layers featuring 3×3 sliding kernels and a stride of one pixel, and a flattened layer that shares weights across all the nodes. The output layer encompasses 50 classes, each representing a different wood species. The Adam algorithm was used to automatically optimize the learning rate. All convolutional layers employed a Rectified Linear Unit (ReLU) activation function with a batch size of 512. After ReLU activation, any input values that were not positive were translated to an output of zero. The final dense layer uses SoftMax activation, thereby producing output values within the zero to 1.0 range. All the models were trained over 22 epochs until the validation loss stabilized. The authors scrutinized the loss progression when assessing the model results. Each dataset was trained for approximately 5 min using a GeForce GTX 1080 GPU (NVIDIA, USA). To evaluate the model accuracy, fine-tuning was applied by building on VGG16 with weights derived from ImageNet (ResNet50). This fine-tuning merges 15 layers of a preexisting fixed model with new additions, facilitating the learning of new weights and enhancing the generalization capacity of the model. Because fine-tuning relies on three-dimensional input models, grayscale conversion was not applied to the images.

Accuracy Evaluation

As delineated in section “Image Processing”, each species was assigned 500 testing patches for D1 and 2,000 testing patches for D2. Two evaluation methods were used in this study. For Evaluation Method 1 (E1), the species from all patches were predicted and used to calculate accuracy, meaning that the predicted outcomes from all patches were considered when determining accuracy. Thus, 100 predictions corresponding to each patch were derived from each image. The accuracy of all predictions was calculated, thereby providing a comprehensive view of the model's performance across the entirety of each image. In Evaluation Method 2 (E2), a sum rule was implemented, as shown in Table 2, wherein the sum of all probabilities associated with patches for each image was maximized. Consequently, the probabilities generated by the CNN, which corresponded to all the species, were summed over the patches to categorize the wood species. In other words, a single prediction value was generated for each image. This single prediction represents the model's assessment of the entire given image. To assess the impact of the pixel size on the accuracy of the CNN model, the extracted files were tested at pixel sizes of 16, 32, 64, 128, and 256. The results revealed no dramatic improvements in the accuracy for pixel sizes greater than 64. Consequently, 64 pixels were selected for the study.

Table 2. Sum Rule for Accuracy Calculation

	Class1	Class1	Class1	...	Class50
Patch1	0.5*	0.03	0.01	...	0.001
Patch2	0.7*	0.01	0	...	0.001
...
Patch99	0.1	0.5*	0.1	...	0.01
Patch100	0.8*	0.1	0.1	...	0.2
Sum of predict value	↓	↓	↓	↓	↓
Image1	57*	3	2	...	0.9

* Predicted class by patch or image

RESULTS AND DISCUSSION

Effect of Evaluation Method on Accuracy

Table 3 presents the accuracy of the predictions derived from both D1 and D2 datasets. Notably, E2 consistently outperformed E1. Consequently, the prediction accuracy was enhanced by aggregating the outcomes of 100 patches extracted from each image. The E2 evaluation used in this study closely mirrors the method utilized by Hafemann *et al.* (2014), who stated that "For the recognition, patch results are combined for the entire image. The straightforward solution is to solely use the central patch of the image for testing, but this yields suboptimal results, as patches are smaller than the images. In this work, we consider the sum rule: the prediction for a given test image is the class that maximizes the sum of the probabilities on all patches of the image." Nevertheless, the prediction accuracy achieved in this study was relatively low (44% for 50 species with 20 images per class) compared to that (97% for 112 species with 20 images per class) reported by Hafemann *et al.* (2014). This discrepancy can be attributed to the authors' distinct patch extraction method, which extracts 100 patches from each image, whereas Hafemann *et al.* (2014) extracted a single patch per training epoch from each image. However, it is not anticipated that a significant improvement in the accuracy will be gained if an exact approach is adopted. Additionally, considerable fluctuation in accuracy was observed,

depending on how the data were split between the training and test sets. This is a common phenomenon in the field of data science, and it underlines the importance of partitioning data for model training and evaluation. Notably, this model must be evaluated using a dataset divided in a manner similar to that of the D1-divided dataset. This partitioning strategy ensures a more reliable assessment of the model's real-world applicability and robustness.

Table 3. The Accuracy of Predictions Derived from Both D1 and D2 Datasets

Number of Classes	Model	Dataset	Evaluation	Accuracy (%)	
				Test Set	Surplus Set
50 (Species)	Hafemann <i>et al.</i> (2014)	D1-divided	E1-Patch	28	30
			E2-Image	44	45
		D2-non-divided	E1-Patch	44	33
			E2-Image	70	52

Effect of Data Allocation Scheme on Accuracy

A comparison of the accuracy between datasets D1 and D2 revealed two noteworthy insights (Fig. 3), in which D2 consistently outperformed D1 across both evaluation methods; however, the accuracy associated with D2 experienced a substantial decrease when tested on surplus data, whereas D1's accuracy remained nearly unchanged. This underlines the fundamental rule of machine learning: the robustness of a model cannot be ensured if the same data are used for training and testing. Therefore, the authors used D1 in all subsequent evaluations and analyses. When the accuracy for each species was observed, the species with the highest classification accuracy include *Pourthiaea villosa* (73%) for D1 under method E1, *Castanea crenata* and *Quercus stenophylla* (100%) for D1 under method E2, *Daphniphyllum teijsmannii* (83%) for D2 under E1, and *Daphniphyllum teijsmannii* (100%) for D2 under E2.

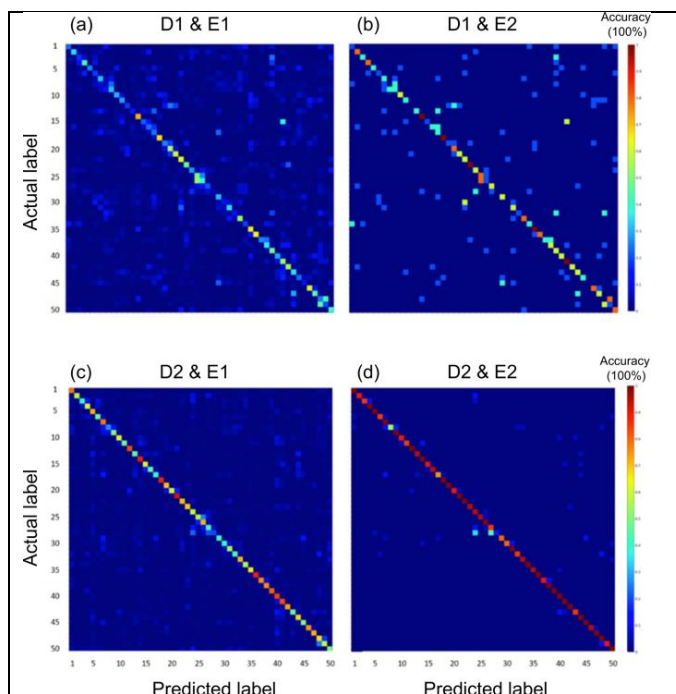


Fig. 3. Comparison of the accuracy between datasets D1 and D2

Accuracy Comparison Among CNN Structures

In line with the suggestion of Hafemann *et al.* (2014), the model was fine-tuned based on VGG16 to compare the accuracy achieved with and without fine-tuning. The results are summarized in Table 4. In all the instances, the highest accuracy was obtained using the fine-tuned model. Using the second evaluation method (E2), the fine-tuned model achieved a testing accuracy of 72% and surplus accuracy of 71%. In machine learning analysis, especially when the internal mechanisms are largely unknown—often referred to as a "black box"—the number of samples or corpus size plays a pivotal role. This principle holds true across fields, but in the realm of scientific data, acquiring a large-scale, or "big data" set is a hard task. Considering these challenges, an achievable sample size of one thousand specimens was used for model construction. This may seem modest compared to models trained on larger databases, but is sufficient for the level of complexity inherent in wood science and for making reliable predictions. Transparency is of paramount importance in this approach.

Accuracy for Genes and Family Prediction

An additional CNN model was developed to predict the genera and families of the wood samples. As shown in Table 1, the 50 wood species investigated in this study spanned 39 genera and 29 families, and the corresponding accurate results are outlined in Table 4. Interestingly, the accuracy of the fine-tuned model using the E2 evaluation method did not improve significantly. When the species were grouped by genus and family, the rate of correct identification increased in some instances, decreased in others, and showed no change on average. Notably, the rate of correct identification increased significantly for the ring-porous species. This likely arises from the fact that ring-porous species within a given genus or family share similar characteristics, thus exhibiting clear distinctions between different classes.

Effect of Original Image Dimensionality on Accuracy

To assess the impact of the original image size on the accuracy of wood species identification, the authors assembled a dataset of images from 10 species, each of which had the same size (3,840 × 3,720 pixels). The dataset was processed according to the methodology depicted in Figs. 1 to 3. An additional dataset corresponding to the same species was curated, albeit with diverse image sizes. It should be noted that standardizing the microscopic magnification across all images is challenging because many images are collected at undisclosed magnifications. As anticipated, the fine-tuned model under the E2 evaluation strategy yielded a higher accuracy (94% for the test set and 83% for the surplus set) when working with standardized image settings compared with those with varied dimensions (86% for the test set and 75% for the surplus set). The attained accuracy of 94% in this study mirrors the findings reported by Hafemann *et al.* (2014), where 97% accuracy was used for classification across 112 species, each represented by 20 images. Consequently, these accuracy results demonstrated a reasonably high level of performance, even when there was considerable size variability among the original images.

Estimation of Practical Accuracy

To gauge the practical accuracy of wood species identification from microscopic images *via* a CNN, the impact of various factors, including the evaluation method, data allocation, parameter tuning, classification target, and image dimensions, was assessed. Surplus data were compiled to examine the robustness of the model.

Table 4. Accuracy Comparison among CNN Structures

Dataset	The Number of Classes	Pixel of Original Image	Evaluation	Model	Accuracy (%)	
					Test Set	Surplus Set
D1-Divided	50 (species)	4 kinds* ¹	E1-patch	VGG16 (Fine-tuning)	51	51
				Hafemann <i>et al.</i> (2014)	28	30
			E2-image	VGG16 (Fine-tuning)	72	71
				Hafemann <i>et al.</i> (2014)	44	45
	39 (genus)	4 kinds* ¹	E1-patch	VGG16 (Fine-tuning)	56	45
				Hafemann <i>et al.</i> (2014)	35	28
			E2-image	VGG16 (Fine-tuning)	76	64
				Hafemann <i>et al.</i> (2014)	54	44
	29 (family)	4 kinds* ¹	E1-patch	VGG16 (Fine-tuning)	55	45
				Hafemann <i>et al.</i> (2014)	29	21
			E2-image	VGG16 (Fine-tuning)	78	64
				Hafemann <i>et al.</i> (2014)	43	31
	10 (species)	4 kinds* ¹	E1-patch	VGG16 (Fine-tuning)	70	64
				Hafemann <i>et al.</i> (2014)	47	43
			E2-image	VGG16 (Fine-tuning)	86	75
				Hafemann <i>et al.</i> (2014)	70	56
1 kind* ²		E1-patch	VGG16 (Fine-tuning)	79	71	
			Hafemann <i>et al.</i> (2014)	55	47	
		E2-image	VGG16 (Fine-tuning)	94	83	
			Hafemann <i>et al.</i> (2014)	74	63	

*¹ Original images of 3,840 × 3,072 pixels, 3,200 × 2,560 pixels, 3,008 × 2,000 pixels, 1,360 × 1,024 pixels were used.

*² Only original images of 3,840 × 3,072 pixels were used.

Despite D2 having a higher testing accuracy than D1, further tests on surplus data revealed a lack of robustness. Fine-tuning based on VGG16 consistently led to higher accuracy than the original parameter setup.

No significant improvement in accuracy was observed when predicting genera or families, as opposed to species. Consistent with the findings of Hafemann *et al.* (2014), increased accuracy was observed when using a single image size. However, the current study found prediction accuracies reaching 86% (for the testing set) and 75% (for the surplus set), despite variations in the dimensions of the original images. To validate if the constructed CNN model is universally applicable, further samples are required. Consequently, the authors launched a website that enables anyone to identify wood species using their own photos. The website employs two configurations of the fine-tuned VGG16 model using D1 and E2: one designed for predicting 50 species, and the other for predicting 10 species from images of equal dimensions. Their accuracies for the test dataset were 72% and 94 %, respectively. The second goal is to facilitate opportunities for many people to attempt tree species identification through this website, and the incorporation of accumulated data into the model will lead to further performance enhancements. By opening the authors' process and providing these resources, it not only adds to the global repository of scientific knowledge, but also facilitates advancements in wood science at the grassroots level.

CONCLUSIONS

This study aimed to estimate the practical accuracy of identifying Japanese wood species from microscopic images using a CNN.

1. Assessments were made based on various factors, including the evaluation methodology, data allocation strategy, CNN structure, classification objective, and original image size. The overall practical accuracy for the identification of the 50 Japanese wood species stands at approximately 70%. Moreover, it was demonstrated that the constructed model could be used to classify images procured at varying magnification levels.
2. To promote the broader use of tree species identification, the authors created a model using VGG16 on their website. Users can view the outputs from two distinct models: one developed with 50 species, and the other with 10 species of uniform dimensions. The accuracy of the test datasets for these models are 72% and 94%, respectively. It is believed that for tree species identification to become universally applicable, it is imperative that the model be adapted to any type of microscope used.

Therefore, this study is considered a continuing process, and the authors aim to further validate its practical accuracy through the collection and analysis of predicted results from the website. It is anticipated that the data accumulated from this interactive platform will contribute significantly to the ongoing refinement and improvement of the authors' identification model.

DECLARATIONS

Availability of Data and Materials

The datasets generated and/or analysed during the current study are available in the database repository, [<http://db.ffpri.affrc.go.jp/woodDB/TWTwDB/home.php>]. Anyone can use the CNN models and information of the TWTw No. used for training, validation and test data set are also available from the URL: <https://inatetsu2nd-woodsprecog-01-home-f1zh5g.streamlit.app/>.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge the use of image data from Forest Research and Management Organization Japanese Wood Identification Database, <http://db.ffpri.affrc.go.jp/woodDB/TWTwDB/home.php>).

The authors would like to acknowledge financial support from JSPS (KAKENHI, No. 26850111).

REFERENCES CITED

- Drakopoulos, F., Baby, D., and Verhulst, S. (2021). "A convolutional neural-network framework for modelling auditory sensory cells and synapses," *Communications Biology* 4, article 827. DOI: 10.1038/s42003-021-02341-5
- Geus, A. R., Silva, S. F., Gontijo, A. B., Silva, F. O., Batista, M. A., and Souza, J. R. (2020). "An analysis of timber sections and deep learning for wood species classification," *Multimedia Tools and Applications* 79, 34513-34529. DOI: 10.1007/s11042-020-09212-x
- Geus, A. R., Backes, A. R., Gontijo, A. B., Albuquerque, G. H. Q., and Souza, J. R. (2021). "Amazon wood species classification: A comparison between deep learning and pre-designed features," *Wood Science and Technology* 55, 857-872. DOI: 10.1007/s00226-021-01282-w
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). "Generative adversarial nets," *Advances in Neural Information Processing Systems* 27, 680-2672. DOI: 10.48550/arXiv.1406.2661
- Hafemann, L. G., Oliveira, L. S., and Cavalin, P. (2014). "Forest species recognition using deep convolutional neural networks," Paper presented at: *22nd International Conference on Pattern Recognition*, Stockholm, Sweden, pp. 1103-1107. DOI: 10.1109/ICPR.2014.199
- He, T., Mu, S., Zhou, H., and Hu, J. (2021). "Wood species identification based on an ensemble of deep convolution neural networks," *Wood Res.* 66, 1-14. DOI: 10.37763/WR.1336-4561/66.1.0114
- Hwang, S. W., and Sugiyama, J. J. (2021). "Computer vision-based wood identification and its expansion and contribution potentials in wood science: A review," *Plant Methods* 17, article 47. DOI: 10.1186/s13007-021-00746-1

- Kırbaç, İ., and Çifci, A. (2022). “An effective and fast solution for classification of wood species: A deep transfer learning approach,” *Ecol. Inform.* 69, article 101633. DOI: 10.1016/j.ecoinf.2022.101633
- Kwon, O., Lee, H. G., Lee, M. R., Jang, S., Yang, S. Y., Park, S. Y., Choi, I. G., and Yeo, H. (2017). “Automatic wood species identification of Korean softwood based on convolutional neural networks,” *Journal of Korean Wood Science Technology* 45(6), 797-808. DOI: 10.5658/WOOD.2017.45.6.797
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE* 86(11), 324-2278. DOI: 10.1109/5.726791
- Lens, F., Liang, C., Guo, Y., Tang, X., Jahanbanifard, M., Silva, F. S. C., Ceccantini, G., and Verbeek, F. J. (2020). “Computer-assisted timber identification based on features extracted from microscopic wood sections,” *IWA Journal* 41(4), 660-680.
- Lopes, D. J. V., Burgreen, G. W., and Entsminger, E. D. (2020). “North American hardwoods identification using machine-learning,” *Forests* 11(3), article 298. DOI: 10.3390/f11030298
- Moulin, J. C., Lopes, D. J. V., Mulin, L. B., Bobadilha, G. D. S., and Oliveira, R. F. (2022). “Microscopic identification of Brazilian commercial wood species via machine-learning,” *Cerne* 28, article e-102978. DOI: 10.1590/01047760202228012978
- Oktaria, A. S., Prakasa, E., Suhartono, E., Sugiarto, B., Prajitno, D. R., and Wardoyo, R. (2019). “Wood species identification using convolutional neural network (CNN) architectures on macroscopic images,” *Journal of Information Technology and Computer Science* 4, 274-283. DOI: 10.25126/jitecs.201943155
- Ravindran, P., Wade, A. C., Owens, F. C., Shmulsky, R., and Wiedenhoef, A. C. (2022). “Towards sustainable North American wood product value chains, part 2: Computer vision identification of ring-porous hardwood,” *Canadian Journal of Forest Research* 52, 1014-1027. DOI: 10.1139/cjfr-2022-0077
- Sherstinsky, A. (2020). “Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network,” *Physica D: Nonlinear Phenomena* 404, article ID 132306. DOI: 10.1016/j.physd.2019.132306
- Sun, Y., Lin, Q., He, X., Zhao, Y., Dai, F., Qiu, J., and Cao, Y. (2021). “Wood species recognition with small data: A deep learning approach,” *International Journal of Computational Intelligence Systems* 14(1), 1451–1460. DOI: 10.2991/ijcis.d.210423.001

Article submitted: January 8, 2024; Peer review completed: February 11, 2024; Revised version received and accepted: February 25, 2024; Published: May 31, 2024.
DOI: 10.15376/biores.19.3.4838-4851