Check for updates

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

Shoujia Liu[1,2], Chang Zheng[1,2], Jiajun Wang[1,2,3], Yang Lu[1,2], Jie Yao[4], Zhiyuan Zou[4], Yafang Yin[1,2] and Tuo He[1,2,5]*

[1]Department of Wood Anatomy and Utilization, Research Institute of Wood Industry, Chinese Academy of Forestry, Beijing, China, [2]Wood Collections, Chinese Academy of Forestry, Beijing, China, [3]National Centre for Archaeology, Beijing, China, [4]Beijing Information Science and Technology University, Beijing, China, [5]Wildlife Conservation Monitoring Center, National Forestry and Grassland Administration, Beijing, China

Introduction: Global illegal trade in timbers is a major cause of the loss of tree species diversity. The Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) has been developed to combat the illegal international timber trade. Its implementation relies on accurate wood identification techniques for field screening. However, meeting the demand for timber field screening at the species level using the traditional wood identification method depending on wood anatomy is complicated, time-consuming, and challenging for enforcement officials who did not major in wood science.

Methods: This study constructed a CITES-28 macroscopic image dataset, including 9,437 original images of 279 xylarium wood specimens from 14 CITES-listed commonly traded tree species and 14 look-alike species. We evaluated a suitable wood image preprocessing method and developed a highly effective computer vision classification model, SE-ResNet, on the enhanced image dataset. The model incorporated attention mechanism modules [squeeze-and-excitation networks (SENet)] into a convolutional neural network (ResNet) to identify 28 wood species.

Results: The results showed that the SE-ResNet model achieved a remarkable 99.65% accuracy. Additionally, image cropping and rotation were proven effective image preprocessing methods for data enhancement. This study also conducted real-world identification using images of new specimens from the timber market to test the model and achieved 82.3% accuracy.

Conclusion: This study presents a convolutional neural network model coupled with the SENet module to discriminate CITES-listed species with their look-alikes and investigates a standard guideline for enhancing wood transverse image data, providing a practical computer vision method tool to protect endangered tree species and highlighting its substantial potential for CITES implementation.

1 Introduction

As an integral ecosystem component, trees assume a pivotal role in purifying our atmosphere, serving as habitats for diverse organisms and actively combating climate change (Rivers, 2017). In the intricate fabric of Earth's ecological system, they are indispensable entities, ensuring the vitality and equilibrium of our planet's ecological balance. When a tree is harvested, the wood can be used in every stage of human society, such as papermaking, construction, and furniture manufacturing (Falk, 2009; Latib et al., 2020). Approximately 73,000 tree species exist worldwide, but excessive logging and timber over-exploitation have resulted in up to 30% of the world's tree species being at risk of extinction, along with biodiversity destruction, soil erosion, and other ecological problems (Brancalion et al., 2018; Cazzolla et al., 2022; IUCN, 2022). This threat is true for some tropical tree species, which are widely used in producing high-value furniture, musical instruments, and handicrafts due to their excellent physical processing properties and beautiful patterns (Wick, 2019; Atikah et al., 2021). Thus, driven by high profits, some tree species are being over-harvested, which has been recorded as extinct/extinct in the wild, i.e., *Lachanodes arborea* (Lambdon and Ellick, 2016). Benefits from illegal logging are estimated to account for 15%–30% of the international timber trade, accounting for $51 to $152 billion a year (INTERPOL, 2021).

In recent decades, the world has witnessed a distressing decline in global tree populations. To prevent international trade from threatening the survival of endangered wildlife, the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) came into effect on 1 July 1975 to ensure that international trade in wild animal and plant specimens does not threaten species survival by subjecting international trade in selected species to certain controls (Appendices I, II, and III), according to their needed degree of protection (Goldsmith, 1978). In November 2022, the 19th Conference of the Parties to the CITES was held in Panama, and numerous tropical tree species were newly listed in CITES appendices. To date, more than 34,310 plant species of 134 genera have been included in the CITES appendices, including approximately 670 tree species, and 80% of these species are internationally traded for their timbers (CITES, 2023). Thus, a fast and accurate wood identification method is needed to support CITES implementation and promote legal logging.

Wood, or secondary xylem, is composed of countless cells of different shapes, sizes, and arrangements, with a complex and anisotropic structure (Shmulsky and Jones, 2019). Species within the same genus has similar appearance and even wood anatomical structure. Traditional wood identification, the most mainstream recognition approach, relies on human examination of the anatomical features of wood samples and refers to an identification standard list of the macroscopic and microscopic characters compiled by the International Association of Wood Anatomists (IAWA) (Committee, 1989; Yin et al., 2022). This task can only be completed accurately by experienced wood anatomists who familiarize with wood anatomy and identification, with the help of identification tools and reference materials. From a general point of view, it is arduous for the traditional wood identification method to reach species-level discrimination by observing anatomical features. To compensate for the shortcomings of traditional wood identification methods and break through the wood identification bottleneck at the species level, some techniques, i.e., DNA barcoding (Jiao et al., 2020), mass spectrometry (Deklerck et al., 2019; Price et al., 2021), and near-infrared spectroscopy (Bergo et al., 2016; Pan et al., 2021), have been developed. However, the lack of reference data and the high cost to establish it have limited applications of these approaches.

Computer vision is an interdisciplinary field at the intersection of computer science and image processing that aims to bridge the gap between human visual perception and machines by endowing computers with the capacity to understand, interpret, and extract knowledge from digital visual data (Liu et al., 2017). Image classification is a fundamental computer vision task that attempts to comprehend an entire image to classify images by assigning them to a specific label. With the rapid development of computer vision research and computer hardware performance, several neural network architectures have been proposed for image classification (Wang et al., 2019). Advancements in deep learning and convolutional neural networks (CNNs) have enabled more accurate and robust object and pattern identification with visual data. It is also very appealing to many wood anatomists and has been widely used in wood classification. The automated wood

identification method combines deep learning and computer vision to extract structural features and detect key information hidden in wood images (Voulodimos et al., 2019; Hwang and Sugiyama, 2021).

CNNs are the most commonly used computer vision-based wood identification models. As research continues, an array of neural network architectures, including distinguished models such as LeNet, AlexNet, ResNet, and GoogLeNet, have been introduced to address the wood image classification task (Kwon et al., 2017; Ravindran et al., 2018; Oktaria et al., 2019). The performance of these convolutional architectures is boosted by increasing their depth while maintaining their gradient information. In several studies, ResNet has shown superior classification performance compared to other models (He et al., 2021; Wu et al., 2021). However, a wood image is a fine-grained texture image with the characteristics of large intraclass variation and small interclass variation. To accurately determine the wood species, spatial and channel information should be exploited more delicately. Recently, the squeeze-and-excitation network (SENet) was proposed to provide the unit with a mechanism to explicitly model dynamic, nonlinear dependencies between channels using global information (Hu et al., 2018). This mechanism enables the network to selectively amplify or suppress specific feature maps, improving model performance. In this way, SENet can be embedded in the CNN, and the model can be trained to achieve better results.

In many cases, obtaining wood images from reliable sources is difficult, especially for globally regulated tree species, resulting in an unsatisfactory number of high-quality images to satisfy the modeling requirements. Therefore, in the case of limited wood image datasets, transfer learning and data augmentation are effective methods for solving the problem of insufficient data volume (Dyk and Meng, 2001; Ravindran et al., 2018). Kırbaş and Çifci (2022) compared the impacts of several deep learning architectures, namely, ResNet-50, Inception V3, Xception, and VGG19, based on the WOOD-AUTH dataset. They found that Xception performs remarkably well in the transfer learning domain. Hengshuo et al. proposed a wood identification algorithm based on an improved residual CNN, which augments the data based on the self-similarity of wood cross-sectional macrostructure and uses an improved residual CNN model, i.e., ResNet101, based on block gradient weighting to extract the features of each sub-image (Su et al., 2021).

Although there have been many wood species identification studies based on computer vision (Ravindran et al., 2018; Ravindran and Wiedenhoeft, 2020), two main factors still limit the development and application of this technique: the self-collected wood image data from different sources before use have not been adequately processed, and the identification accuracy is mainly affected by the sensitivity of the model to the slight variability in the wood species. Room for improvement in the models used in the existing studies remains.

This study aimed to develop a fast and reliable computer vision-based deep learning model by exploiting spatial and channel information to discriminate CITES-listed tree species from their look-alikes. The specific aims of this study were to (1) construct a CITES-28 (14 commonly traded CITES-listed tree species and 14 of their look-alikes) wood transverse surface image dataset, while concurrently investigating the optimal image data processing approach from the perspective of data enhancement; (2) establish a state-of-the-art SE-ResNet model by embedding the SENet module in a CNN (ResNet); and (3) discriminate CITES-listed tree species from their look-alikes using SE-ResNet in real-world identification.

2 Materials and methods

2.1 Data preparation and augmentation

In this section, we first explain how to prepare and enhance the wood image data in the experimental preparation stage from the perspective of data enhancement. Second, based on the enhanced data, we select the currently commonly used CNN for model training. Finally, the model performance is evaluated for tree species identification.

2.1.1 Image dataset collection

In this study, 279 verified specimens of 14 commercially important CITES-listed species and 14 of their look-alike species that were often mixed with CITES-listed species in trade were collected from the Wood Collection of the Chinese Academy of Forestry (CAFw), the USDA Forest Products Laboratory Wood Collection of Madison (MADw), and the Samuel J. Record Collection (SJRw). The transverse surfaces of the wood samples were sanded at grits of 180, 240, 400, 800, and 1,000 to obtain a clear surface for image acquisition. Their macroscopic transverse images are shown in Figure 1. The nonoverlapping images of 2,048 × 2,048 pixels, representing 6.35 × 6.35 mm of tissue, were taken with a XyloTron (Ravindran et al., 2020). A total of 9,437 original images were collected to build a CITES-28 dataset of 14 CITES-listed tree species and 14 of their look-alikes based on previous studies (Table 1) (Ravindran et al., 2018; Yin et al., 2022).

2.1.2 Dataset partitioning and patch dataset creation

We divided the dataset before creating patches to avoid dividing the same image into both the training and testing sets. The CITES-28 dataset of the original 9,437 images was divided into 80%/10%/10% training/validation/test splits at the image level. To ensure that the errors are representative of the entire dataset, 10-fold cross-validation for each model was used and accuracy is reported as the average over the 10 folds. An image dataset with a high imbalance results in poor classification performance; hence, six kinds of patches were extracted from the CITES-28 dataset images in this study. More overlap exists in patches of classes (species) with low quantities of images to maintain a balanced distribution of classes (species) in our dataset. The details of the patch dataset used for training and testing are listed in Table 2.

2.1.3 Image turning

We constructed a specific data augmentation method to fully train the model according to the data characteristics, as shown in Equation 1.
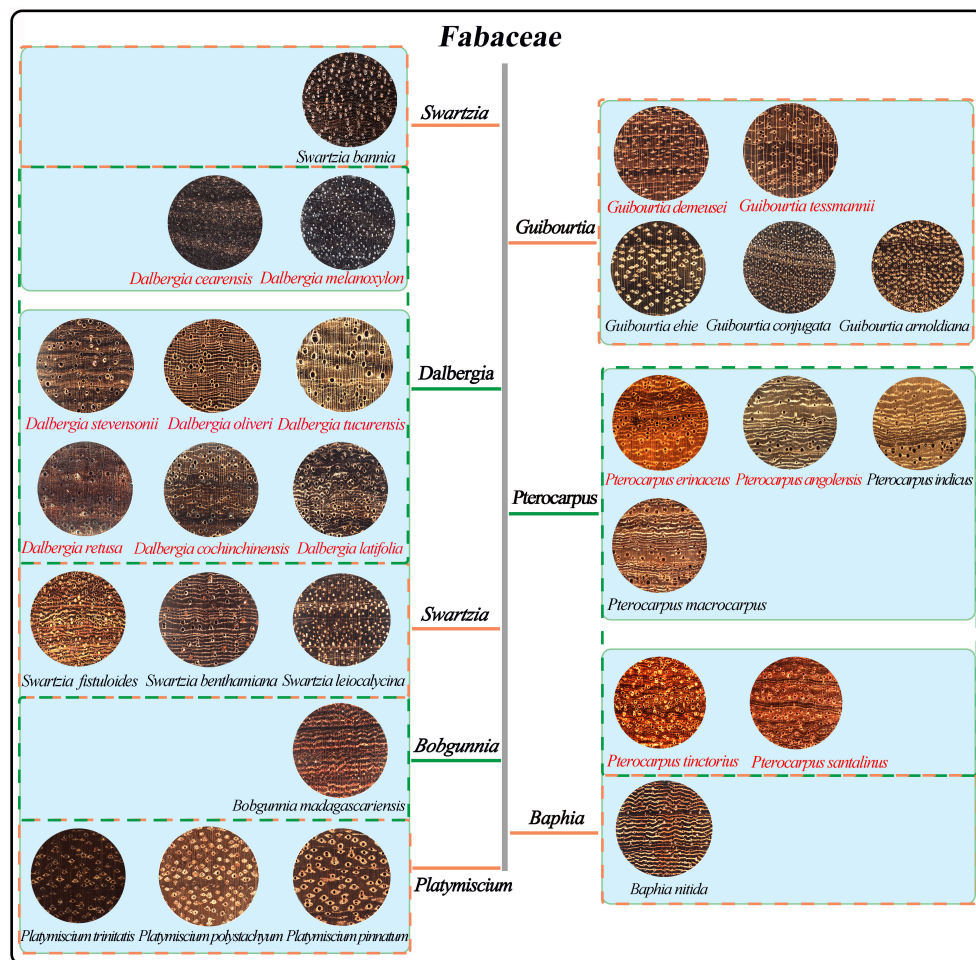
FIGURE 1

The transverse section of wood species in this research. The CITES-listed wood species and their look-alikes in the same blue frame are often confused (species in red text are CITES-listed). Species in the same dashed box are in the same genus.

TABLE 1　Detailed information of the CITES-28 image dataset.

Label	Species (class)	Protection level	Number of collected images	Number of xylarium specimens	Number of images/ xylarium specimens
0	*Dalbergia cearensis*	CITES II	321	14	23
1	*Dalbergia cochinchinensis*	CITES II	263	7	38
2	*Dalbergia latifolia*	CITES II	762	21	36
3	*Dalbergia melanoxylon*	CITES II	313	11	28
4	*Dalbergia oliveri*	CITES II	414	6	69
5	*Dalbergia retusa*	CITES II	698	16	44
6	*Dalbergia stevensonii*	CITES II	498	11	45
7	*Dalbergia tucurensis*	CITES II	461	12	38
8	*Platymiscium pinnatum*	/	201	10	20

(Continued)

TABLE 1 Continued

Label	Species (class)	Protection level	Number of collected images	Number of xylarium specimens	Number of images/ xylarium specimens
9	*Platymiscium polystachyum*	/	193	16	12
10	*Platymiscium trinitatis*	/	134	6	22
11	*Swartzia bannia*	/	122	5	24
12	*Swartzia benthamiana*	/	106	6	18
13	*Swartzia fistuloides*	/	111	4	28
14	*Swartzia leiocalycina*	/	115	7	16
15	*Bobgunnia madgascariensis*	/	110	5	22
16	*Pterocarpus erinaceus*	CITES II	373	9	41
17	*Pterocarpus santalinus*	CITES II	196	4	49
18	*Pterocarpus tinctorius*	CITES II	263	6	44
19	*Pterocarpus angolensis*	CITES II	805	19	42
20	*Pterocarpus indicus*	/	1,163	29	40
21	*Pterocarpus macrocarpus*	/	600	18	33
22	*Baphia nitida*	/	123	5	25
23	*Guibourtia demeusei*	CITES II	154	4	39
24	*Guibourtia tessmannii*	CITES II	249	7	36
25	*Guibourtia arnoldiana*	/	294	9	33
26	*Guibourtia ehie*	/	337	9	37
27	*Guibourtia conjugata*	/	58	3	19
	Total		**9,437**	**279**	**~34**

The meaning of the symbol "/" represents the species are not protected by CITES.
The bold values represents the total for each column.

TABLE 2 Prediction accuracies of two established models.

Model	Rotation mode	Image compensation	Accuracy (%)
SE-ResNet	Fixed	Y	**99.65**
	Fixed	N	99.55
	Random	Y	99.63
	Random	N	99.47
ResNet	Fixed	Y	**99.45**
	Fixed	N	99.41
	Random	Y	99.28
	Random	N	99.25

Y represents use 0 as a compensation item to complete the processed image.
N denotes that the rotated image is not processed.
The bold values represent the highest accuracy of SE-ResNet and ResNet.

$$aug(xi) \ = \ R(xi) \ + \ C \qquad (1)$$

where *aug* is the enhancement method to be passed by the sample input to the model training. *R* is the image operation method for wood pictures proposed in this paper. During operation, image samples smaller than the input requirements of the model may be generated; thus, this paper introduces compensation factor *C* to compensate for the phenomenon of missing samples caused by the data enhancement operation.

In the actual test application process, controlling the acquisition angle of the test equipment and other factors is difficult. Therefore, we design a rule-based data enhancement rule with rotation as the core when building the data enhancement method. In addition, considering that the test angle accepted by the test equipment cannot be guaranteed in the actual test process, we propose a random rotation strategy to expand the diversity of training data further. The data enhancement operation can be represented by
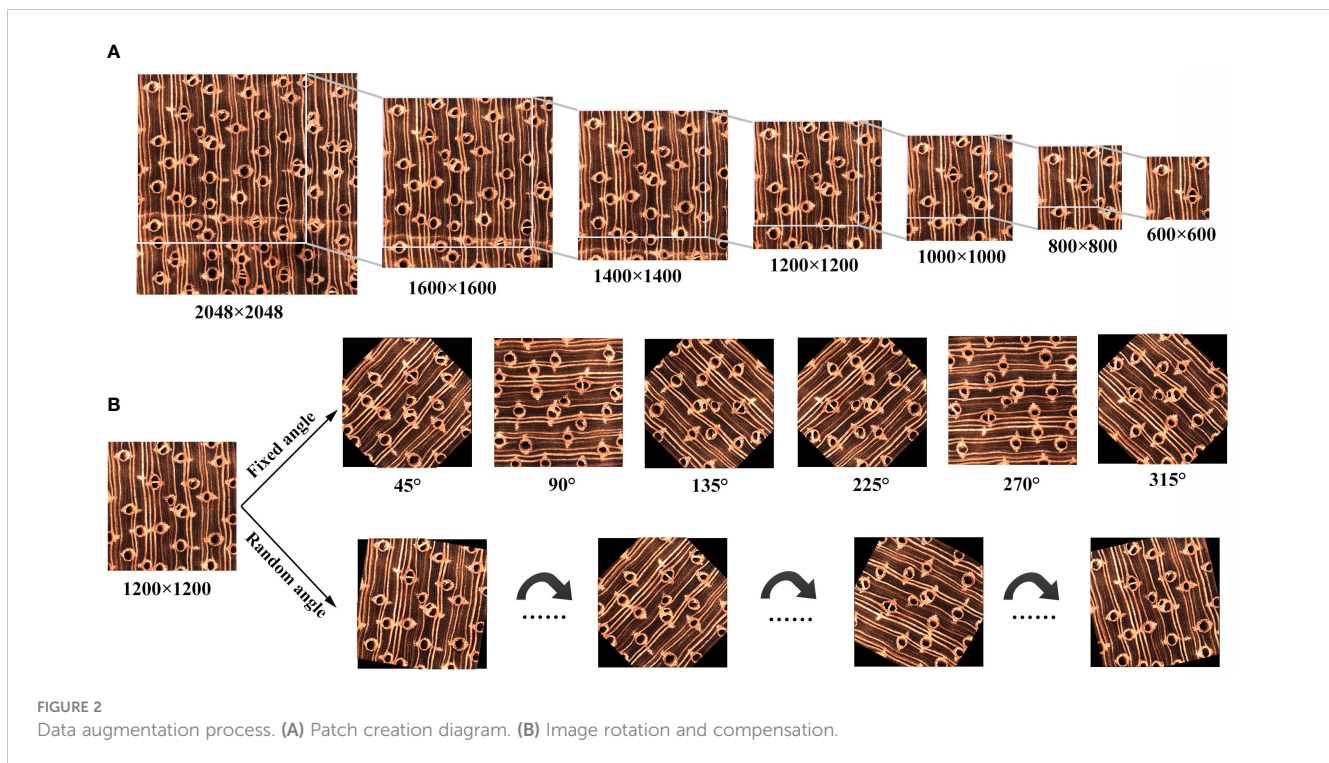
FIGURE 2
Data augmentation process. **(A)** Patch creation diagram. **(B)** Image rotation and compensation.

Equation 2.

$$R = R_{random} \cap R_{rule} \qquad (2)$$

where R_{random} represents the random rotation strategy and R_{rule} represents the rule rotation strategy.

The data enhancement process is shown in Figure 2. When we enhanced the data based on fixed rotation, the image was rotated to the center with rotation angles of {45°, 90°, 135°, 225°, 270°, 315°}. When we enhanced the data based on a random rotation strategy, the image was rotated to the center, and the rotation angle was randomly sampled between 0° and 360°. After obtaining the rotated image, to meet the requirements that the model can input only the square image region, we cut the rotated image. We used 0 (black) as a compensation item to make up the exact area.

2.2 Architecture and training of convolutional neural networks

2.2.1 Convolutional neural network architecture

Based on the enhanced data, we selected ResNet, a widely used network, as the backbone. As He et al. (2016) discovered, a multilayer deep neural network can produce unexpected results, and the training accuracy drops as the layers increase, which is technically known as vanishing gradients. To address this problem, ResNet was proposed to help build a deeper neural network by utilizing skip connections or shortcuts to jump over some layers. SENet, as a classical attention mechanism, can be embedded in the CNN. In this study, we selected SE-ResNet to carry out our experiments.
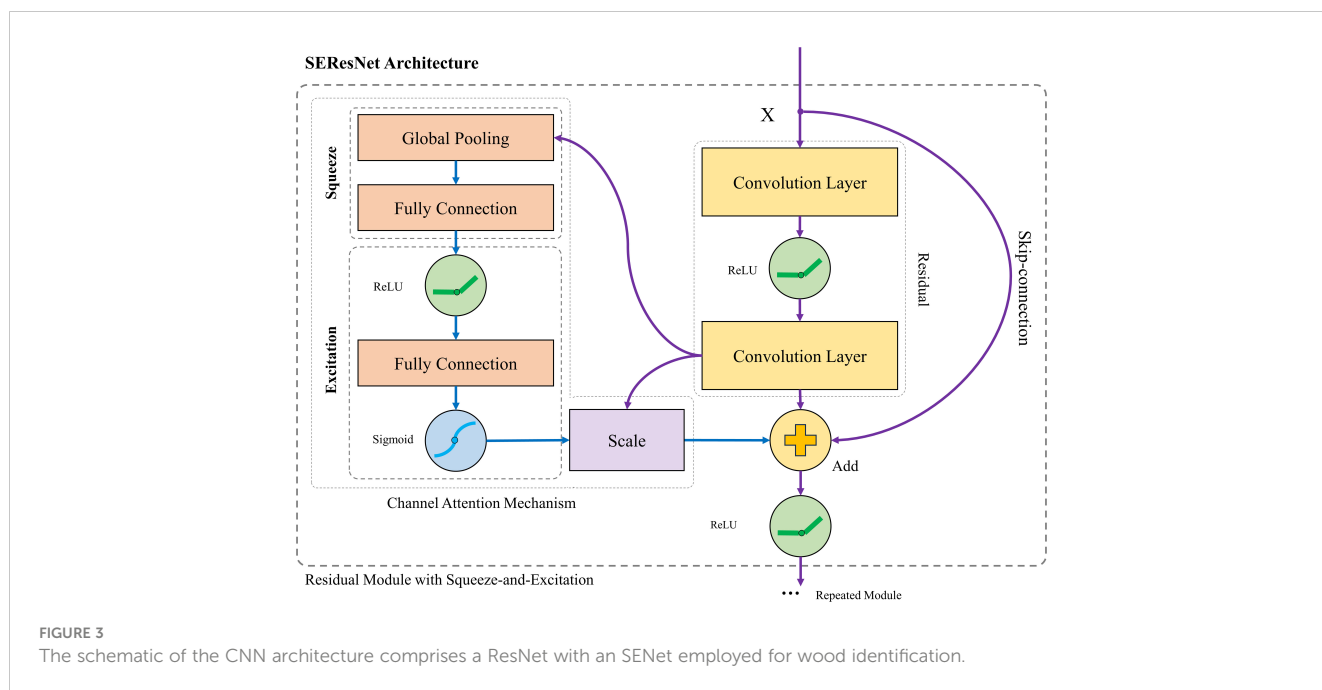
To better obtain the feature expression of wood images in neural networks, more detailed analysis and processing were carried out on the features among image channels and the depth extraction of two-dimensional features. SENet provides the unit with a mechanism to explicitly model dynamic, nonlinear dependencies between channels using global information (Hu et al., 2018). SENet can ease the learning process and significantly enhance the representational power of the network. The SE-ResNet module includes a block, global average pooling, a fully connected (FC) layer, and an activation function layer with ReLU and sigmoid. The channel attention module is added after the residual module. The schematic architecture of the SE-ResNet model is shown in Figure 3.

Equation 3 shows that SENet mines and analyzes the weight information of different channels by studying the relationship between channels. First, it squeezes global spatial information into a channel descriptor by using global average pooling to generate channelwise statistics, which is ρ in Equation 3. Then, it learns to recalibrate the feature adaptively through two FC layers and uses two activation functions to learn nonmutually exclusive relationships. After obtaining the scalars, channelwise multiplication between the scalar and the feature map is carried out to obtain the final output.

$$s = \sigma(W_2 \delta(W_1 \rho(x))) \qquad (3)$$

2.2.2 Optimization objective

We complete the forward pass and update the network by backpropagation. Globally, the loss function over iterations is still minimized until the loss converges, as shown in Equation 4, where θ

FIGURE 3
The schematic of the CNN architecture comprises a ResNet with an SENet employed for wood identification.

is the model parameter, X is the model input, and Y is the corresponding label.

Given an augmented selection, the model is optimized using perturbed features. The feature extraction ability is enhanced during the training process to obtain outstanding performance. In this sense, the augmentations help the model acquire more knowledge of potential input samples, increasing classification performance for real scenes.

$$\widehat{\theta} = argminL_{\theta}(\theta|aug(X), Y) \qquad (4)$$

2.3 Model training

Model training is carried out in two phases. In the first phase, the model pretrained on the ImageNet dataset is studied for image classification via transfer learning. In the second phase, the model is trained to build classifiers for wood identification with the training set and tested at every epoch with the validation set. The initial learning rate of the model is 0.001, the momentum value is set to 0.9, and 24 epochs are trained sequentially. The epoch with the best result in the test set is taken to save the model. Stochastic gradient descent (SGD) is used to optimize the model. The model trained with the following hardware specifications: CPU Intel Core i9-14900K 6.0 GHz, 24 GB of RAM NVIDIA RTX 4090, and 96 GB of GPU. The experimental software environment is Ubuntu 20.04, Python 3.7.13, PyTorch 1.12.1, and NVIDIA CUDA 10.2.

In this study, the following process investigated the optimal image processing method (image rotation mode and patch size) for model accuracy. First, the model was trained with six different image patches extracted from the original image dataset (600 pixels × 600 pixels, 800 pixels × 800 pixels, 1,000 pixels × 1,000 pixels, 1,200 pixels × 1,200 pixels, 1,400 pixels × 1,400 pixels, and 1,600 × 1,600 pixels) to

determine the most appropriate size. For the most appropriate patch size dataset, we rotated the patches with two strategies, fixed rotation and random rotation, to achieve the expansion of the image dataset. Compensated or uncompensated processing was performed for images obtained with different rotations. Then, image datasets obtained from different processing strategies were used to train the model and compare the impact of different data enhancement methods on model wood identification performance.

2.4 Evaluation

The performance of the trained models was evaluated using a test set. The highest identification accuracy of all models based on test set images is reported. A confusion matrix is given to better understand the classification results of wood species given by the model and to analyze the causes of species discrimination errors in terms of wood anatomy. The confusion matrix contains information about the true and predicted values of the classification and reflects the wood species classification results. Accuracy usually describes model performance on all sample categories and is used when all sample categories are equally important. The higher the value is, the better the performance of the classification model. Accuracy is calculated using true positives (TPs), false positives (FPs), true negatives (TNs), and false negatives (FNs), which are shown in Equation 5.

$$Accuracy \ (A) = \frac{TN + TP}{TN + TP + FP + FN} \qquad (5)$$

To further test the generalization ability of the trained model, for each species, we captured 10 images from each sample purchased from the Guangdong Yuzhu Timber Market to conduct real-world identification.
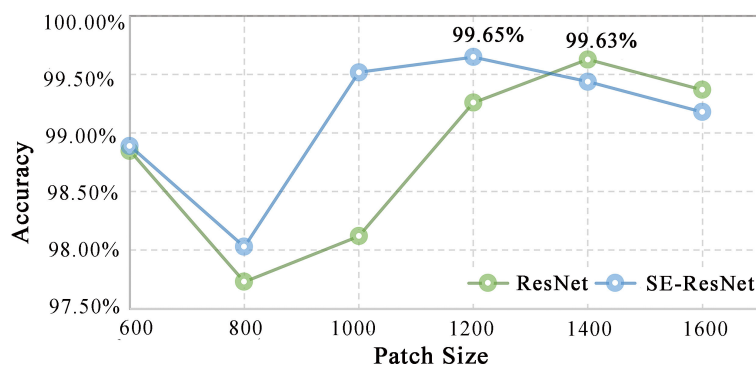
FIGURE 4
The model performance when using images of different patch sizes.

3 Results and discussion

3.1 Appropriate image data augmentation for wood species identification models

3.1.1 Image cropping

Cropping is the common means in the present study. Figure 4 displays the model performance of different patch sizes in this experiment. This figure shows that the ResNet and SE-ResNet models can achieve high identification accuracy when the patch size is larger than 1,200 pixels × 1,200 pixels, which aligns with previous research (He et al., 2020). The appropriate perceptual field of view size is an important factor affecting the classification performance of the model. The computational cost of high-resolution images is too high; thus, the image size must be adjusted but kept within the range where the desired features can be extracted. In addition, information loss should be accounted for when choosing the size of the wood image. For macroscopic images of wood cross-sections, the size of the image feature field of view needs to be balanced with the number of images. Therefore, to obtain as much image data as possible with low overlap while ensuring model accuracy, we consider a patch size of 1,200 pixels × 1,200 pixels to be the optimal size for model training.

3.1.2 Image rotation

The ResNet and SE-ResNet models were established based on the most suitable patch size. The validity of different image rotation methods on wood identification models was compared for the first time. Table 3 shows that the fixed rotation method reached higher accuracy for the ResNet (99.65%) and SE-ResNet (99.45%) models. Moreover, these models exhibited relatively high accuracy after using 0 as a compensation item to complete the cropped image; both values were above 99.28%. The results showed that using a fixed rotation method and image compensation to obtain adequate training data is feasible. The accuracy and loss curves of the SEResNet model during the training and test process are shown in Figure 5. It showed that learning rate decays by 10 times and tended to be flat when training to 15, respectively. At the same epoch, the test set reached the highest accuracy with a loss value of less than 0.05 (Figure 5B). It showed that this model has strong generalization and stability ability and is able to conduct wood species identification.

Some existing publicly accessible datasets provide numerous images that can be used in computer vision research, but no more attention is given to preprocessing images from different sources before use. Geus et al. (2021) conducted an experiment that applied a rotation of 1° to each image within the training set, covering a range from −15° to 15°. Nevertheless, this experiment did not give specific conclusions about wood image preprocessing methods, and no further research has been carried out. In this experiment, the image undergoes segmentation, followed by rotation over a comprehensive 360° range and subsequent compensation. The results serve to validate the efficacy of the proposed preprocessing approach in substantially expanding the available dataset.

TABLE 3 Image dataset details with the different patch sizes.

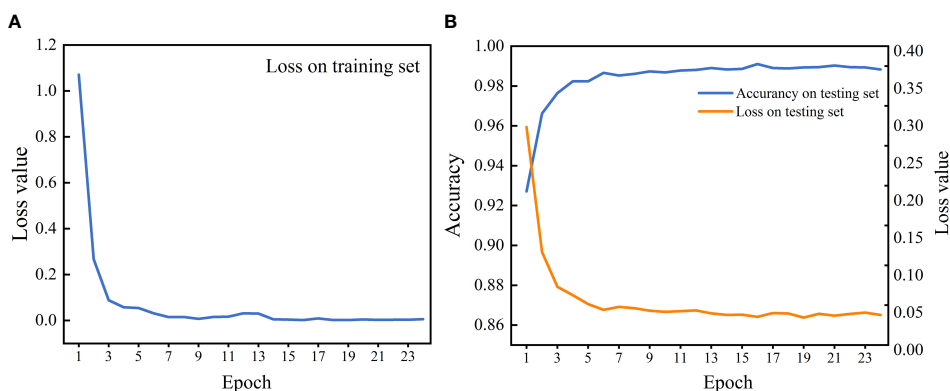| Patch size	Number of train patch	Number of val patch	Number of test patch	Number of total patch
1,600	32,776	3,790	3,889	40,455
1,400	32,157	4,296	4,002	40,455
1,200	32,408	4,029	4,018	40,455
1,000	32,204	3,988	4,263	40,455
800	32,286	3,857	4,312	40,455
600	32,215	4,170	4,070	40,455

FIGURE 5
The loss value and accuracy of the SEResNet model across each epoch. **(A)** The model training. **(B)** The model test.

3.2 Model accuracy improvement by adding an attention channel

CNN models commonly used in previous studies, such as AlexNet, InceptionV3, and DenseNet, have shown strong feature extraction capabilities when processing two-dimensional images (Geus et al., 2021; Wagle and Harikrishnan, 2021); however, the model performance is unsatisfactory when targeting wood image data. The wood section image contains features with strong regularity. In contrast, the variations between different wood species are subtle. In addition, the difference in wood section images is mainly reflected in factors such as texture, color, and control distribution. The attention mechanism can extract key features of similar images better. In this experiment, the channel attention mechanism is introduced to improve wood species identification accuracy at the species level.

As shown in Table 3, the model accuracy of SE-ResNet is generally higher than that of ResNet under various rotation and compensation methods. Compared with ResNet, the performance of SE-ResNet was better and suitable for identifying the CITES-listed wood species with their look-alikes. ResNet incorporates SENet, which enhances the capacity of the network to learn identification keys by adaptively recalibrating the channelwise feature responses. The attention mechanism in computer vision is inspired by the human attention mechanism, which imitates the manner in which people focus more on specific information in an image while ignoring the rest (Lu et al., 2023). As a module that can affect the performance of the model, the attention mechanism can focus limited attention on only the most essential information to save computational resources and obtain the most effective information quickly. In particular, when the original model is underfitting with fewer parameters and cannot fully learn the training data rules, adding the attention module enhances the expressive power of the model, improving the underfitting problem and accuracy.

3.3 Discrimination of CITES-listed species from their look-alikes

In this study, the SE-ResNet model obtained the highest accuracy for wood species identification on a dataset with fixed and complementary rotation. The confusion matrix in Figure 6 shows the classification results of the model for each wood species in the test set. Overall, the classification accuracy of the 28 tree species reached 99.65%, with 20 wood species correctly classified. Surprisingly, the predictions made by the SE-ResNet model for five *Guibourtia* species were perfect. When discriminating CITES-listed *Dalbergia* species, they were identified with an accuracy of over 95% with this model. The CITES-listed *Dalbergia* species was completely distinguished from eight look-alike species, including *Platymiscium*, *Swartzia*, and *Bobgunnia*. However, approximately 3.23% of the *Dalbergia cearensis* images were misclassified as *Pterocarpus angolensis*.

Within *Pterocarpus*, all the CITES-listed *Pterocarpus* species were completely discriminated from their look-alikes, except for *Pterocarpus angolensis*, which was confused with *Pterocarpus indicus* and *Pterocarpus macrocarpus*. For the closely related *Pterocarpus* species, *Pterocarpus angolensis*, *Pterocarpus erinaceus*, *Pterocarpus macrocarpus*, and *Pterocarpus indicus* were misjudgments between each other. These four species have large wood anatomical similarities, such as the axial parenchyma arrangement and ray type, which make them difficult to distinguish completely and accurately on the basis of macroscopic images of a single cross-section. Not surprisingly, *Pterocarpus santalinus* and *Pterocarpus tinctorius* appeared to be misidentified, with 5.88% of the images of *Pterocarpus tinctorius* misclassified as *Pterocarpus santalinus*. With the traditional methods, it is extremely difficult to distinguish between these two species because of their highly similar macroscopic and microscopic structural characteristics. Therefore, *Pterocarpus tinctorius* was listed in CITES Appendix II in 2019 due to its over-exploitation and extreme similarity with *Pterocarpus santalinus*, listed in CITES Appendix II in 2007.

The vast majority of wood and wood products have lost the key identification characteristics, such as leaves and bark, which increases the difficulty of wood identification. In particular, wood species of the same genus or even closely related genera have high similarities in appearance and anatomical characteristics. Not surprisingly, compared to human vision-based methods, the results demonstrated in this experiment can provide higher identification accuracy (Wiedenhoeft et al., 2019). Simultaneously, the deep learning model effectively eradicates human subjectivity
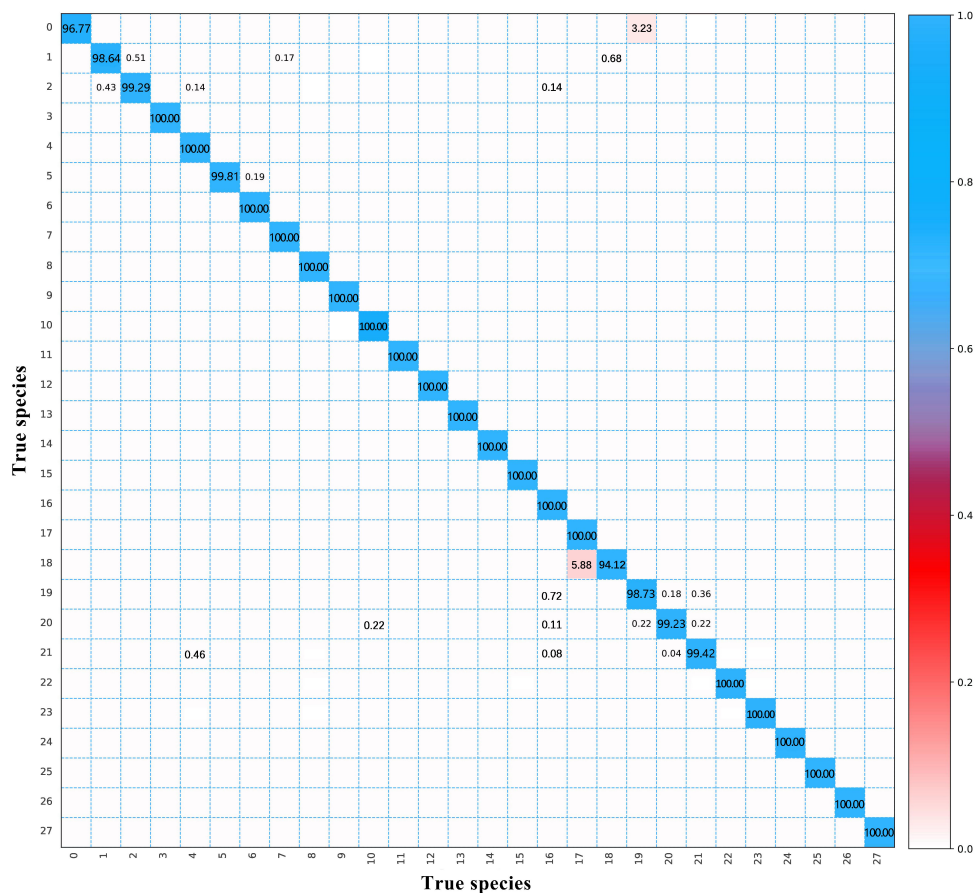
FIGURE 6
The classification results of the SE-ResNet model with the test dataset. 0—Dalbergia cearensis; 1—Dalbergia cochinchinensis; 2—Dalbergia latifolia; 3—Dalbergia melanoxylon; 4—Dalbergia oliveri; 5—Dalbergia retusa; 6—Dalbergia stevensonii; 7—Dalbergia tucurensis; 8—Platymiscium pinnatum; 9—Platymiscium polystachyum; 10—Platymiscium trinitatis; 11—Swartzia bannia; 12—Swartzia benthamiana; 13—Swartzia fistuloides; 14—Swartzia leiocalycina; 15—Bobgunnia madgascariensis; 16—Pterocarpus erinaceus; 17—Pterocarpus santalinus; 18—Pterocarpus tinctorius; 19—Pterocarpus angolensis; 20—Pterocarpus indicus; 21—Pterocarpus macrocarpus; 22—Baphia nitida; 23—Guibourtia demeusei; 24—Guibourtia tessmannii; 25—Guibourtia arnoldiana; 26—Guibourtia ehie; 27—Guibourtia conjugate.

TABLE 4 The identification performance of the models for independent images.

Label	Species (class)	SE-Resnet		ResNet	
		Accuracy%	Misidentified labels	Accuracy%	Misidentified labels
0	*Dalbergia cearensis*	100	/	100	/
1	*Dalbergia cochinchinensis*	80	*Dalbergia retusa* *Dalbergia stevensonii*	0	*Pterocarpus santalinus* *Dalbergia latifolia* *Dalbergia stevensonii*
2	*Dalbergia latifolia*	100	/	100	/
3	*Dalbergia melanoxylon*	100	/	100	/
4	*Dalbergia oliveri*	80	*Dalbergia stevensonii*	100	/
5	*Dalbergia retusa*	100	/	80	*Dalbergia oliveri*
6	*Dalbergia stevensonii*	100	/	0	*Dalbergia tucurensis* *Dalbergia retusa*
7	*Dalbergia tucurensis*	80	*Dalbergia stevensonii*	70	*Pterocarpus indicus* *Dalbergia oliveri*

(Continued)

TABLE 4   Continued

Label	Species (class)	SE-Resnet		ResNet	
		Accuracy%	Misidentified labels	Accuracy%	Misidentified labels
8	*Platymiscium pinnatum*	40	*Platymiscium polystachyum* *Platymiscium trinitatis*	70	*Platymiscium trinitatis*
9	*Platymiscium polystachyum*	20	*Platymiscium pinnatum*	40	*Platymiscium pinnatum*
10	*Platymiscium trinitatis*	20	*Dalbergia melanoxylon* *Dalbergia retusa*	50	*Dalbergia stevensonii* *Platymiscium polysta*
11	*Swartzia bannia*	100	/	60	*Dalbergia stevensonii* *Dalbergia oliveri*
12	*Swartzia benthamiana*	60	*Dalbergia cochinchinensis* *Dalbergia stevensonii*	70	*Pterocarpus angolensis*
13	*Swartzia fistuloides*	100	/	100	/
14	*Swartzia leiocalycina*	40	*Platymiscium trinitatis* *Swartzia fistuloides*	10	*Swartzia bannia* *Platymiscium trinitatis*
15	*Bobgunnia madgascariensis*	100	/	90	*Dalbergia cearensis*
16	*Pterocarpus erinaceus*	100	/	100	/
17	*Pterocarpus santalinus*	90	*Pterocarpus macrocarpus*	90	*Dalbergia retusa*
18	*Pterocarpus tinctorius*	90	*Pterocarpus santalinus*	100	/
19	*Pterocarpus angolensis*	100	/	100	/
20	*Pterocarpus indicus*	100	/	90	*Dalbergia oliveri*
21	*Pterocarpus macrocarpus*	80	*Pterocarpus angolensis* *Pterocarpus indicus*	50	*Dalbergia oliveri*
22	*Baphia nitida*	100	/	100	/
23	*Guibourtia demeusei*	100	/	90	*Pterocarpus tinctorius*
24	*Guibourtia tessmannii*	100	/	10	*Guibourtia demeusei*
25	*Guibourtia arnoldiana*	40	*Guibourtia ehie*	0	*Guibourtia ehie*
26	*Guibourtia ehie*	100	/	100	/
27	*Guibourtia conjugata*	100	/	100	/
	Total	**82.3**		**70.4**	

The meaning of the symbol "/" represents the species don't misidentified as other wood species.
The bold values represents the average accuracy of SE-ResNet and ResNet.

and assumes the role of a professional with greater efficiency and precision.

3.4 Prospect of wood species identification based on a deep learning model

Table 4 presents the discriminative results of the SE-ResNet model for images collected from the timber markets. For completely unfamiliar samples, the recognition accuracies of SE-Resnet and Resnet are 82.3% and 70.4%, respectively; the generalization ability of SE-Resnet model is superior. For SE-Resnet, the classification results show that images of 16 species were fully identified and 6 species had 8 or 9 out of 10 images accurately identified. In this regard, the CITES-listed tree species of *Pterocarpus* and *Guibourtia* were all

distinguished from their look-alikes. *Dalbergia* spp., *D. cearensis*, and *D. cochinchinensis* were also accurately identified. However, this model performed worse when discriminating *Platymiscium* spp., with images of all three species at a level of poor stability. Although the number of samples is large, the effective area of the image that can be acquired for each sample is small and there is a relatively small number of images of *Platymiscium* used for modeling. Such a phenomenon also appears in *Swartzia benthamiana* with *Swartzia leiocalycina*. It demonstrated that even if more images could be obtained through data enhancement strategies, fully covering the variability within species would not have been possible.

Based on this result, we considered that the model proposed in this experiment could identify 22 tree species accurately and consistently. The SE-ResNet model has great potential for generalization, and the model gives better recognition for wood

species with more specimens. Thus, it is recommended that more specimens of the same species be collected to train the model to cover as much variability as possible and that, where possible, many specimens be used to test the generalization ability of the model in future studies, which can enhance the accuracy, generalization ability, and applicability of the trained model.

Our experiment explored the effect of image cropping and rotation on the performance of deep learning models for wood image identification and added an attention channel module into the ResNet-50 model. It was found that fixed rotation and cropping are effective data enhancement methods for wood images. In addition, we added only the attention channel module into the models, and other attention mechanisms, such as multi-head self-attention, self-attention, and convolutional block attention modules, were not tested in wood species identification (Woo et al., 2018; Guo et al., 2022). Follow-up research should explore more data enhancement methods, such as mirroring and scaling. The generative adversarial network (GAN) technique is a new computer vision method. The GAN-based oversampling technique not only increases the minority class representation to solve class imbalance problems but also may help to prevent overfitting (Sampath et al., 2021). However, GANs have not been reported in wood image recognition.

Numerous practical tools have been developed to address the specific field application requirements, exemplified by solutions such as MyWood-ID (Tang et al., 2018) and XyloPhone (Wiedenhoeft, 2020). Considering the comprehensive research conducted in this paper, an intelligent wood identification system was devised and successfully deployed for on-site inspections carried out by customs officials. Compared to other emerging techniques (NIRS, DART-MS, and DNA barcode), the computer vision-based wood identification approach demonstrates considerable promise in achieving species-level precision in field screening for wood species.

At this research stage, the dataset images are mainly processed on the surface of the wood by both knifing and sanding in the laboratory so that the anatomical characteristics of the wood can be fully exposed before the images are acquired. The experimentally acquired images after fine sanding are idealized, and the models trained based on these image data may not match the practical application. Therefore, exploring the deep learning models with wood images of rough surfaces is necessary to reduce the workload of wood surface treatment by processing rough surface images.

4 Conclusion

The traditional wood species identification method has a history of nearly a century, forming a complete identification process and norms, and still occupies a dominant position. However, in the new era of demand, traditional methods need to be complemented by new technologies to achieve species-level wood species identification. Faced with the current conservation pressure of endangered tree species, a more accurate, convenient, and economically friendly method is urgently needed for the identification of wood species. Computer vision is the most feasible technique for traded wood

species identification coupled with deep learning, especially for the field inspection at import and export ports.

The results of this study showed that the 1,200 pixel × 1,200 pixel patch size can be applied as the best choice for the training model and that the number of wood images can be effectively expanded by image cropping and rotation. In addition, the channel attention mechanism (SENet) module is added to the CNN structure to identify CITES-listed tree species and their look-alikes with an accuracy of over 99%. It shows a relatively satisfactory performance in real-world identification. For the public, the model does not require extensive knowledge of wood anatomy and experience in species identification, which greatly reduces the complexity of the traditional wood identification process.

This work not only provides a CNN model along with added attention channels for successful identification but also provides a standard guideline for image data enhancement when conducting wood species identification. Further studies are needed to explore the interspecific wood anatomy features with deep learning models and feature visualization. The results of this study show that the wood transverse image dataset coupled with the SE-ResNet model can accurately discriminate CITES-listed species from their look-alikes to combat illegal timber trade and contribute to tree species conservation.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

SL: Data curation, Writing – original draft, Writing – review & editing, Methodology, Visualization. CZ: Methodology, Software, Writing – review & editing. JW: Data curation, Writing – review & editing. YL: Data curation, Writing – review & editing. JY: Methodology, Writing – review & editing. ZZ: Software, Writing – review & editing. YY: Funding acquisition, Writing – review & editing. TH: Funding acquisition, Writing – review & editing.

Funding

Acknowledgments

Mr. Yupei Wei for their help with the figures in this paper, and Dr. Yu Sun for his contributions to image collection.

Conflict of interest

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Atikah, T., Wardani, W., and Surya, D. (2021). "Population, economic value and the trade of Dalbergia latifolia in Lombok, West Nusa Tenggara, Indonesia," in *In IOP Conference Series: Earth and Environmental Science*, Vol. 762. 012015. doi: 10.1088/1755-1315/762/1/012015

Bergo, M. C., Pastore, T. C., Coradin, V. T., Wiedenhoeft, A. C., and Braga, J. W. (2016). NIRS identification of *Swietenia macrophylla* is robust across specimens from 27 countries. *IAWA J.* 37, 420–430. doi: 10.1163/22941932-20160144

Brancalion, P. H. S., Almeida, D. R. A. D., Vidal, E., Molin, P. G., Sontag, V. E., Souza, S. E. X. F., et al. (2018). Fake legal logging in the Brazilian Amazon. *Sci. Adv.* 4, eaat1192. doi: 10.1126/sciadv.aat1192

Cazzolla, G. R., Reich, P. B., Gamarra, J. G., Crowther, T., Hui, C., Morera, A., et al. (2022). The number of tree species on Earth. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2115329119. doi: 10.1073/pnas.2115329119

CITES (2023). *Appendices I, II and III*. Available online at: https://cites.org/eng/app/appendices.php (Accessed June 25, 2023).

Committee, I. H. (1989). IAWA list of microscopic features for hardwood identification. *IAWA J.* 10, 219–332. doi: 10.1163/22941932-90000349

Deklerck, V., Mortier, T., Goeders, N., Cody, R. B., Waegeman, W., Espinoza, E., et al. (2019). A protocol for automated timber species identification using metabolome profiling. *Wood Sci. Technol.* 53, 953–965. doi: 10.1007/s00226-019-01111-1

Dyk, D. A. V., and Meng, X. L. (2001). The art of data augmentation. *J. Comput. Graph Stat.* 10, 1–50. doi: 10.1198/10618600152418584

Falk, B. (2009). Wood as a sustainable building material. *For Prod J.* 59, 6–13. doi: 10.5176/2301-394x_ace18.43

Geus, A. R. D., 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 Sci. Technol.* 55, 857–872. doi: 10.1007/s00226-021-01282-w

Goldsmith, E. I. (1978). The convention on international trade in endangered species of wild fauna and flora. *J. Med. Primatol.* 7, 122–124. doi: 10.1159/000459796

Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., et al. (2022). Attention mechanisms in computer vision: A survey. *Comput. Visual media* 8, 331–368. doi: 10.1007/s41095-022-0271-y

He, T., Liu, S. J., Lu, Y., Zhang, Y. G., Jiao, L. C., and Yin, Y. F. (2021). iWood: An automated wood identification system for endangered and precious tree species using convolutional neural networks. *Scientia Silvae Sinicae.* 57, 152–159. doi: 10.11707/j.1001-7488.20210915

He, T., Lu, Y., Jiao, L. C., Zhang, Y. G., Jiang, X. M., and Yin, Y. F. (2020). Developing deep learning models to automate rosewood tree species identification for CITES designation and implementation. *Holzforschung.* 74, 1123–1133. doi: 10.1515/hf-2020-0006

He, K., Zhang, X., Ren., S., and Song, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA. 770–778. doi: 10.1109/CVPR.2016.90

Hu, J., Shen, L., and Sun, G. (2018). Squeeze-and-excitation networks. *Proc. Proc. IEEE Conf. Comput. Vision Pattern Recognit.* 2018, 7132–7141. doi: 10.1109/CVPR.2018.00745

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, 1–21. doi: 10.1186/s13007-021-00746-1

INTERPOL (2021). *Forestry Crime Fact sheet 2021*. Available online at: https://www.interpol.int/content/download/17367/file/Forestry%20Crime%20Fact%20sheet%202021.pdf (Accessed May 15, 2023).

IUCN (2022). *The IUCN Red List of Threatened Species*. Available online at: https://www.iucnredlist.org (Accessed December 15, 2022).

Jiao, L. C., Lu, Y., He, T., Guo, J., and Yin, Y. F. (2020). DNA barcoding for wood identification: Global review of the last decade and future perspective. *IAWA J.* 41, 620–643. doi: 10.1163/22941932-bja10041

Kırbaş, İ., and Çifci, A. (2022). An effective and fast solution for classification of wood species: A deep transfer learning approach. *Ecol. Inform.* 69, 101633. doi: 10.1016/j.ecoinf.2022.101633

Kwon, O., Lee, H. G., Lee, M. R., Jang, S., Yang, S. Y., Park, S. Y., et al. (2017). Automatic wood species identification of Korean softwood based on convolutional neural networks. *Korean J. Wood Sci. Technol.* 45, 797–808. doi: 10.5658/WOOD.2017.45.6.797

Lambdon, P. W., and Ellick, S. (2016). *Lachanodes arborea* (The IUCN Red List of Threatened Species) (Accessed 06 June 2024).

Latib, H. A., Liat, L. C., Ratnasingam, J., Law, E. L., Azim, A. A. A., Mariapan, M., et al. (2020). Suitability of paulownia wood from Malaysia for furniture application. *BioResources.* 15, 4727–4737. doi: 10.15376/biores.15.3.4727-4737

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., and Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing* 234, 11–26. doi: 10.1016/j.neucom.2016.12.038

Lu, S., Liu, M., Yin, L., Yin, Z., Liu, X., and Zheng, W. (2023). The multi-modal fusion in visual question answering: a review of attention mechanisms. *PeerJ Comput. Sci.* 9, e1400. doi: 10.7717/peerj-cs.1400

Oktaria, A. S., Prakasa, E., and Suhartono, E. (2019). Wood species identification using convolutional neural network (cnn) architectures on macroscopic images. *J. Inf Comput. Sci.* 4, 274–283. doi: 10.25126/jitecs.201943

Pan, X., Li, K., Chen, Z., and Yang, Z. (2021). Identifying wood based on near-infrared spectra and four gray-level co-occurrence matrix texture features. *Forests.* 12, 1527. doi: 10.3390/f12111527

Price, E. R., Miles-Bunch, I. A., Gasson, P. E., and Lancaster, C. A. (2021). *Pterocarpus* wood identification by independent and complementary analysis of DART-TOFMS, microscopic anatomy, and fluorescence spectrometry. *IAWA J.* 42, 397–418. doi: 10.1163/22941932-bja10064

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, 1–10. doi: 10.1186/s13007-018-0292-9

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. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.01015

Ravindran, P., and Wiedenhoeft, A. C. (2020). Comparison of two forensic wood identification technologies for ten meliaceae woods: computer vision versus mass spectrometry. *Wood Sci. Technol.* 54, 1139–1150. doi: 10.1007/s00226-020-01178-1

Rivers, M. (2017). The Global Tree Assessment–red listing the world's trees. *BGjournal.* 14, 16–19. Available at: https://www.jstor.org/stable/26369172.

Sampath, V., Maurtua, I., Martín, J. J. A., and Gutierrez, A. (2021). A survey on generative adversarial networks for imbalance problems in computer vision tasks. *J. Big Data.* 8, 27. doi: 10.1186/s40537-021-00414-0

Shmulsky, R., and Jones, P. D. (2019). ""Tree growth and production of woody tissue."," in *The Forest Products and Wood Science* (Wiley, Chichester), 1–20. doi: 10.1002/9781119426400

Su, H. S., Lv, J., Ding, Z. P., Tang, Y. J., Chen, X. D., Zhou, Q., et al. (2021). Wood identification algorithm based on improved residual neural network. *Chin. J. Sci. Silvae Sin.* 57, 147–154. doi: 10.11707/j.1001-7488.20211215

Tang, X. J., Tay, Y. H., Siam, N. A., and Lim, S. C. (2018). "MyWood-ID: Rapid and robust automated macroscopic wood identification system using smartphone with macro-lens," in *In Proceedings of the 2018 International Conference on Computational Intelligence and Intelligent Systems—CIIS 2018.* 37–43.

Voulodimos, A., Doulamis, N., Doulamis, A., and Protopapadakis, E. (2019). Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* 2018, 1–13. doi: 10.1155/2018/7068349

Wagle, S. A., and Harikrishnan, R. (2021). Comparison of plant leaf classification using modified alexNet and support vector machine. *Trait Signal.* 1, 38. doi: 10.18280/ts.380108

Wang, W., Yang, Y., Wang, X., Wang, W., and Li, J. (2019). Development of convolutional neural network and its application in image classification: a survey. *Opt Eng.* 8, 040901–040901. doi: 10.1117/1.OE.58.4.040901

Wick, E. (2019). "Impacts of the Convention on International Trade in Endangered species of Wild Fauna and Flora (CITES) on various stakeholders in the music industry," in *Swiss Wood Solution*. (Swiss: Swiss Wood Solution).

Wiedenhoeft, A. C. (2020). The XyloPhone: toward democratizing access to high-quality macroscopic imaging for wood and other substrates. *IAWA J.* 41, 699–719. doi: 10.1163/22941932-bja10043

Wiedenhoeft, A. C., Simeone, J., Smith, A., Parker-Forney, M., Soares, R., and Fishman, A. (2019). Fraud and misrepresentation in retail forest products exceeds US forensic wood science capacity. *PloS One* 14, e0219917. doi: 10.1371/journal.pone.0219917

Woo, S., Park, J., Lee, J. Y., and Kweon, I. S. (2018). "Cbam: Convolutional block attention module," in *In Proceedings of the European conference on computer vision (ECCV)*. 3–19.

Wu, F., Gazo, R., Haviarova, E., and Benes, B. (2021). Wood identification based on longitudinal section images by using deep learning. *Wood Sci. Technol.* 55, 553–563. doi: 10.1007/s00226-021-01261-1

Yin, Y. F., Jiao, L. C., He, T., and Jiang, X. M. (2022). *Identification manual of endangered and precious timber species common in trades* (Beijing: Science Press).