

Original papers

Advanced wood species identification based on multiple anatomical sections and using deep feature transfer and fusion

Kallil M. Zielinski ^a ^{*}, Leonardo Scabini ^a, Lucas C. Ribas ^b, Núbia R. da Silva ^c, Hans Beeckman ^d, Jan Verwaeren ^e, Odemir M. Bruno ^{a,f}, Bernard De Baets ^f 

^a São Carlos Institute of Physics, University of São Paulo, Brazil

^b Institute of Biosciences, Humanities and Exact Sciences, São Paulo State University, Brazil

^c Federal University of Catalão, Brazil

^d Royal Museum for Central Africa, Belgium

^e BIOVISM, Dept. of Data Analysis and Mathematical Modelling, Ghent University, Belgium

^f KERMIT, Dept. of Data Analysis and Mathematical Modelling, Ghent University, Belgium



ARTICLE INFO

Keywords:

Wood species identification
Texture analysis
Transfer learning
Convolutional neural networks
Feature extraction

ABSTRACT

Wood is a versatile and renewable resource, widely used across industries, yet the increasing demand has led to illegal logging with severe environmental, social, and economic consequences. To reduce illegal wood trade and its associated threats to biodiversity, robust methods for wood species identification and accurate datasets are crucial. In recent years, there have been significant advances in this area, but many current techniques face challenges such as high costs, the need for skilled experts for data interpretation, and the lack of good datasets for professional reference. Therefore, most of these methods, and certainly the wood anatomical assessment, may benefit from tools based on Artificial Intelligence. In this paper, we apply two transfer learning techniques with Convolutional Neural Networks (CNNs) to a multi-view Congolese wood species dataset including sections from different orientations and viewed at different microscopic magnifications. We explore two feature extraction methods in detail, namely Global Average Pooling (GAP) and Random Encoding of Aggregated Deep Activation Maps (RADAM), for efficient and accurate wood species identification. Our results indicate superior accuracy on diverse datasets and anatomical sections, surpassing the results of other methods. Our proposal represents a significant advancement in wood species identification, offering a robust tool to support the conservation of forest ecosystems and promote sustainable forestry practices.

1. Introduction

Wood is a versatile and renewable resource that can be produced in a sustainable way. It is widely used in many industries such as construction, furniture and paper production. The global demand for wood has led to the emergence of illegal logging and trade, having environmental, social, and economic repercussions. Illegal wood trade represents a significant portion of global wood trade, with percentages increasing in regions such as Southeast Asia, Central Africa, and South America (May and Global Financial Integrity, 2017). This illicit trade, worth billions of dollars annually, also threatens ecosystems due to the over-exploitation of rare and protected species. To combat this issue, various protection measures, such as the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) (United Nations, 2024), and policy measures like the European

Union Timber Regulation (EUTR) and the U.S. Lacey Act have been implemented (ITTO, 2020).

The effective implementation of the above policies and regulations also requires efficient methods for identifying wood species as well as robust datasets. Currently, wood species identification is primarily done through wood anatomical analysis, which involves the examination of tissue and cell diagnostic features using various imaging tools such as hand lenses, light or electronic microscopes, 2D and 3D scans, among others. Also, the International Association of Wood Anatomists (IAWA) has developed a list of standardized microscopic diagnostic features that can be used to identify hardwood species based on anatomical patterns, such as vessels, rays, parenchyma, and fibers (Wheeler et al., 1989). Although this approach is widely applied, readily available, and cost-effective, it can sometimes fail to distinguish between closely related taxa or determine the exact species (Dormontt et al., 2015; Gasson, 2011).

* Corresponding author.

E-mail address: kallil@usp.br (K.M. Zielinski).

<https://doi.org/10.1016/j.compag.2024.109867>

Received 12 April 2024; Received in revised form 24 October 2024; Accepted 21 December 2024

Available online 20 January 2025

0168-1699/© 2024 Published by Elsevier B.V.

Alternative methods for wood species identification have been gradually developed, including DNA analysis, Near Infrared spectroscopy, and Direct Analysis in Real Time (DART) mass spectrometry (Braga et al., 2011; Hassold et al., 2016; Pastore et al., 2011; Price et al., 2021; Jiao et al., 2018). These methods show promising results but are still hindered by factors such as high costs, the need for skilled experts for data interpretation, and the lack of reference datasets. Recently, pattern recognition techniques employing machine vision for automated wood species identification have emerged as a feasible and attractive solution. This approach is less dependent on expert knowledge and can leverage existing datasets containing high-quality microscopy images (Nithaniyal et al., 2014; Hanssen et al., 2011).

The state-of-the-art in wood species identification has seen significant progress through the incorporation of automated classification techniques based on macroscopic and microscopic images. Several studies (Zhao et al., 2014; Guang-Sheng and Peng, 2013; Khalid et al., 2011) have indeed reported promising results. However, these studies either focus on a limited number of species or rely on morphological wood features, which are dependent on segmentation and may consequently yield variable outcomes.

Texture analysis has emerged as a promising technique, as it can describe the spatial organization of pixels and the variation of patterns in an area on the surface of the studied object. Filho et al. (2014), Wang et al. (2013) used texture attributes derived from macroscopic images for wood species identification. Martins et al. (2012), Cavalin et al. (2013) employed texture features to identify wood species from the Brazilian flora using microscopic transverse cross-sections. These studies made use of Local Phase Quantization (LPQ), Local Binary Patterns (LBP) and gray-level co-occurrence matrices as feature descriptors.

Various computer vision models have been employed to automate wood species identification using digital imagery of anatomical sections. These models typically involve collecting a representative dataset of labeled digital images, applying a feature extraction procedure, and training a machine learning classification algorithm. Martins et al. (2013) achieved an accuracy of 86% using LBP as a feature descriptor combined with Support Vector Machines (SVMs). Filho et al. (2014) used a strategy of dividing the image into sub-images, classifying them independently using SVMs, and fusing the class probabilities through a fusion rule, achieving an accuracy of 97.77% for 41 different species.

Recent studies have also employed deep convolutional neural networks (CNNs) for wood species identification. Ravindran et al. (2018) obtained an accuracy of 87.4% using CNNs on a dataset of 2303 macroscopic images from the Meliaceae family. Another work (Ravindran et al., 2020) proposed a CNN model with a ResNet34 backbone and two linear layers to identify 12 common wood species in the United States based on macroscopic imagery of transverse sections reaching an accuracy of 97.7%. Lens et al. (2020) proposed four different pre-trained CNN architectures, achieving a similar accuracy of over 98% on 2240 images from 112 species using the ResNet101 backbone. Wu et al. (2021) utilized two highly effective CNNs (ResNet50 and DenseNet121) for hardwood lumber identification, reaching an accuracy of 98.2% on 11 common hardwood species classification tasks.

The application of transfer learning has become increasingly effective in wood species identification. This method uses models pre-trained on large datasets, allowing for fine-tuning for specific wood textures. For instance, Tajbakhsh et al. (2020) showed that transfer learning can enhance the performance of deep learning models even when limited labeled data are available. Zhao et al. (2019) employed transfer learning with a pre-trained CNN model for wood species identification, achieving an accuracy of 95.2% on a dataset of 1832 macroscopic images from 32 species. This approach not only yielded impressive classification results but also reduced the time and computational resources required for training the model. Recent work (Herrera-Poyatos et al., 2024) adopted this approach for wood species identification with NIR spectroscopy and CNNs. Nieradzki et al. (2024) used a YOLO network to detect key vessel elements in microscopic wood images. Zheng et al.

(2024) employed a database of wood species containing two anatomical sections to classify 15 different wood species using a pre-trained backbone followed by a Region CNN (R-CNN) method.

Despite significant advances in wood species identification, the African continent, particularly the Congo Basin, remains underrepresented. To address this gap, we leverage the timber species dataset from the Democratic Republic of Congo (DRC) provided by the Belgian Royal Museum for Central Africa (RMCA) (Biodiversity, 2019), focusing on texture features extracted from distinct microscopic cross sections, which have been shown to have strong discriminative abilities in works by da Silva et al. (2017), da Silva et al. (2022). The present work applies pre-trained CNN models to this dataset to enhance wood species identification. While previous methods, such as the LPQ approach (Ojansivu and Heikkilä, 2008) combined with random forest classifiers, have shown promising results (da Silva et al., 2022), we propose two approaches that use pre-trained models to further improve the efficiency and accuracy of wood species identification.

The first approach employs a serial activation map fusion using Global Average Pooling (GAP) on a backbone CNN model that has been pre-trained on ImageNet (Deng et al., 2009). This allows us to benefit from a richer set of feature representations learned on diverse images. The second approach involves the use of RADAM (Random encoding of Aggregated Deep Activation Maps), a feature extractor based on randomized auto-encoders that has demonstrated state-of-the-art accuracy on texture recognition tasks (Scabini et al., 2023). By incorporating RADAM into our wood species identification pipeline, we aim to further optimize the classification performance by exploiting the advanced learning capabilities of this approach.

Through the implementation of the above approaches, we seek to improve upon the results achieved by da Silva et al. (2022) and contribute to the development of more efficient and more accurate wood species identification models, ultimately promoting the protection and conservation of forest ecosystems.

2. Materials and methods

This section presents the materials and methods used in our study, including the dataset, the feature extraction methods adopted and the parameters of the experimental configurations employed. The following subsections provide a detailed overview of these components.

2.1. Datasets

This research uses the timber species dataset from the DRC, which was assembled by da Silva et al. (2017). This dataset includes 77 different Congolese timber species, which are listed in Table 1 along with the number of images per class. Each species has images from three distinct anatomical planes: tangential, transversal and radial. The sections were cut with a sliding microtome, dehydrated in a graded ethanol series (50%, 75%, 96% and 100%), and fixed with Euparal. RGB images were acquired using a light microscope (Olympus BX60) using CellB software at 2.5× magnification. Each image has a resolution of 1000 × 1000 pixels corresponding to 1388.88 × 1388.88 μm. This approach offers a comprehensive view of each species, facilitating a more accurate identification. In summary, the dataset is composed of 805 images for each anatomical plane. Fig. 1 displays an illustrative representation of the images from each anatomical section in this dataset, which shows the different visual characteristics present in each anatomical plane.

The dataset contains an average of 10 images per class, which is relatively low for machine learning tasks. To overcome this limitation, da Silva et al. (2022) applied data augmentation techniques to increase the number of images per class. Starting from the original dataset consisting of 1000 × 1000 pixel images, they created three additional datasets with different image dimensions and transformations:

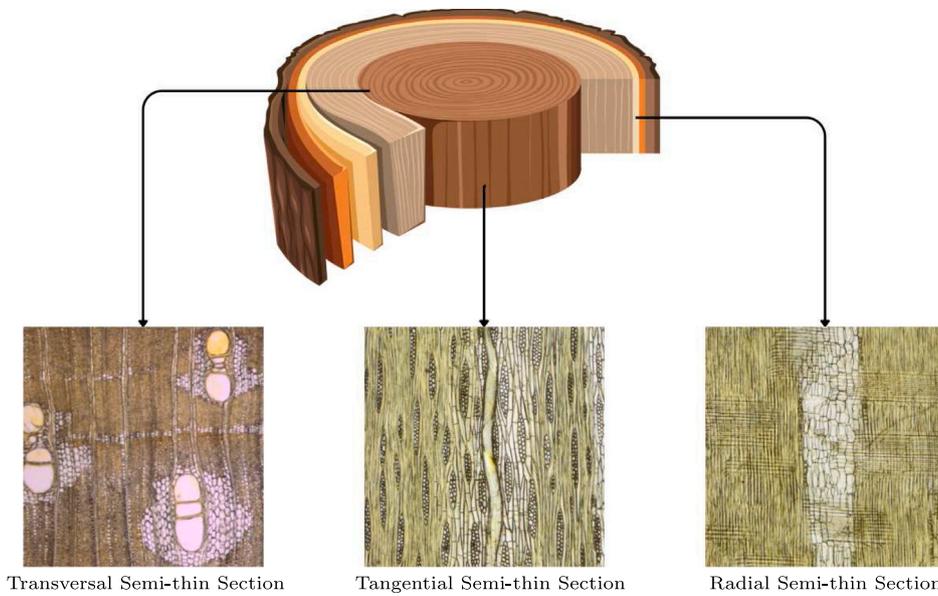


Fig. 1. Representation of each anatomical section of a wood sample from the species *Afzelia africana*. The arrows point from the respective regions on the sample (transversal, tangential, and radial) to their corresponding images. Each image offers unique structural characteristics that are crucial for wood species identification.

Table 1
Names of the species and number of images for each anatomical section of the dataset.

Species	Images	Species	Images
Afzelia africana	9	Afzelia bella	5
Afzelia bipindensis	8	Afzelia cuanzensis	8
Afzelia pachyloba	8	Albizia adianthifolia	17
Albizia antunesiana	10	Albizia ferruginea	14
Alstonia boonei	12	Amphimas ferrugineus	8
Amphimas pterocarpoides	9	Anthonotha macrophylla	7
Antiaris toxicaria	12	Antrocaryon nannanii	17
Autranella congolensis	8	Beilschmiedia congolana	10
Brachystegia laurentii	7	Canarium schweinfurthii	13
Ceiba pentandra	6	Celtis gomphophylla	11
Chrysophyllum africanum	4	Chrysophyllum lacourtianum	8
Copaifera mildbraedii	13	Cordia platythyrsa	8
Cynometra alexandri	15	Cynometra hankei	10
Diospyros crassiflora	10	Drypetes gossweileri	10
Ekebergia capensis	8	Entandrophragma angolense	20
Entandrophragma candollei	13	Entandrophragma cylindricum	14
Entandrophragma utile	17	Erythrophleum suaveolens	6
Ficus mucoso	8	Funtumia africana	15
Gilbertiodendron dewevrei	11	Guibourtia arnoldiana	8
Guibourtia demeusei	9	Hallea stipulosa	17
Holoptelea grandis	12	Iringia grandifolia	14
Khaya anthotheca	14	Klainedoxa gabonensis	9
Leplaea cedrata	15	Leplaea laurentii	20
Leplaea thompsonii	5	Lophira alata	4
Lovoa trichilioides	11	Mammea africana	10
Milicia excelsa	12	Milletia laurentii	10
Morus mesozygia	7	Musanga cecropioides	12
Nauclea diderrichii	12	Nesogordonia kabingaensis	8
Newtonia leucocarpa	7	Ongokea gore	10
Pentaclethra eetveldeana	7	Pentaclethra macrophylla	9
Pericopsis elata	5	Petersianthus macrocarpus	11
Piptadeniastrum africanum	12	Pouteria aningeri	8
Prioria balsamifera	12	Prioria oxyphylla	14
Pterocarpus soyauxii	17	Pterocarpus tinctorius	10
Pycnanthus angolensis	4	Scorodophleus zenkeri	8
Staudtia kamerunensis	13	Terminalia superba	9
Tessmannia africana	13	Tieghemella heckelii	9
Triplochiton scleroxylon	10	Zanthoxylum gillettii	7
Zanthoxylum lemairei	12		

- **1000 × 500:** Each image from the original dataset was split into two halves, doubling the total number of images;
- **500 × 500:** The 1000 × 500 images were further split into four sections, quadrupling the number of images from the original dataset;
- **500 × 500-OGRN:** Similar to the previous dataset, but with transformations: the first image remains unaltered (O), the second is smoothed with a Gaussian Filter (G), the third is rotated 90 degrees (R) and the fourth has added salt-and-pepper noise with a density of 0.05 (N).

These augmented datasets not only increase the quantity of training data but also introduce variations that enhance the robustness of machine learning models. Fig. 2 demonstrates the transformations applied to the original 1000 × 1000 images.

2.2. Feature extraction methods

Texture analysis using CNNs typically adopts end-to-end models that leverage a pre-trained backbone, followed by fine-tuning of the architecture specific to the texture recognition problem at hand. In contrast, this paper employs two approaches that use the backbone only as a feature extractor with a dedicated classifier applied to the resulting features. Notably, this process eliminates the need for fine-tuning of the architecture.

2.2.1. GAP

The first approach considered involves using GAP computed over the activation maps produced by the last layer of the backbone, which are subsequently fed to an SVM classifier (Cortes and Vapnik, 1995). The function representing the pre-trained backbone is denoted as f , and the input image from each of the anatomical planes (tangential, transversal, and radial) as X_T , X_{Tr} , and X_R , resp., with $X_T, X_{Tr}, X_R \in \mathbb{R}^{w_0 \times h_0 \times 3}$. Features are extracted from the last layer of the backbone as follows:

$$Z_T = f(X_T), \quad Z_{Tr} = f(X_{Tr}), \quad Z_R = f(X_R). \quad (1)$$

For each wood sample, there are three feature vectors corresponding to the three anatomical planes. These features can be combined in two ways:

- **1000 × 1000 (Original):** This dataset contains 805 images per anatomical section with a resolution of 1000 × 1000 pixels;

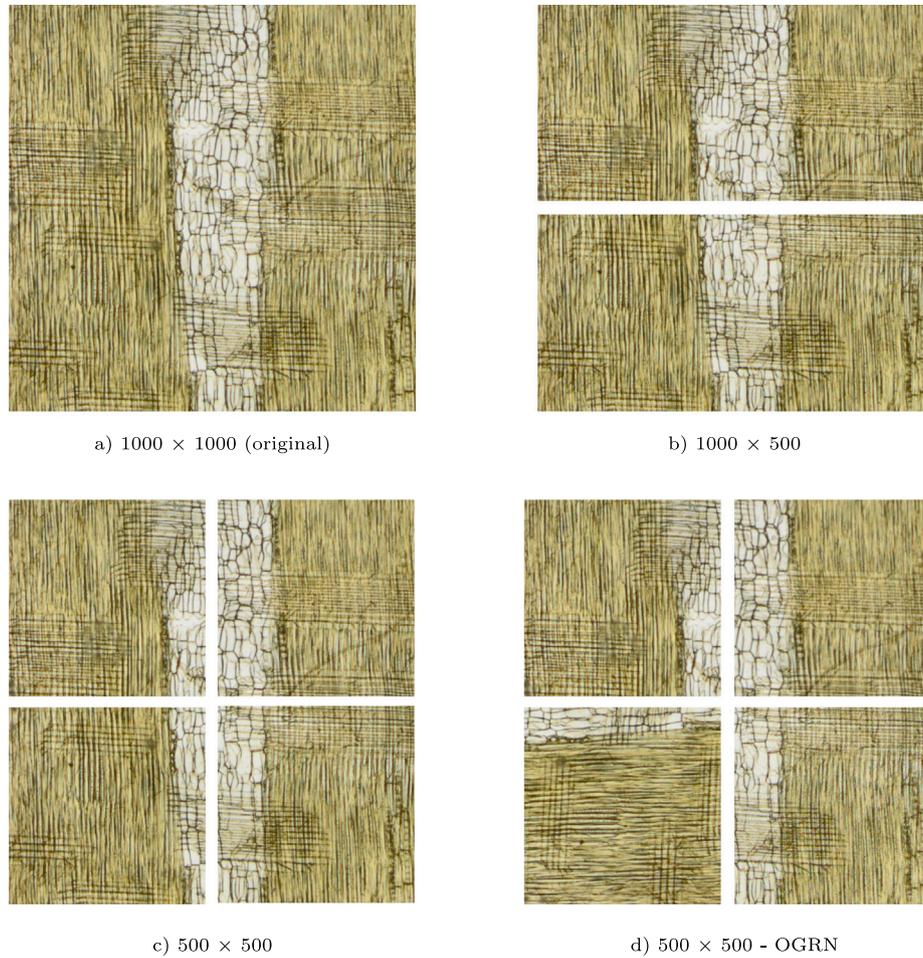


Fig. 2. Illustration of the data augmentation techniques employed for each sub-dataset. (a) 1000 × 1000 original dataset, (b) 1000 × 500 dataset with each image halved, (c) 500 × 500 dataset with each original image split into four equal parts, and (d) 500 × 500 - OGRN dataset demonstrating the applied transformations including original (O), Gaussian smoothing (G), 90-degree rotation (R), and salt-and-pepper noise addition (N).

- 1. Serial Feature Concatenation (SFC):** We concatenate the feature vectors end-to-end. This can be defined mathematically as:

$$Z_s = [Z_T, Z_{Tr}, Z_R]. \quad (2)$$

- 2. Parallel Feature Merging (PFM):** We perform an element-wise summation of the feature vectors:

$$Z_p = Z_T \oplus Z_{Tr} \oplus Z_R. \quad (3)$$

Finally, the combined features (Z_s or Z_p) are used as input to an SVM classifier. Let h denote the SVM, then the predicted class label is given by:

$$\hat{y} = h(Z_s) \quad \text{or} \quad \hat{y} = h(Z_p). \quad (4)$$

2.2.2. RADAM

The second approach adopted is RADAM (Scabini et al., 2023), which significantly differs from GAP. Instead of considering only the last layer of a pre-trained backbone, RADAM takes into account multiple activation maps at different depths in the architecture. This approach allows to capture various levels of texture characteristics, from simple to complex features.

In the RADAM method, the features are combined through a specific sequence of steps, including a Randomized Autoencoder (RAE) (Kasun et al., 2013; Cambria et al., 2013). A RAE is a type of neural network that consists of one hidden layer and aims to reproduce its input as target output, i.e., $Y = X$. The input weights of the RAE are randomly

generated using a Linear Congruential Generator (LCG) (Knuth, 1997). After training the RAE using the least-squares method, the output weights are used as features in our approach, as described below.

Each step of the RADAM method is described as follows. The function representing the pre-trained backbone is denoted as f , and its intermediate activation maps at different depths i as f_i . Each depth level corresponds to the output of a specific convolutional block in the backbone. For each input image of the three anatomical planes (X_T , X_{Tr} and X_R) and each depth i , the corresponding activation map is extracted:

$$A_{T,i} = f_i(X_T), \quad A_{Tr,i} = f_i(X_{Tr}), \quad A_{R,i} = f_i(X_R). \quad (5)$$

Next, the depthwise $2p$ -norm of the activation maps is computed, which can be denoted as g . The computed norms for the three anatomical planes at depth i are given by:

$$N_{T,i} = g(A_{T,i}), \quad N_{Tr,i} = g(A_{Tr,i}), \quad N_{R,i} = g(A_{R,i}). \quad (6)$$

These norms are then concatenated along the third dimension (channel dimension z_i). However, since the spatial dimensions (w_i and h_i) of the activation maps might differ across different depths, the spatial dimension of all activation maps are resized to the same spatial size using bilinear interpolation. The concatenations of the normalized activation maps for the three anatomical planes are denoted as:

$$C_T = \bigoplus_{i=1}^n N_{T,i}, \quad C_{Tr} = \bigoplus_{i=1}^n N_{Tr,i}, \quad C_R = \bigoplus_{i=1}^n N_{R,i}, \quad (7)$$

where \bigoplus denotes the concatenation operation along the third dimension for each computed anatomical section norm, with i varying from 1

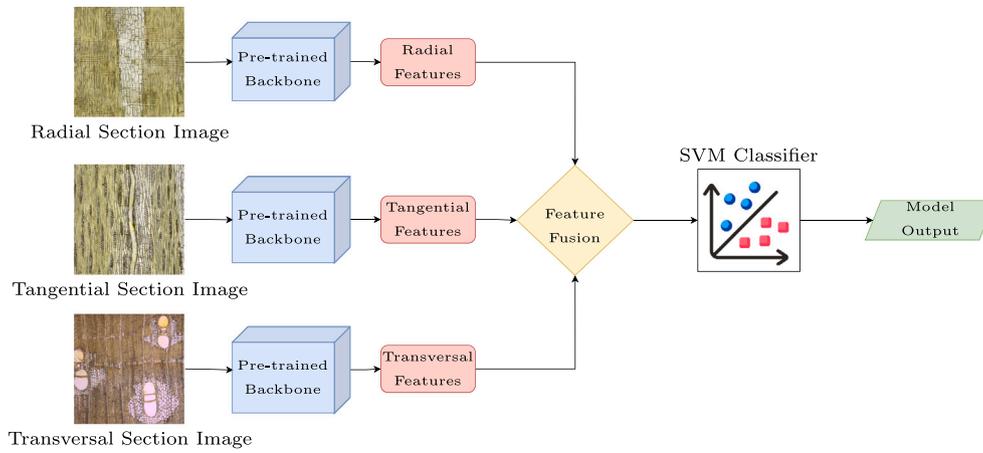


Fig. 3. Methodological process employed in this study. Anatomical sections are input to a pre-trained backbone to extract feature blocks. These blocks are then fused and classified using an SVM to produce the final wood species prediction.

Table 2

Number of parameters for each backbone, respective GFLOPs, and average image processing time for GAP and RADAM methods.

Backbone	N. of params (millions)	GFLOPs	Avg. image processing time (ms)	
			GAP	RADAM
ResNet-18	11.7	1.8	2.565	9.757
ResNet-50	25.5	4.1	5.979	24.735
ConvNext Large	230	34.4	35.509	239.738
ConvNext XLarge	392	60.9	61.105	316.939

to n , and n representing the number of depths considered in the RADAM approach.

Next, a set of m RAEs are applied to the concatenated activation maps C_T , C_{Tr} and C_R . The projection obtained by the output weights of each RAE is considered as the encoded representation of the activation map. This representation is denoted as r_j for the j th RAE:

$$Z_{T,i} = r_j(C_T), \quad Z_{Tr,j} = r_j(C_{Tr}), \quad Z_{R,j} = r_j(C_R). \quad (8)$$

After repeating this process for m RAEs, the representations are combined using the PFM strategy, summing the features for each anatomical section image:

$$Z_T = \sum_{j=1}^m Z_{T,j}, \quad Z_{Tr} = \sum_{j=1}^m Z_{Tr,j}, \quad Z_R = \sum_{j=1}^m Z_{R,j}. \quad (9)$$

In a last step before inputting these features into a classifier, the features of all three anatomical sections are combined using the SFC or PFM strategy just as for GAP (see Eqs. (2) and (3)). Finally, these combined features serve as input to an SVM classifier, which performs the final classification task.

In order to facilitate the understanding of the overall process, Fig. 3 shows the methodology adopted in this paper. The individual anatomical sections (radial, tangential and transversal) are first input through the pre-trained backbone, resulting in their respective feature blocks. These features are then combined through SFC or PFM, forming a unified feature representation. This feature representation is then passed through an SVM classifier for final wood species classification. The output of the model is the predicted wood species, derived from the combined and classified features of the three anatomical sections.

2.3. Experimental configuration

This section describes the experimental setup, including the libraries, parameters of the methods and classification algorithms used in this study.

2.3.1. Hardware and software

All experiments were performed on a machine with a GTX 1080ti, Intel Core i7-7820X 3.60 GHz processor (using 8 threads), and 64 GB of RAM. The feature extraction methods were implemented using PyTorch (Paszke et al., 2019). To investigate the performance of some backbones on our dataset, the library Pytorch Image Models (TIMM) (Wightman, 2019) was employed, which includes a wide-range of pre-trained computer vision backbones. The classification process was realized through the use of the SVM implementation in the Scikit-learn (Pedregosa et al., 2011) library.

2.3.2. Backbones

Experiments were made with four different pre-trained backbones for our feature extraction methods. Specifically, ResNets (He et al., 2016) (18 and 50) were used, both of which provide five distinct depths for feature extraction. In addition to ResNets, two variants of the ConvNext architecture (Liu et al., 2022) were considered: ConvNext-Large-In21k and ConvNext-XLarge-In21k. Both are pre-trained on the ImageNet-21k dataset, and have four different feature extraction depths. To further detail the complexity of each backbone, Table 2 outlines their number of parameters and GFLOPs.

The average image processing time for both GAP and RADAM was measured using a batch size of 1, with the sampling based on the average processing time for all images across a complete epoch. The experiments were conducted on the GPU detailed in Section 2.3.1. These results highlight the computational efficiency of each backbone when used with the respective feature extraction methods. Notably, while we compare dataset sizes of 1000×1000 , 1000×500 , and 500×500 , the processing time remained consistent, as each image is resized to 224×224 before input into the CNN. Consequently, the observed increase in total processing time when comparing larger datasets (e.g., 1000×1000 vs. 1000×500) is proportional to the number of images rather than their original dimensions, since the feature extraction process is applied to each image independently.

2.3.3. RADAM and feature extraction configuration

The feature extraction methods in this study are applied to images with a 224×224 resolution to reduce the computational cost. In addition, the RADAM method incorporates the use of 4 RAEs, following the configuration suggested by the authors (Scabini et al., 2023), which provides a balance between performance and computational cost.

The performance of the different backbones was evaluated across three combinations of anatomical sections: Transversal only, Transversal + Tangential and Transversal + Tangential + Radial. Moreover, for each of these section combinations, the performance of both feature fusion strategies described in Section 2.2, i.e., Serial Feature Concatenation (SFC) and Parallel Feature Merging (PFM) was also compared.

Table 3

Accuracies (expressed as percentages) and standard deviations of SVM using GAP features for each individual anatomical section, dataset and backbone. The best results on each dataset are highlighted in bold.

Dataset	Section	Backbone			
		Resnet18	Resnet50	Convnext Large	Convnext XLarge
Original	Tr	67.00±0.90	61.70±0.50	79.90±1.10	82.20±1.10
	T	74.50±0.60	67.30±0.80	87.30±0.80	88.20±0.60
	R	57.30±0.90	52.40±0.70	75.40±1.20	77.60±0.80
1000 × 500	Tr	85.70±0.50	82.70±0.40	96.10±0.30	96.30±0.40
	T	82.60±0.40	82.20±0.20	95.00±0.60	95.40±0.40
	R	75.30±0.50	73.30±0.40	91.00±0.40	91.50±0.40
500 × 500	Tr	91.00±0.30	90.20±0.40	98.10±0.10	98.70±0.20
	T	88.30±0.40	89.60±0.30	97.40±0.20	97.50±0.10
	R	76.90±0.30	79.00±0.20	93.10±0.20	92.90±0.30
500 × 500-OGRN	Tr	78.40±0.50	74.90±0.80	91.50±0.40	92.60±0.50
	T	79.60±0.70	78.50±0.30	91.50±0.50	93.50±0.40
	R	60.60±0.50	61.70±0.50	82.80±0.50	83.70±0.50

2.3.4. Classification process

The SVM was used as the primary classification algorithm for this study with a linear kernel and $C = 1$, and no hyperparameter tuning was done. For each of the datasets detailed in Section 2.1, a 10-fold cross-validation strategy was used, and the classification process was repeated 10 times to ensure a comprehensive evaluation of the SVM's performance across datasets. Furthermore, the average accuracy and its standard deviation was computed over the 10 repetitions. These were the key metrics for assessing the performance of our proposed models. This provides a balanced view of the model's performance, taking into account both the model's correctness and consistency across different runs.

3. Results

This section present and discuss the results obtained with each of the two feature extraction models.

3.1. GAP results

In the first phase of the investigation with GAP, the approach was explored by analyzing the accuracies achieved for each individual anatomical section, across all datasets and backbones. This examination provides detailed insights into the strength and performance of GAP when applied to separate anatomical sections and aids in identifying the contribution of each section toward the overall accuracy in species identification. The results for each anatomical section (Transversal (Tr), Tangential (T), and Radial (R)) are presented in Table 3, across all four datasets: Original (1000 × 1000), 1000 × 500, 500 × 500 and 500 × 500-OGRN and for each backbone in Table 2.

The results indicate that the Tr and T anatomical sections provide more discriminative power for GAP in comparison to the radial (R) section. This could be attributed to the unique cellular arrangement and structures visible in the transversal and tangential sections, which may provide more distinctive features for the identification process. The radial section, while still providing valuable insights, seems to exhibit a less discriminative performance, possibly due to its inherent similarity among various species.

There are also some interesting insights that can be derived from the classification results of the individual anatomical sections.

The results for each anatomical section show different misclassification patterns. For instance, for the transversal section (Fig. 4a), *Chrysophyllum africanum* tends to be confused with *Pentaclethra eetveldeana*. For the tangential section (Fig. 4b), confusion between *Klainedoxa gabonensis* and *Lophira alata* is frequent. Lastly, the radial section (Fig. 4c) demonstrates lower classification accuracy, with *Newtonia leucocarpa* being misclassified due to less distinct anatomical characteristics. These results indicate that certain anatomical sections, such as the transversal and tangential sections, provide more discriminative features for species identification, while the radial section poses more challenges for

the model. Misclassifications can often be attributed to similarities in cellular structures, such as vessel arrangement in the transversal section and parenchyma structures in the tangential section.

Building upon these observations, a further analysis on the use of combinations of anatomical sections was done, specifically transversal with tangential (Tr + T) and also all sections combined (Tr + T + R). The goal is to investigate whether the complementary information from multiple sections can enhance the accuracy, and if so, to what extent. Incorporating more than one anatomical section allows the model to harness a broader spectrum of wood features, potentially improving the overall classification performance. The corresponding results are listed in Table 4.

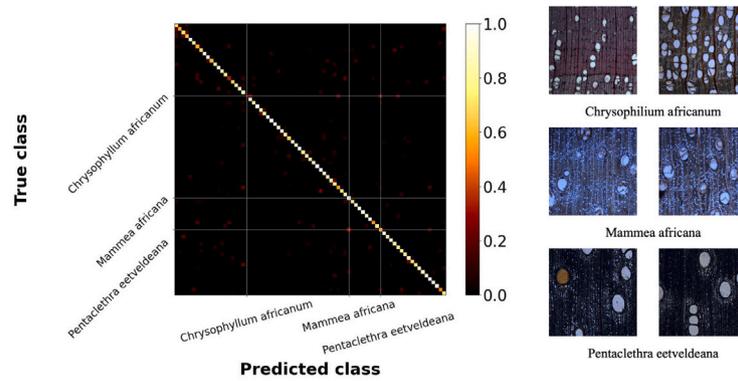
The discussion of the results begins with the model's performance on the original dataset, which has fewer images per class compared to the other datasets. As the results suggest, this dataset yielded the lowest accuracy among all datasets, with the highest accuracy of 93.2 ± 0.6 obtained when all the anatomical sections are considered using the PFM fusion strategy and the ConvNext XLarge backbone.

In contrast, the accuracy of the model dramatically improved when increased the number of images per class in the datasets. For instance, on the 1000 × 500 dataset, it reached an accuracy of 99.3 ± 0.2 (all anatomical sections, SFC fusion strategy and ConvNext XLarge backbone), and on the 500 × 500 dataset, it reached an even higher performance of 99.8 ± 0.0 (all anatomical sections, SFC fusion strategy and ConvNext Large). Notably, on the 500 × 500-OGRN dataset, which contains the same number of images but incorporates noise, the model maintained a high accuracy of 98.2 ± 0.2 when using the ConvNext XLarge backbone. This result underlines the observation that increasing the number of samples per class in a dataset significantly enhances the performance of the model.

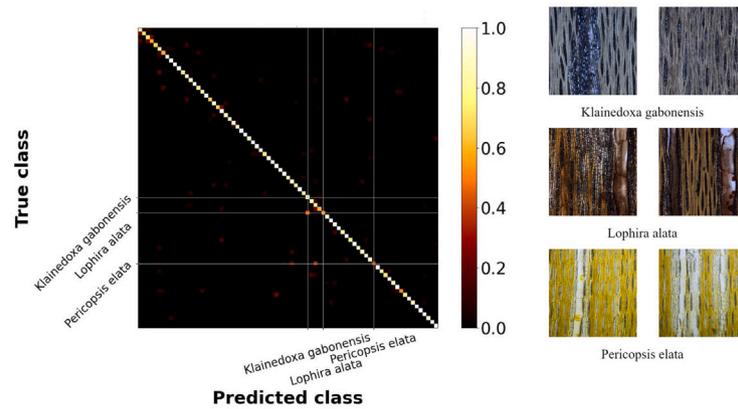
Another factor that has a positive impact on the model performance is the inclusion of more anatomical sections in the analysis. This is clear from the results obtained on the original dataset using the ConvNext Large backbone. When we considered the transversal section (see Table 3) and the ResNet18 backbone, the accuracy was 67.0 ± 0.90 . However, the accuracy increased to 85.7 ± 0.5 upon the inclusion of the tangential section and further increased to 87.4 ± 0.8 when also the radial section was incorporated.

In terms of feature fusion strategies, one can see that SFC generally outperforms PFM. However, it should be noted that the SFC strategy results in a larger feature vector due to concatenation, which may, sometimes, be a disadvantage as it can reduce the efficiency of the classification step. On the other hand, PFM allows for a better control of the number of features regardless of the number of terms involved in the summations.

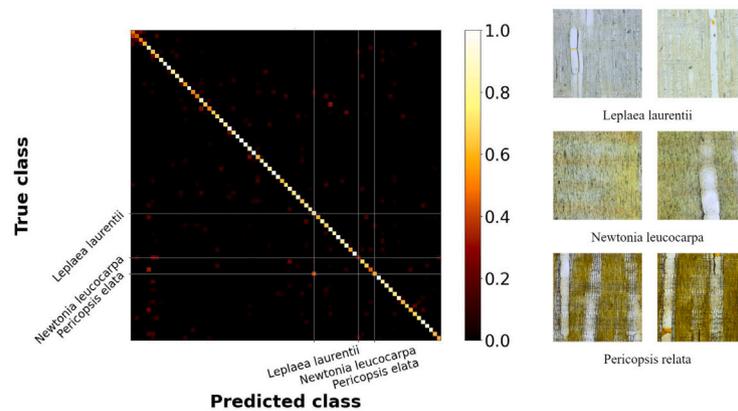
Finally, the comparison of different backbones reveals some interesting insights. Although ResNet18 generally outperforms ResNet50, especially when using SFC, the ConvNext backbones display superior performance across all datasets, regardless of the combination of



(a) Confusion matrix and class samples for the GAP results on the transversal anatomical section.



(b) Confusion matrix and class samples for the GAP results on the tangential anatomical section.



(c) Confusion matrix and class samples for the GAP results on the radial anatomical section.

Fig. 4. Confusion matrices for the GAP results across the three anatomical sections: (a) Transversal, (b) Tangential, and (c) Radial.

sections and the fusion strategy considered. Given that they have a considerably larger size, better pre-training and a more advanced architecture design compared to the ResNets, this was an expected result.

3.2. RADAM results

This section details the results obtained with the RADAM feature extraction method. Similar to GAP, the performance was evaluated across different datasets starting with the analysis of individual anatomical sections. The results are listed in Table 5.

The results obtained with RADAM mirror the previous observations for GAP. Both the Tr and T sections appear to have stronger discriminative capabilities compared to the R section. The superior performance of these sections may be attributed to their distinctive cellular structures and arrangements that RADAM effectively captures, thereby providing more characteristic features for the wood species identification task.

The results for each anatomical section in the RADAM approach highlight distinct patterns of misclassification. For the transversal section (Fig. 5a), *Chrysophyllum africanum* is frequently confused with *Pentaclethra eetveldeana*, possibly due to similarities in vessel arrangement. The tangential section (Fig. 5b) shows accurate identification of

Table 4

Accuracies (expressed as percentages) and standard deviations of SVM using GAP features for different combinations of anatomical sections, and each fusion strategy, dataset and backbone. The best results on each dataset are highlighted in bold.

Dataset	Combination	Fusion Strategy	Backbone			
			ResNet18	ResNet50	ConvNext Large	ConvNext XLarge
Original	Tr+T	SFC	85.70±0.50	76.90±0.60	91.20±0.70	91.80±0.70
		PFM	84.70±0.80	78.30±0.60	90.5±0.60	91.30±0.40
	Tr+T+R	SFC	87.40±0.90	80.80±0.50	92.50±0.50	93.10±0.70
		PFM	86.40±0.90	83.0 ±0.50	92.10±0.60	93.20±0.60
1000 × 500	Tr+T	SFC	95.10±0.50	93.00±0.50	98.90±0.20	99.00±0.30
		PFM	92.80±0.60	93.20±0.50	98.70±0.40	98.60±0.40
	Tr+T+R	SFC	96.80±0.40	94.50±0.40	99.20±0.30	99.30±0.20
		PFM	95.00±0.40	94.90±0.40	98.80±0.20	98.90±0.30
500 × 500	Tr+T	SFC	97.80±0.10	97.50±0.10	99.80±0.10	99.80±0.00
		PFM	96.00±0.20	97.00±0.10	99.50±0.10	99.60±0.10
	Tr+T+R	SFC	97.20±0.30	98.30±0.20	99.80±0.00	99.80±0.10
		PFM	97.00±0.20	97.70±0.10	99.60±0.10	99.70±0.10
500 × 500-OGRN	Tr+T	SFC	92.40±0.40	89.5±0.40	97.20±0.30	97.90±0.20
		PFM	89.60±0.50	89.80±0.50	96.30±0.30	97.50±0.20
	Tr+T+R	SFC	93.40±0.40	92.50±0.50	97.70±0.20	98.20±0.20
		PFM	91.10±0.40	92.50±0.50	97.00±0.30	97.90±0.20

Table 5

Accuracies (expressed as percentages) and standard deviations of SVM using RADAM features for each individual anatomical section, dataset and backbone. The best results on each dataset are highlighted in bold.

Dataset	Section	Backbone			
		ResNet18	Resnet50	ConvNext Large	ConvNext XLarge
Original	Tr	77.00±0.50	77.20±0.70	86.10±0.70	87.30±0.70
	T	82.00±0.40	83.70±0.70	90.30±0.90	90.50±0.60
	R	69.70±0.80	68.90±0.60	81.70±0.90	83.60±0.90
1000 × 500	Tr	94.80±0.40	94.30±0.50	97.90±0.30	98.30±0.30
	T	91.00±0.50	91.90±0.30	96.90±0.30	96.70±0.50
	R	88.80±0.40	88.70±0.70	95.00±0.30	95.30±0.30
500 × 500	Tr	97.30±0.20	97.50±0.20	99.10±0.10	99.40±0.10
	T	95.50±0.10	96.70±0.20	99.00±0.10	98.90±0.10
	R	90.80±0.30	91.90±0.20	96.80±0.20	96.20±0.10
500 × 500-OGRN	Tr	87.50±0.20	87.60±0.40	94.70±0.30	94.80±0.40
	T	88.20±0.50	90.00±0.30	95.50±0.30	96.40±0.30
	R	75.30±0.50	78.10±0.50	88.70±0.30	88.70±0.60

Leplaea laurentii, but with confusion between *Pericopsis elata* and *Leplaea laurentii*, likely due to similar parenchyma structures. Finally, the radial section (Fig. 5c) reveals significant misclassification between *Albizia adianthifolia* and *Newtonia leuocarpa*, possibly due to close ray and tracheid patterns. These results suggest that while the RADAM method is robust, certain anatomical sections, particularly the radial section, exhibit challenges in species differentiation. By leveraging information from multiple sections, classification accuracy can be improved across species.

These observations led further investigation into the combination of anatomical sections, specifically Tr + T and Tr + T + R, and also testing different feature fusion strategies, with the intention of complementing information from different sections to improve the classification accuracy. The results are shown in Table 6.

The discussion of the results begins with the model's performance on the original dataset, reaching its highest accuracy of 94.8 ± 0.5 when all anatomical sections are considered using the PFM feature fusion strategy and ConvNext XLarge.

However, a remarkable performance improvement was observed as the number of images per class in the datasets increased. Specifically, the model accuracy increased to 99.50 ± 0.2 on the 1000×500 dataset (all anatomical sections, SFC strategy and ConvNext XLarge), while on the 500×500 dataset it reached a nearly perfect score of 99.9 ± 0.0 (multiple configurations) and finally on the 500×500 -OGRN dataset the model maintained a good performance of 98.7 ± 0.2 (multiple configurations). These results, similar to the previous section, confirm the model's enhanced performance with an increase in the number of samples per class in the dataset.

Furthermore, the inclusion of more anatomical sections in the analysis also increased the accuracy of the RADAM model. For instance,

considering only the transversal section in the original dataset using the ResNet18 backbone yielded an accuracy of 77.0 ± 0.5 (see Table 5, which increased to 90.2 ± 0.7 when including the tangential section, and further improved to 91.2 ± 0.6 when incorporating the radial section, using the SFC strategy.

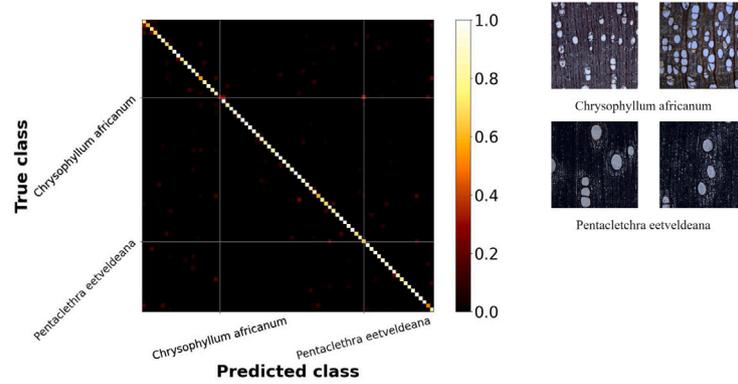
As for the feature fusion strategies, in general, SFC continues to be better than PFM, but we can note that there are some cases, for example, the ConvNext Large backbone on the original dataset with all anatomical sections, where PFM outperforms SFC.

Lastly, the results show a similar comparison of the backbones as in the previous section, where in many cases ResNet18 outperforms ResNet50, but in general the ConvNexts delivered the best performance across all datasets regardless of the combination of sections and the fusion strategy considered.

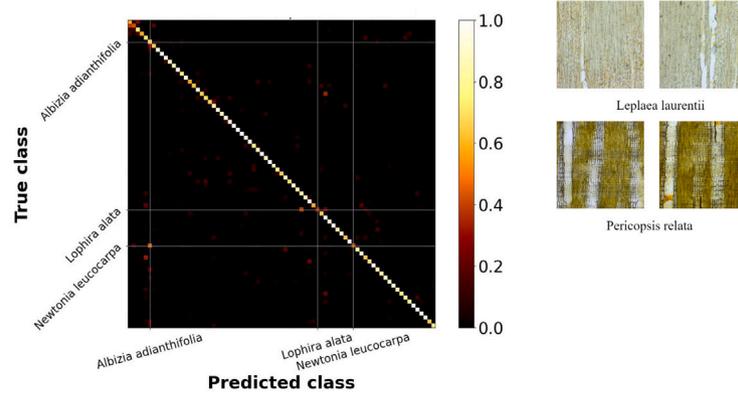
3.3. Overall comparison

This subsection presents the main results for the four distinct datasets, taking into account different combinations of anatomical sections for the methods introduced in this study, as well as the approach proposed by da Silva et al. (2022), which applies LPQ on each section, uses a random forest classifier on each section separately and further concatenates the probability matrices obtained in the classification process as input to a logistic regression model. A comprehensive comparison of these results is presented in Table 7.

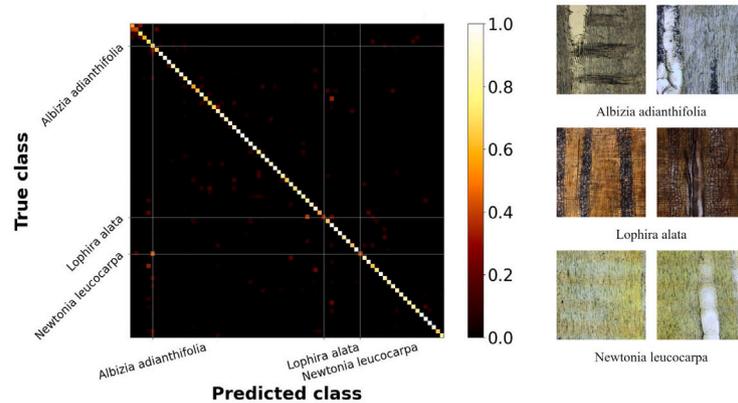
An observation that can be made from this comparison is that both methods proposed in this study consistently surpass the performance of the approach by da Silva et al. (2022), regardless of the dataset employed or the combination of anatomical sections considered. Furthermore, among our proposed methods, RADAM exhibits superior



(a) Confusion matrix for RADAM results and class samples on the transversal anatomical section.



(b) Confusion matrix for RADAM results and class samples on the tangential anatomical section.



(c) Confusion matrix for RADAM results and class samples on the radial anatomical section.

Fig. 5. Confusion matrices for RADAM results across the three anatomical sections: (a) Transversal, (b) Tangential, and (c) Radial.

performance compared to GAP, demonstrating its robustness across different datasets, anatomical sections and feature fusion strategies.

3.4. Leave-k-trees-out

In this study, the method's performance was also analyzed using the leave-k-trees-out cross-validation strategy, as used by da Silva et al. (2022). This cross-validation approach involves segregating all samples from a single tree for each species into the test set, thereby ensuring complete independence between the training and test sets in terms of specimens. Consequently, 165 samples were allocated for testing and 640 for training.

Table 8 presents a comparison among GAP, RADAM and the method proposed by da Silva et al. (2022). All methods were evaluated under the leave-k-trees-out cross-validation scheme. For our methods, the combination of all three anatomical sections and the SFC strategy were used.

The results show a notable performance disparity. While the previous method exhibited accuracies ranging from 22.00% to 30.00% across different datasets, both GAP and RADAM demonstrated significantly higher accuracies, with RADAM slightly outperforming GAP. These results suggest that our method, although showing a slight inferior performance compared to the traditional cross-validation configuration, still surpasses the performance of da Silva et al. (2022).

Table 6

Accuracies (expressed as percentages) and standard deviations of SVM using RADAM features for different combinations of anatomical sections, and each fusion strategy, dataset and backbone. The best results on each dataset are highlighted in bold.

Dataset	Combination	Fusion Strategy	Backbone			
			ResNet18	ResNet50	ConvNext Large	ConvNext XLarge
Original	Tr+T	SFC	90.20±0.70	89.00±0.60	93.50±0.50	94.00±0.60
		PFM	88.60±0.70	88.20±0.60	92.60±0.50	93.60±0.50
	Tr+T+R	SFC	91.20±0.60	90.60±0.40	94.00±0.70	94.30±0.70
		PFM	90.90±0.80	89.60±0.80	94.70±0.60	94.80±0.60
1000 × 500	Tr+T	SFC	98.10±0.40	98.00±0.30	99.40±0.30	99.30±0.30
		PFM	97.70±0.30	97.90±0.40	99.20±0.30	99.20±0.30
	Tr+T+R	SFC	98.50±0.40	98.60±0.30	99.60±0.20	99.50±0.20
		PFM	97.70±0.40	98.10±0.40	99.40±0.20	99.40±0.20
500 × 500	Tr+T	SFC	99.40±0.10	99.50±0.10	99.90±0.00	99.90±0.00
		PFM	98.90±0.10	99.30±0.10	99.90±0.10	99.80±0.00
	Tr+T+R	SFC	99.50±0.10	99.60±0.10	99.90±0.00	99.90±0.10
		PFM	99.10±0.10	99.30±0.10	99.80±0.00	99.80±0.00
500 × 500-OGRN	Tr+T	SFC	96.10±0.30	95.70±0.30	98.70±0.20	98.70±0.20
		PFM	94.70±0.40	95.20±0.40	98.50±0.30	98.30±0.20
	Tr+T+R	SFC	96.30±0.30	96.20±0.30	98.70±0.20	98.70±0.20
		PFM	95.40±0.40	95.60±0.40	98.50±0.30	98.60±0.20

Table 7

Accuracy (expressed as percentages) and standard deviations of the three methods for different datasets and different anatomical section combinations. For GAP and RADAM, we used the ConvNext XLarge backbone and the SFC strategy when using more than one anatomical section. The best results on each dataset are highlighted in bold.

Dataset	Anatomical Section	Method		
		GAP	RADAM	da Silva et al. (2022)
Original	Tr	82.2±1.10	87.3±0.70	56.0±2.00
	T	88.2±0.60	90.5±0.60	42.0±2.00
	R	77.6±0.80	83.6±0.90	42.0±2.00
	Tr + T	91.8±0.70	94.0±0.60	62.0±4.00
	Tr + T + R	93.1±0.70	94.3±0.70	66.0±2.00
1000 × 500	Tr	96.3±0.40	98.3±0.30	71.0±2.00
	T	95.4±0.40	96.7±0.50	71.0±1.00
	R	91.5±0.40	95.3±0.30	52.0±1.00
	Tr + T	99.0±0.30	99.3±0.30	85.0±2.00
	Tr + T + R	99.3±0.20	99.5±0.20	91.0±2.00
500 × 500	Tr	98.7±0.20	99.4±0.10	75.0±2.00
	T	97.5±0.10	98.9±0.10	69.0±1.00
	R	92.9±0.30	96.2±0.10	54.0±1.00
	Tr + T	99.8±0.00	99.9±0.00	86.0±2.00
	Tr + T + R	99.8±0.10	99.9±0.10	95.0±1.00
500 × 500-OGRN	Tr	92.6±0.60	94.8±0.40	38.0±2.00
	T	93.5±0.40	96.4±0.30	34.0±1.00
	R	83.7±0.50	88.7±0.60	27.0±1.00
	Tr + T	97.9±0.20	98.7±0.20	48.0±2.00
	Tr + T + R	98.2±0.20	98.7±0.20	62.0±3.00

In summary, the leave-k-trees-out cross-validation approach has demonstrated the robustness of GAP and RADAM in the context of significant sample variability when characterizing different trees/specimens.

4. Conclusions

In this study, we investigated the efficacy of two feature extraction techniques (GAP and RADAM) used in conjunction with an SVM classifier for the purpose of wood species identification through texture analysis on images from three different anatomical sections. Our results indicate that SVM, when employing either GAP or RADAM, surpasses the results of da Silva et al. (2022), demonstrating their effectiveness in the context of wood species identification through texture analysis techniques. Our research also revealed that combining multiple anatomical sections can significantly enhance the performance.

Among the two methods proposed in this study, RADAM consistently demonstrated superior performance across different datasets, anatomical section combinations and feature fusion strategies. This highlights the robustness of RADAM in handling diverse conditions and reinforces its potential as an effective tool, not only for the recognition of macroscopic textures, but also for wood species identification using microscopic images.

Table 8

Accuracy comparison of GAP, RADAM and the method of da Silva et al. (2022), analyzed with the leave-k-trees-out cross validation strategy. For our methods, we used all three anatomical sections and the SFC strategy.

	Original	1000 × 500	500 × 500	500 × 500-OGRN
da Silva et al. (2022)	30.00	28.00	27.00	22.00
GAP	92.17	91.87	95.03	93.96
RADAM	93.98	94.28	95.03	94.12

Our findings suggest that RADAM can potentially automate wood species identification, reducing the need for specialized anatomical knowledge and enabling more efficient and scalable methods. This is important for the timber industry, where accurate and fast identification of wood species is critical for supply chain monitoring and regulation. However, while being a promising approach, the reliance of the methods on microscopic images introduces challenges in practical applications. In real-world scenarios, capturing high-quality microscopic images at scale may be impractical, especially dealing with large physical timber boards. The main limitation of this study thus lies in obtaining precise microscopic details on the images.

To address this problem, future research should explore the adaptation of the methods studied to macroscopic images, which are easier to capture in industrial environments. Additionally, other factors such as image resolution or the use of other neural network architectures such

as Vision Transformers (ViTs) should be explored.

In summary, the work proposed here represents a significant step forward in the field of wood species identification. Through exploiting the computational efficiency of pre-trained neural network models and advanced feature extraction techniques such as RADAM, we provided a robust and efficient approach to wood species identification, opening doors to more sustainable and responsible forestry practices.

CRedit authorship contribution statement

Kallil M. Zielinski: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Formal analysis, Data curation, Conceptualization. **Leonardo Scabini:** Writing – review & editing, Visualization, Validation, Methodology, Conceptualization. **Lucas C. Ribas:** Writing – review & editing, Validation. **Núbia R. da Silva:** Data curation. **Hans Beeckman:** Writing – review & editing, Validation. **Jan Verwaeren:** Writing – review & editing, Supervision. **Odemir M. Bruno:** Supervision. **Bernard De Baets:** Writing – review & editing, Validation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

K. M. Zielinski acknowledges support from the São Paulo Research Foundation (FAPESP), Brazil (Grant #2022/03668-1) and Higher Education Personnel Improvement Coordination (CAPES), Brazil (Grant #88887.631085/2021-00). L. Scabini acknowledges funding from FAPESP, Brazil (Grants #2021/09163-6 and #2023/10442-2). L. C. Ribas acknowledges support from FAPESP, Brazil (grants #2023/04583-2). O. M. Bruno acknowledges support from CNPq, Brazil (Grant #307897/2018-4) and FAPESP, Brazil (grants #2018/22214-6 and #2021/08325-2). The authors are also grateful to the NVIDIA GPU Grant Program. B. De Baets received funding from the Flemish Government under the “Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.compag.2024.109867>.

Data availability

The authors do not have permission to share data.

References

Biodiversity, 2019. Collections of the royal museum of central africa. URL <https://www.biodiversity.be/>.

Braga, J.W.B., Pastore, T.C.M., Coradin, V.T.R., Camargos, J.A.A., da Silva, A.R., 2011. The use of near Infrared Spectroscopy to Identify solid wood Specimens of *Swietenia Macrophylla*0 (Cites Appendix II). *IAWA J. 32* (2), 285–296.

Cambria, E., Huang, G.-B., Kasun, L.L.C., Zhou, H., Vong, C.M., Lin, J., Yin, J., Cai, Z., Liu, Q., Li, K., Leung, V.C., Feng, L., Ong, Y.-S., Lim, M.-H., Akusok, A., Lendasse, A., Corona, F., Nian, R., Míche, Y., Gastaldo, P., Zunino, R., Decherchi, S., Yang, X., Mao, K., Oh, B.-S., Jeon, J., Toh, K.-A., Teoh, A.B.J., Kim, J., Yu, H., Chen, Y., Liu, J., 2013. Extreme learning machines [trends & controversies]. *IEEE Intell. Syst. 28* (6), 30–59. <http://dx.doi.org/10.1109/MIS.2013.140>.

Cavalin, P.R., Kapp, M.N., Martins, J., Oliveira, L.E.S., 2013. A multiple feature vector framework for forest species recognition. In: *Proceedings of the 28th Annual ACM Symposium on Applied Computing*. ACM, New York, NY, USA, pp. 16–20.

Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn. 20* (3), 273–297.

da Silva, N.R., Deklerck, V., Baetens, J.M., den Bulcke, J.V., Ridder, M.D., Rousseau, M., Bruno, O.M., Beeckman, H., Acker, J.V., Baets, B.D., Verwaeren, J., 2022. Improved wood species identification based on multi-view imagery of the three anatomical planes. *Plant Methods 18* (1).

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: A large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. pp. 248–255.

Dormont, E.E., Boner, M., Braun, B., Breulmann, G., Degen, B., Espinoza, E., Gardner, S., Guillery, P., Hermanson, J.C., Koch, G., Lee, S.L., Kanashiro, M., Rimbawanto, A., Thomas, D., Wiedenhoeft, A.C., Yin, Y., Zahnen, J., Lowe, A.J., 2015. Forensic timber identification: It's time to integrate disciplines to combat illegal logging. *Biol. Cons. 191*, 790–798.

Filho, P.L.P., Oliveira, L.S., Nisgoski, S., Britto, A.S., 2014. Forest species recognition using macroscopic images. *Mach. Vis. Appl. 25* (4), 1019–1031.

Gasson, P., 2011. How precise can wood identification be? Wood anatomy's role in support of the legal timber trade, especially cites. *IAWA J. 32* (2), 137–154.

Guang-Sheng, C., Peng, Z., 2013. Wood cell recognition using geodesic active contour and principal component analysis. *Optik 124* (10), 949–952.

Hanssen, F., Wischniewski, N., Moreth, U., Magel, E.A., 2011. Molecular identification of *fitzroya cupressoides*, *sequoia sempervirens*, and *Thuja plicata* wood using taxon-specific rDNA-ITS primers. *IAWA J. 32* (2), 273–284.

Hassold, S., Lowry, P.P., Bauert, M.R., Razafintsalama, A., Ramamonjisoa, L., Widmer, A., 2016. DNA barcoding of Malagasy rosewoods: Towards a molecular identification of CITES-listed *Dalbergia* species. *PLOS One 11* (6), e0157881.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. CVPR, pp. 770–778.

Herrera-Poyatos, D., Herrera-Poyatos, A., Montes, R., de Palacios, P., Esteban, L.G., Iruela, A.G., Fernández, F.G., Herrera, F., 2024. Deep learning methodology for the identification of wood species using high-resolution macroscopic images. *arXiv: 2406.11772*.

ITTO, 2020. Biennial Review and Assessment of the World Timber Situation. Yokoyama, Japan. URL <https://www.itto.int/>.

Jiao, L., Yu, M., Wiedenhoeft, A.C., He, T., Li, J., Liu, B., Jiang, X., Yin, Y., 2018. DNA barcode authentication and library development for the wood of six commercial *pterocarpus* species: the critical role of xylarium specimens. *Sci. Rep. 8* (1), 1945.

Kasun, L., Zhou, H., Huang, G.-B., Vong, C.-M., 2013. Representational learning with ELMs for big data. *IEEE Intell. Syst. 28*, 31–34.

Khalid, M., Yusof, R., Mohd Khairuddin, A.S., 2011. Tropical wood species recognition system based on multi-feature extractors and classifiers. In: *2011 2nd International Conference on Instrumentation Control and Automation*. IEEE, pp. 6–11.

Knuth, D.E., 1997. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*, vol. 2, third ed. Addison-Wesley Longman Publishing Co., Inc.

Lens, F., Liang, C., Guo, Y., Tang, X., Jahanbanifard, M., da Silva, F.S.C., Ceccantini, G., Verbeek, F.J., 2020. Computer-assisted timber identification based on features extracted from microscopic wood sections. *IAWA J. 41* (4), 660–680.

Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., Xie, S., 2022. A convnet for the 2020s. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 11976–11986.

Martins, J., Oliveira, L.S., Nisgoski, S., Sabourin, R., 2013. A database for automatic classification of forest species. *Mach. Vis. Appl. 24* (3), 567–578.

Martins, J.G., Oliveira, L.S., Sabourin, R., 2012. Combining textural descriptors for forest species recognition. In: *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*. IEEE, pp. 1483–1488.

May, C., 2017. *Global Financial Integrity, 2017. Transnational Crime and the Developing World*. Global Financial Integrity.

Nieradzki, L., Sieburg-Rockel, J., Helmling, S., Keuper, J., Weibel, T., Olbrich, A., Stephani, H., 2024. Automating wood species detection and classification in microscopic images of fibrous materials with deep learning. *Microsc. Microanal. 30* (3), 508–520. <http://dx.doi.org/10.1093/mam/ozae038>, URL <http://dx.doi.org/10.1093/mam/ozae038>.

Nithaniyal, S., Newmaster, S.G., Ragupathy, S., Krishnamoorthy, D., Vassou, S.L., Parani, M., 2014. DNA Barcode Authentication of Wood Samples of Threatened and Commercial Timber Trees within the Tropical Dry Evergreen Forest of India. *PLoS ONE 9* (9), e107669.

Ojansivu, V., Heikkilä, J., 2008. Blur Insensitive Texture Classification Using Local Phase Quantization. pp. 236–243.

Pastore, T.C.M., Braga, J.W.B., Coradin, V.T.R., Magalhães, W.L.E., Okino, E.Y.A., Camargos, J.A.A., de Muñiz, G.I.B., Bressan, O.A., Davrieux, F., 2011. Near infrared spectroscopy (NIRS) as a potential tool for monitoring trade of similar woods: Discrimination of true mahogany, cedar, andiroba, and curupixá. *Holzforchung 65* (1), 73–80.

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., Chintala, S., 2019. PyTorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems 32 (NeurIPS)*. pp. 8024–8035, URL <http://papers.neurips.cc>.

- 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., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830, URL <http://www.jmlr.org>.
- Price, E.R., Miles-Bunch, I.A., Gasson, P.E., Lancaster, C.A., 2021. Pterocarpus wood identification by independent and complementary analysis of DART-TOFMS, microscopic anatomy, and fluorescence spectrometry. *IAWA J.* 42 (4), 397–418.
- Ravindran, P., Costa, A., Soares, R., Wiedenhoeft, A.C., 2018. Classification of CITES-listed and other neotropical Meliaceae wood images using convolutional neural networks. *Plant Methods* 14 (1), 25.
- Ravindran, P., Thompson, B.J., Soares, R.K., Wiedenhoeft, A.C., 2020. The XyloTron: Flexible, open-source, image-based macroscopic field identification of wood products. *Front. Plant Sci.* 11.
- Scabini, L., Zielinski, K.M., Ribas, L.C., Gonçalves, W.N., De Baets, B., Bruno, O.M., 2023. RADAM: Texture recognition through randomized aggregated encoding of deep activation maps. *Pattern Recognit.* 143, 109802.
- da Silva, N.R., De Ridder, M., Baetens, J.M., Van den Bulcke, J., Rousseau, M., Martinez Bruno, O., Beeckman, H., Van Acker, J., De Baets, B., 2017. Automated classification of wood transverse cross-section micro-imagery from 77 commercial Central-African timber species. *Ann. For. Sci.* 74 (2), 30.
- Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J.N., Wu, Z., Ding, X., 2020. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. *Med. Image Anal.* 63, 101693.
- United Nations, 2024. Convention on international trade in endangered species of wild fauna and flora: Appendices I, II, and III. URL <https://cites.org/sites/default/files/eng/disc/CITES-Convention-EN.pdf> (Accessed: 1 October 2024).
- Wang, H.-j., Qi, H.-n., Wang, X.-F., 2013. A new Gabor based approach for wood recognition. *Neurocomputing* 116, 192–200.
- Wheeler, E., Baas, P., Gasson, P., 1989. IAWA List of Microscopic Features for Hardwood Identification. *IAWA J. / Int. Assoc. Wood Anat.* 10, 219–332.
- Wightman, R., 2019. PyTorch Image Models (timm). <https://github.com/rwightman/pytorch-image-models>.
- Wu, F., Gazo, R., Haviarova, E., Benes, B., 2021. Wood identification based on longitudinal section images by using deep learning. *Wood Sci. Technol.* 55 (2), 553–563.
- Zhao, P., Dou, G., Chen, G.-S., 2014. Wood species identification using feature-level fusion scheme. *Optik* 125 (3), 1144–1148.
- Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., Gao, R.X., 2019. Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* 115, 213–237.
- Zheng, C., Liu, S., Wang, J., Lu, Y., Ma, L., Jiao, L., Guo, J., Yin, Y., He, T., 2024. Opening the black box: explainable deep-learning classification of wood microscopic image of endangered tree species. *Plant Methods* 20 (1).