

# Technical framework of advanced volumetric sub- $\mu$ -CT wood imaging integrated with adaptable deep-learning-based wood species classification: initial evaluation on softwood and hardwood data

Jannik STEBANI <sup>1,2,3,\*</sup>, Tim LEWANDROWSKI <sup>4</sup>, Kilian DREMEL <sup>1</sup>, Simon ZABLER <sup>1,5</sup> and Volker HAAG <sup>4</sup>

<sup>1</sup> Fraunhofer IIS Development Center X-ray Technology EZRT, Am Hubland, 97074 Würzburg, Germany

<sup>2</sup> Lehrstuhl für Röntgenmikroskopie, Faculty of Physics, University of Würzburg, Josef-Martin-Weg 63, 97074 Würzburg, Germany

<sup>3</sup> Experimental Physics 5, Faculty of Physics, University of Würzburg, Am Hubland, 97074 Germany

<sup>4</sup> Thünen Institute of Wood Research, Leuschnerstraße 91d, G-21031 Hamburg, Germany

<sup>5</sup> Faculty of Computer Science, Deggendorf Institute of Technology (THD), Dieter-Görlitz-Platz 1, 94469 Deggendorf, Germany

\*Corresponding author; email: [jannik.stebani@iis.fraunhofer.de](mailto:jannik.stebani@iis.fraunhofer.de)

ORCID iDs: Stebani: 0009-0004-9631-9928; Lewandrowski: 0000-0002-2376-3436; Dremel: 0009-0007-6359-8758; Zabler: 0000-0003-1632-6673; Haag: 0000-0002-5913-6485

Accepted for publication: 26 June 2025; published online: 5 August 2025

**Summary** – In an age of increasing surveillance and control of markets and trade flows, new methods for species identification of commercial timbers and other woody plant products are needed. Traditional plant anatomy has generated a vast body of knowledge on taxonomical identification. In highly processed and mixed materials such as charcoal, fibreboard board and paper, where DNA and chemotaxonomic composition is regularly destroyed or altered, the anatomical structure is often still intact. Comprehensive studies in the field of state-of-the-art  $\mu$ CT analysis have shown that the use of sub-micrometre CT scanning technology can be very suitable for wood identification and other structural investigations. The current state of CT technology (resolution, working speed, feasibility) opens up various options for use in combination with other relevant key technologies, in particular machine learning and other forms of digitisation in the context of wood research. Thus, the primary objective of this methodologically oriented study is to design and implement, for the first time, an automated and modular classification pipeline that integrates machine learning with volumetric high-resolution sub- $\mu$ CT imaging for wood identification. To evaluate this approach, we conducted initial experiments using a simplified basal case, applying the method to sub- $\mu$ CT scans of softwood (*Pinus sylvestris*) and hardwood (*Acer pseudoplatanus*).

**Keywords** – CT wood imaging, EUDR, machine learning, neural network, timber trade, wood classification, wood anatomy.

## Introduction

The latest developments and research in the field of digital image processing and analysis, and machine learning in the form of deep learning, open up new research opportunities and applications for wood identification based on structural wood anatomy. Traditional wood anatomy, against other methods like genetic testing, isotope analysis or mass spectroscopy, has retained a major position in the field of wood identification (Beeckman *et al.* 2020). It is

a rapid, cost-effective and widely applied method and will continue to be used alongside other wood identification methods. Wood anatomy is highly significant for the strict implementation of the laws governing the timber trade in the context of CITES, EUDR and FLEGT (Koch & Haag 2015; Helmling *et al.* 2016, 2018; Haag *et al.* 2018, 2020; Zahnen *et al.* 2020; Lewandrowski *et al.* 2022), the enforcement of which is impeded by insufficient capacities for forensic timber identification (Dormontt *et al.* 2015). In some cases, it is the only viable option when other methods are not applicable due to process-related changes in the wood (chemical modification, destruction of DNA). Various teams worldwide are currently engaged in transferring traditional wood anatomy into the digital age in new fields of application. The main focus here is currently the automation of wood identification by interpreting it as a computer vision problem and solving it algorithmically. Mostly, the algorithm class of deep neural networks is considered to solve the problem due to their recent breakthrough performance (Krizhevsky *et al.* 2012; He *et al.* 2015; Bello *et al.* 2021; Dosovitskiy *et al.* 2021; Sarker 2021) on many tasks, including wood science (Lens *et al.* 2020; Hwang & Sugiyama 2021; Silva *et al.* 2022). Rapid and cost-efficient automated systems are expected to complement traditional means of wood identification methods. The first machine vision-based wood identification systems were concerned primarily with the analysis of macroscopic transverse sections of commercial timbers. Subsequent development of these techniques also enables the usage of micrograph data from well-established microscopy. This allows a more detailed differentiation of cellular features, although thus far limited to those found in the cross-sections (Hermanson & Wiedenhoef, 2011; Ravindran *et al.*, 2018; He *et al.*, 2020; Hwang & Sugiyama, 2021; Koch *et al.*, 2021). In general, the advances in computing hardware make the inclusion of data in computer vision problems with increasing resolution and dimensionality feasible.

Parallel to the development in digitisation and the use of artificial intelligence in the field of wood research, there has also been substantial technical progress in computed tomography (CT). This technique, as a nondestructive imaging methodology (Withers *et al.* 2021), is well-suited to generate data for subsequent computer vision-based processing with the aim of automated wood identification. For over a decade, CT has been successfully used both macroscopically and microscopically in the field of wood identification and dendrochronology (Bird *et al.* 2008; Grabner *et al.* 2009; Van Den Bulcke *et al.* 2009; Haneca *et al.* 2012; Hubau *et al.* 2013; Haag *et al.* 2018). Recent studies in the micro- ( $\mu$ CT) and nano-resolution range on the sub-cellular level are also increasingly being reported (Hwang *et al.* 2020; Koddenberg *et al.* 2021; Patera *et al.* 2021; Duncan *et al.* 2022). The investigations show that  $\mu$ CT-based analyses have long since attained the status of an established state-of-the-art technology in the field of wood identification and are regularly applied (Haag *et al.* 2022). Although there are ample reports about ongoing wood scientific research with either CT or computer vision in the literature, the combination of volumetric state-of-the-art  $\mu$ CT with computer vision in the form of deep neural networks is not widely reported. Current approaches use either the CT data acquisition method in isolation (Kobayashi *et al.* 2019; Gejdoš *et al.* 2023; Stelzner *et al.* 2023; Dierickx *et al.* 2024) for wood scientific purposes, focus on macroscopic planar slices (Vacek *et al.* 2024) or use deep learning methods on macroscopic or microscopic planar light images of wood samples (He *et al.* 2020; Fabijańska *et al.* 2021; Kirbaş & Çifci 2022; Silva *et al.* 2022; Bello 2023) without the usage of volumetric CT.

Thus, volumetric sub- $\mu$ CT in conjunction with automated computer vision-based wood identification using deep neural networks remains an underreported combination approach with considerable benefits for wood science and downstream timber industrial applications. Thus, the present study aims to methodologically describe a modular end-to-end pipeline encompassing machine learning in the form of deep neural networks (DNN) and volumetric high-resolution sub- $\mu$ CT technology for the field of wood identification, and to carry out preliminary evaluations with specifically recorded softwood and hardwood data. Based on the technical developments in the CT space and the machine learning space, we expect large benefits via precise and rapid wood identification by combining said technologies. Distinct softwood (*Pinus sylvestris*) and hardwood (*Acer pseudoplatanus*) species were selected as the basis for the proof-of-concept study. Thus, the following three points can be highlighted as the key contributions of this study.

First, we described and evaluated a technical framework encompassing an automated pipeline for processing and classifying volumetric, high-resolution  $\mu$ CT scans of wood samples. The deep learning core model was based on a

modular residual network. Further, we compared derived dimensionality-wise formulations of said pipeline that match traditional planar 2D (micrograph-like) and pseudo-volumetric triaxial 2.5D (multi-cross-section-like) to the full volumetric 3D formulation.

We further tested the pipeline's robustness end-to-end with six informative data transformation protocols. Each protocol entailed the computation of *in-silico* image corruptions with increasing magnitude to gain insights about a possible failure mode regime with respect to input data alterations. Knowledge about the breakdown of the deep learning classifier due to noise and imaging defects is paramount for possible downstream real-world applications.

Lastly, we provide the purpose-built Python and PyTorch-based implementation of the model, pipeline and evaluation tooling as a documented open-source package. This enables the reproduction and verification of the experiments detailed in this work. The software package may expedite the realisation of subsequent machine intelligence-based wood classification projects. Main merits of the code infrastructure are an adaptable structure with built-in facilities for 2D, 2.5D and 3D data, as well as the human-readable text-file and command-line interface-based configuration of all training, inference, and robustness experiments. This may help non-deep learning experts in performing machine vision experiments using local in-house data with hyperparameter exploration and comprehensive robustness testing, since machine learning and deep learning pipelines are powerful tools that can be difficult to set up and configure (Talaei Khoei *et al.* 2023).

In summary, this study was conducted with the specific objective of enhancing the potential of  $\mu$ CT technology and deep neural network programming in the field of (forensic) wood identification.

## Materials and methods

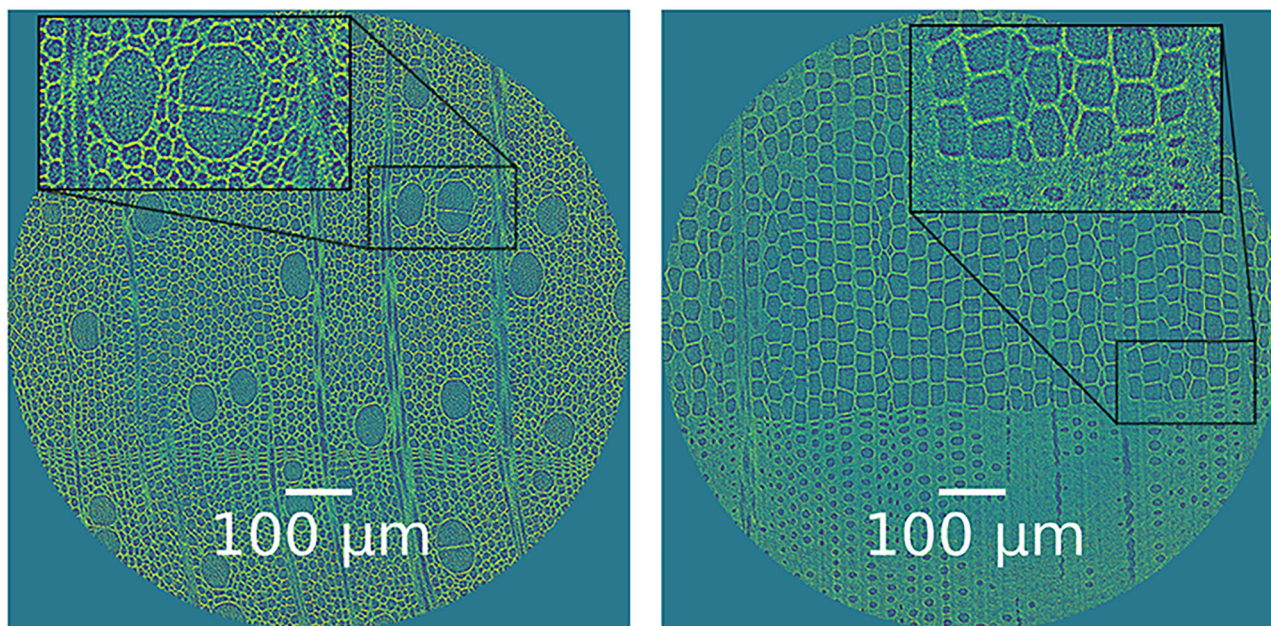
### MATERIAL, SAMPLE AND $\mu$ CT DATASET DESCRIPTION

For the present study, sycamore maple (*Acer pseudoplatanus*) and Scots pine (*Pinus sylvestris*) were selected. The criteria for selecting these timbers were to have timbers with a homogeneous, yet diverse cellular structure while possessing characteristic features for reliable identification at the generic level. The distinct structure of the wood species was deliberately chosen to describe and implement a clear baseline for this conceptual proof-of-principle investigation. The samples of *Pinus sylvestris* and *Acer pseudoplatanus* were collected from vouchered reference material from a single tree instance of the respective species of the scientific wood collection (RBHw) of the Thünen Institute of Wood Research in Hamburg. The X-ray images in this study were taken with the Click-CT (Zabler *et al.* 2020), a sub-micrometre CT scanner developed at the Fraunhofer Development Center for X-ray Technology. The Click-CT delivers high image quality and resolution, particularly for organic structures, as it operates at low X-ray energies and is "detector-based", meaning it achieves a very high resolution due to small detector pixels. The detector consists of an optical microscope that photographs a scintillator plate, which converts the X-ray photons into visible light. The photons are generated by a micro-focal X-ray source. The sample is positioned close to the detector screen and far away from the focal spot to minimize blur, as the effect of the focal spot on resolution is determined by its projected size along the ray path through the sample. The samples were manually cut and split into rectangular bars with a width of approximately 3 mm and a length of 1 cm. The bars were attached upright to a small sample holder with hot glue. Similar to light microscopy, the possible image resolution is determined by the choice of the objective lens. Click-CT offers the possibility to choose between three resolution levels (4 $\times$ , 10 $\times$ , 20 $\times$ ), which lead to different pixel sizes on the detector.

The measurements in this study were performed at 10 $\times$  magnification with a detector pixel size of 650 nm and, due to the geometric magnification, a resulting raw voxel size of 482 nm. Two different settings were used for the scan duration. The exposure time for two of the data sets used (one *Acer pseudoplatanus*, one *Pinus sylvestris*) was 10 hours and comprised 1800 double-averaged projections, each with an exposure time of 10 seconds. For the remaining data sets, the exposure time was reduced to 2 hours. Here, 1500 projections were recorded with an exposure time of 5 s.

The X-ray anode voltage was 40 kV, and the emission current was 200  $\mu$ A for all scans. The scans were performed as region-of-interest (ROI) scans covering a raw volume of  $2300 \times 2300 \times 1900$  volume picture elements (voxels) with a physical size of 1.1 mm  $\times$  1.1 mm  $\times$  0.9 mm inside the bars.

The balanced data set consisted of 20 separate samples with 10 instances belonging to the *Pinus sylvestris* class and 10 instances belonging to the *Acer pseudoplatanus* class. The sample mounting inside the scanner device resulted in differing intra-class orientations. For the *Pinus sylvestris* class, six instances belonged to the transversal orientation group and four instances belonged to the axial-like orientation group. For the *Acer pseudoplatanus* class, six instances appertained to the transversal orientation group and four instances appertained to the axial-transversal group. We explicitly recorded the volumetric datasets with the specified orientations to obtain raw data with canonical cross-section orientations, in turn enabling us to conduct group-wise skill transfer experiments. For this, we refer to the section on dataset splitting in the context of cross-validation and the results concerning the model skill transfer. In Figure 1, representative image slices of both classes from the recorded Click-CT image data are displayed. The colour-coding in the figure is based on the basal CT image contrast. This method involves irradiating the sample with X-rays, which pass through it and are captured by a planar detector on the other side. Using a rotating mechanism, multiple X-ray irradiation measurements are taken. With the help of a tomographic reconstruction algorithm (Willeminck & Noël 2019), cross-sectional slice images can be computed from the raw data. Inside the object, the X-rays pass through structures, tissues or materials with different characteristics. These characteristics influence the X-ray attenuation coefficient  $\mu_{\text{material}}$ . In the samples utilised in this study, structures may correspond to typical wood microstructure elements like cell walls, resin canals and vessels. Using the computed tomography technique, the scanner-adjacent computer reconstructs a volumetric spatial map of voxel intensities  $I(x, y, z)$  that in turn are a function  $h$  of the attenuation coefficient  $\mu$ :  $I(x, y, z) = h(\mu(x, y, z))$ . In planar or volumetric visualisations of the CT image data, it is this mapping of attenuation coefficient data that is encoded in the pixel- or voxel-wise image intensity values (Withers *et al.* 2021). In typical visualisation schemes, low attenuation values like air appear darker, while high attenuation values from dense regions or materials appear brighter. However, many mappings of intensity values to colour and brightness are possible. In Figure 1, lighter yellow pixels encode a higher attenuation while darker violet pixels encode a lower attenuation. The visual appearance, i.e. the relative contrast of a slice image or volumetric rendering can be adjusted by a process called windowing. Here, the range of values displayed is adjusted to improve the visibility of desired structures. The window width sets the range of attenuation values that are displayed between the extreme edge values. The window level setting centres the colour-scale mapping around the indicated value. The joint tuning of these parameters enables the highlighting of specific structures or materials. Interactively modifiable examples thereof can be retrieved in the supplementary showcase repository. Visualisations of the voxel intensity value distributions of the recorded datasets can be inspected in Figure 1 in the supplementary material. Another visualisation of a line profile of  $I(x, y = \tilde{y}, z = \tilde{z})$  for a sample dataset region-of-interest is provided in supplementary Figure 2. Moreover, the CT image data can be compared to stained sections of the two wood species in Figure 2. From the comparison of Figure 1 and Figure 2 it is evident that substantial wood species anatomy (rays, vessels and further cell structures) is well-represented in the CT image data. The main differences between the selected timbers are that sycamore maple is a hardwood with a more complex structure due to the larger number of different cell types (vessels and different axial parenchyma types), while Scots pine is a softwood consisting of up to 95 % tracheids. Sycamore maple (Figure 2A,B) is diffuse-porous with some vessel multiples, commonly in short (2–3 vessels) radial rows. The vessels are fairly small (diameter 44–65–80  $\mu$ m; Richter & Dallwitz 2015) and relatively evenly distributed within a growth ring. In addition to the presence (sycamore maple) vs. absence (Scots pine) of vessels, the two timbers can be reliably differentiated based on the nature of the rays alone. Contrary to the low and mostly uniseriate rays of pine, most rays of maple are multiseriate, and the large rays are usually over 500  $\mu$ m (up to 1000  $\mu$ m) high. As a special characteristic of maple, the vessels feature helical thickenings throughout the body of both narrow and wide elements (Fig. 2B).

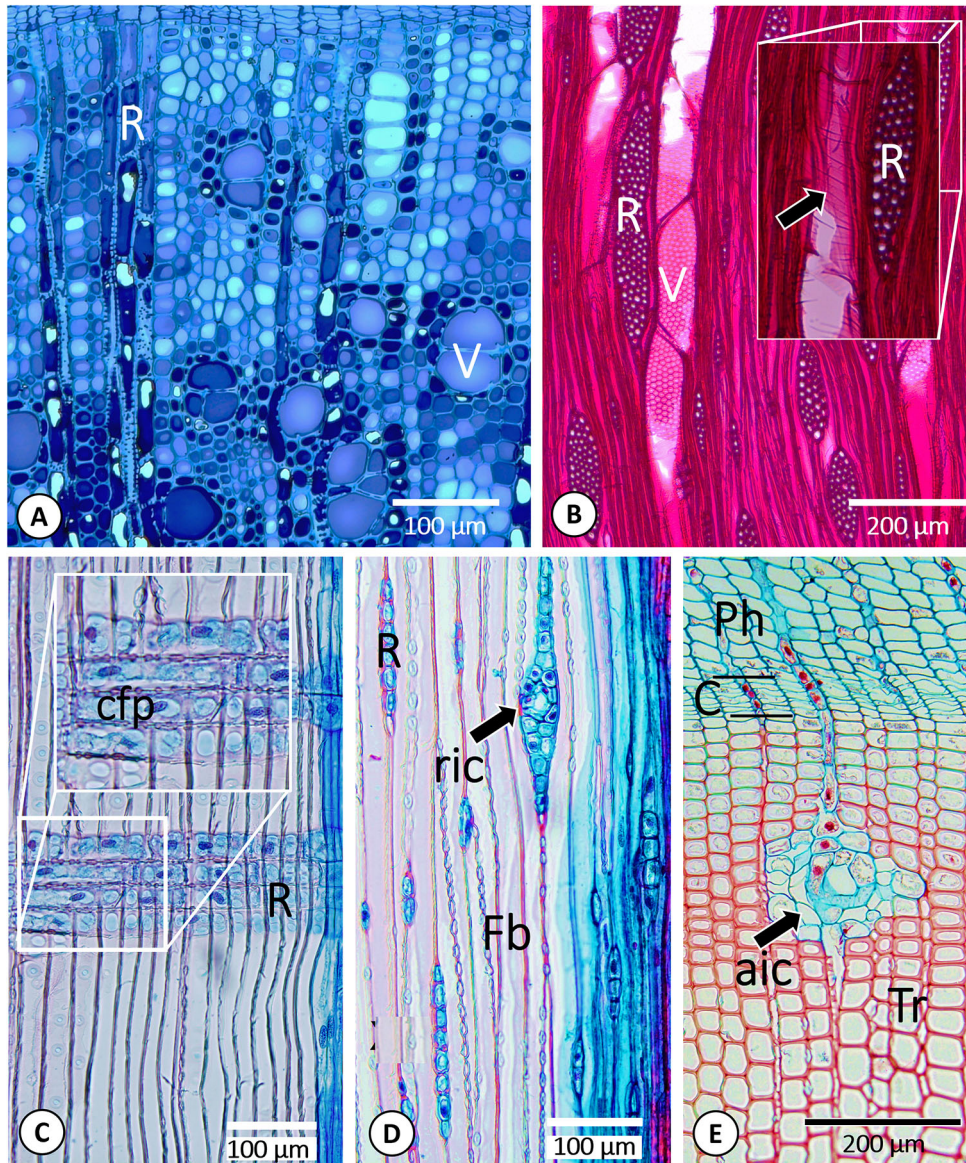


**Fig. 1.** Region-of-interest sub- $\mu$ -CT based illustration of the basal wood species *Acer pseudoplatanus* (left) and *Pinus sylvestris* (right) used for the experiments. Scale bars = 100  $\mu$ m. Inset axes show magnified subregions for better detail assessment.

Special cellular features of Scots pine are the presence of intercellular (resin) canals in radial (Fig. 2D) and axial directions (Fig. 2E), and the characteristic window-like cross-field pitting between ray cells and tracheids (Fig. 2C). The latter are typical for the genus *Pinus* (subsection *Pinus*) and are useful to distinguish species of this subsection from those of other subsections and also from other genera within the family Pinaceae.

#### SOFTWARE FOR DEEP LEARNING MODEL AND DATA PIPELINE

For the classification of the wood samples, we implemented a purpose-built data processing pipeline and deep learning package as open-source software with Python, using PyTorch (Paszke *et al.* 2019) with other libraries of the scientific Python ecosystem (Pedregosa *et al.* 2011; Harris *et al.* 2020; Cardoso *et al.* 2022). At its core, the package provides functionality for automated loading and processing of volumetric 3D and planar scan data as well as infrastructure code for the training and extensive test-time evaluation of wood species classification models. For reproducibility and as an added benefit for the community to use and extend, we release the code as an installable, documented Python package in the openly accessible online repository woodnet as part of the study. We intend to support the project concerning possible feature requests from the community. The adaptable structure of the code allows potential users to load local in-house data and perform experiments using the tiling processing module with full FOV or smaller tiles. The package architecture is not restricted to the processing of volumetric CT scans — data originating from macroscopic images, micrographs, multiple cross-sections and other single- or multi-channel sources can be handled in principle. The package may be used as a library with the publicly hosted application programming interface (API) documentation or as a stand-alone application. Training, inference, and robustness experiments can be configured deeply via human-readable text-based configuration files and conducted with a command-line interface. Also, the built-in tooling for cross-validation, model ensemble recording during training experiments and robustness evaluation with subsequent aggregation of sizable result data is a potential use of the package. Especially, model ensembles and robustness testing in the context of uncertainty quantification in machine learning (Hüllermeier & Waegeman 2021) are further applications supported by the package. With the triad of open-



**Fig. 2.** *Acer pseudoplatanus*: A. Transverse section with rays (R) and Vessels (V) stained with toluidine blue; B. Tangential section with rays (R), Vessels (V) and vessel with helical thickenings (arrow in blow up) stained with safranin; *Pinus sylvestris*: C. Radial section with rays (R) and pinoid window-like cross-field pitting (cfp) in box stained with safranin and astra blue; D. Tangential section with uniseriate rays (R) and ray with radial intercellular (resin) canal (ric) stained with safranin and astra blue; E. Transverse section with phloem (Ph), cambium (C), tracheid ground tissue (Tr) and axial intercellular canal (aic) stained with safranin and astra blue. Scale bars = 100 µm for A, C, D and 200 µm for B and E.

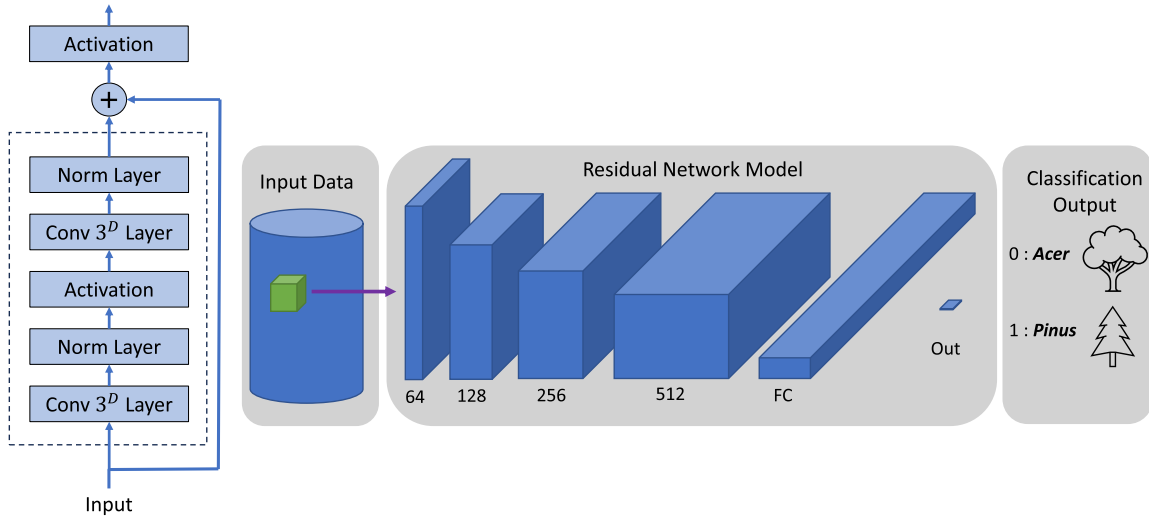
source package, documentation and conceptual showcase applications demonstrated in the repository woodnet-showcase with its supplemental data (Stebani *et al.* 2024), we hope to achieve two outcomes. Firstly, the establishment of advanced sub- $\mu$ -CT imaging as a method in machine learning-supported wood forensics and wood anatomical studies. And secondly, to encourage more research groups to experiment with machine vision and thus embark on the mission to generate a large-scale consensus data corpus.

## DEEP NEURAL NETWORK DESCRIPTION

The primary aim of this proof-of-concept study with methodological emphasis was the implementation of a purpose-built deep learning classification pipeline and corresponding evaluation with a conceptual binary classification problem consisting of sub- $\mu$ CT scans of hardwood and softwood samples. For this, we implemented the modular ResNet architecture (RESidual NETwork) (He *et al.* 2016) as the deep learning backbone. ResNets belong to the larger class of convolutional neural networks (CNN) that were specifically designed for computer vision problems, such as image or CT scan classification. CNN are particularly effective for vision-based classification due to their ability to leverage spatial hierarchies through the convolution operation. The convolution operation, as the core building block of a CNN, applies a set of learnable filters (kernels) to the input data. The filters slide (convolve) across the input data and compute feature maps consisting of salient patterns like edges, textures, other geometry in earlier layers, and higher-level concepts in later layers. The feature maps result from the structured multiplication and summation of the learnable filter weights and the local regions of the input. The CNN itself can be written as a function composition, mapping from input images or volumes to class labels. Their strength in vision-related tasks lies in the ability to adaptively learn salient patterns from input data, making them a hierarchical feature detector and classifier. As demonstrated by equation (1), a generic  $l$ -layer neural network can be described as a composite function consisting of convolutional or fully connected blocks with learnable weight parameters (functions  $a^l$ ) and activation functions  $\sigma^l$  operating on the input pattern  $x$ .

$$\text{NN}(x) = \sigma^l \circ a^l \circ \sigma^{l-1} \circ a^{l-1} \circ \dots \circ \sigma^2 \circ a^2 \circ \sigma^1 \circ a^1(x) \quad (1)$$

For the ResNet, further elements like pooling, residual connections and normalisation layers enter the architecture. Using the ResNet-CNN backbone, we compared the implementations of the dimension-differing 2D planar, 2.5D triaxial and 3D volumetric approaches with respect to the classification performance and robustness against input data corruption. For the 2D formulation, a basic single-channel variant operating on an image slice was used. For the 2.5D triaxial formulation, a 2D-analogous variant with three input channels was used. Orthogonal planes were concatenated along the third channel-wise dimension for this variant. A fully volumetric variant with 3D convolutions was employed for the 3D task. A sketch of a single residual block module of the network (left) and a full network overview (right) is provided in Fig. 3. The first convolution operation utilised a larger ( $7 \times 7(\times 7)$ ) convolution kernel with an isotropic stride of 2. The stride denoted the step size with which the convolution kernel traverses the input data. The feature map output of this first layer was fixed to 64. A dimensionality-reducing maximum pooling layer was placed between the first convolutional layer and the other model blocks. The pooling operation reduced the computational load and the number of parameters and makes the model invariant to small translations of the input. Further model blocks consisted of the twice-repeated application of the convolution operation, normalisation operation and nonlinear activation. This twice-repeated CONV  $\rightarrow$  NORM  $\rightarrow$  NONLIN section was accompanied by the ResNet-typical bypassing skip connection (Fig. 3, left). The kernel size of the intra-block convolutions was fixed to ( $3 \times 3$ ) for the 2D and 2.5D formulation and ( $3 \times 3 \times 3$ ) for the 3D case. The convolutional section of the models operated with a fixed cascade of increasing feature map counts defined by the sequence (64, 128, 256, 512). The normalisation operation was performed by batch normalisation layers (Ioffe & Szegedy 2015). The rectified linear unit (ReLU) (Fukushima 1969) was chosen as the nonlinear activation function inside the network (symbolized by  $\sigma$  in equation (1)). The convolutional part of the network was followed by a fully connected (FC) classification layer of the dimensionality (512,  $N_c$ ). Here,  $N_c$  denotes the number of output classes. For the binary classification of the wood species *Pinus* and *Acer*,  $N_c = 1$  was used. Compatibility between the final convolutional layer and the FC layer irrespective of the input data spatial size was ensured by an adaptive average pooling layer. The raw outputs of the



**Fig. 3.** Sketch of the deep neural network architecture with the principal component (residual network block) on the left side and full classifier network with the convolutional encode cascade and the fully connected classification head (FC) on the right. For the planar and triaxial experiments, convolution operations with dimensionality  $D = 2$  were used, while the volumetric experiments used  $D = 3$ .

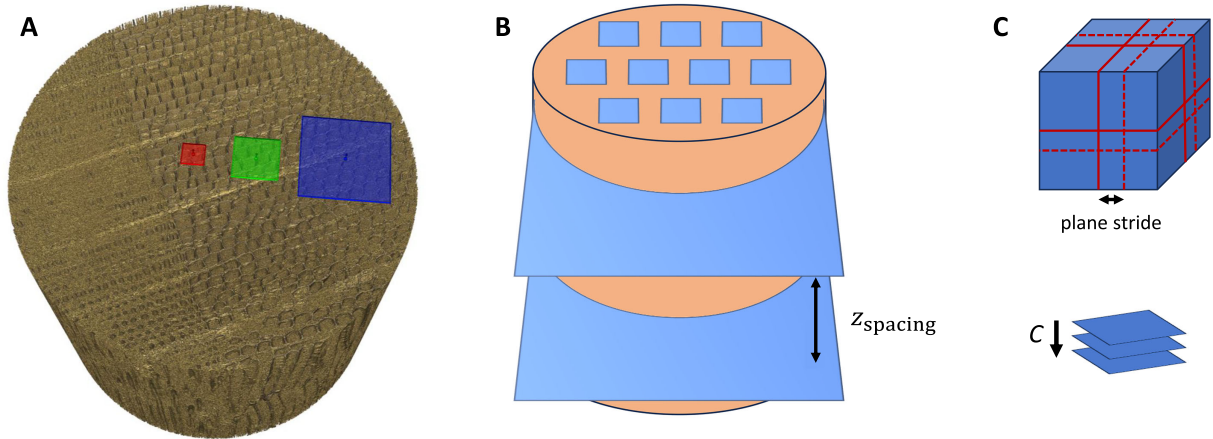
FC layer were processed with the final nonlinear function. For the binary *Acer-Pinus* classification task, the logistic function described in equation (2) was used.

$$f(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

All the architectural choices, functions and parameters listed above can be extracted and possibly adapted using the model definition files in the open-source Python package released as part of this work.

#### DATA PREPROCESSING

The basal raw data was down-sampled to double the voxel edge length along all axes, effectively yielding datasets of dimensionality ( $1150 \times 1150 \times 900$ ) voxels with voxel edge lengths of 964 nm. This approach is equivalent to software binning (Friedrichsdorf *et al.* 2019) and was primarily implemented to alleviate the computational and memory demands associated with large volumetric data inputs. We observed that the software-binned data retained sufficient microstructure details with increased global contextual details related to wood anatomical structures. Additionally, the down-sampling process mitigates the intrinsic noise present in  $\mu$ CT images and improves the signal-to-noise ratio (Irie *et al.* 2022). The experiments were conducted using patches or chunks extracted from the dataset instances. The deep learning model operated on square planar image patches (2D case, size  $s \times s$ ), three square orthogonal patches concatenated along the channel dimension (2.5D case, size  $3 \times s \times s$ ) from a chunk or directly on cuboid subvolume chunks (3D case, size  $s \times s \times s$ ) for training and prediction purposes. Every formulation was also trained and evaluated with varying edge length of the cuboid patches/chunks. We experimented with the sizes  $s \in \{64, 128, 256\}$  voxels. An illustrative rendering of the patches or sub-volume chunk faces in relation to the full scan volume is given in Fig. 4A. For the 2D case, data elements (image patches) were sampled from equidistant  $z$  planes separated by the user-set  $z_{spacing}$  parameter along the symmetry axis. In plane, the patches were generated according to a pre-computed pattern that maximized the patch count yield per  $z$  plane (Fig.4B). Subvolume chunks were created from the in-plane patch patterns by extension along the  $z$  axis. For the 2.5D case, data elements (three orthogonal planes) were sampled from subvolume chunks. The sampling density inside a chunk was adjusted via a plane stride parameter



**Fig. 4.** Illustration of the patch or subvolume extraction scheme. A. Rendering of 64 (red), 128 (green) and 256 (blue) voxel patch faces in relation to the full *cylindrical* CT scan volume. B. Display of the in-plane patch sampling and equidistant planes for the 2D approach. C. Display of the orthogonal planes sampling from a subvolume chunk for the triaxial 2.5D approach. Multiple combinations of orthogonal planes separated by the plane stride parameter were extracted and concatenated along the new channel dimension prior to insertion into the model.

(Fig. 4C). Concatenation of the three patches along the channel dimension then produced a single 2.5D input data element. For the 3D case, simply full subvolume chunks were extracted according to the pattern. Depending on actual values for  $z_{\text{spacing}}$  and plane stride parameters, this extraction pattern yielded total dataset element counts of 720 (3D formulation with largest  $s$ ) to over  $3 \times 10^5$  (2D formulation with minimal  $s$ ). We adopted global voxel intensity normalization as the primary static data preprocessing scheme. For this, the voxel-wise radiodensity measurements were rescaled according to equation (3) to achieve an approximate Gaussian distribution with zero mean and unit standard deviation.

$$\tilde{v}_i = \frac{v_i - \langle v \rangle}{\sigma(v)} \quad (3)$$

In the above equation, the voxel value  $v_i$  is rescaled by subtraction of the mean voxel value  $\langle v \rangle$  and division by the overall voxel value standard deviation  $\sigma(v)$ . The values for  $\langle v \rangle$  and  $\sigma(v)$  were computed from the respective training dataset.

## Dataset splitting for cross-validation

Dataset splitting into training and validation data is a necessary preliminary step in a deep learning experiment. As detailed in the following paragraph about the training procedure, training data is utilised to fit the deep learning model, while the validation data is used to evaluate the model. The cross-validation (CV) technique was employed to obtain a robust estimate of the automated pipeline and model classification capabilities (Hastie *et al.* 2001). In the basal setting, stratified  $k$ -fold cross validation was utilized such that approximately the same percentage of both target classes (i.e., *Pinus* and *Acer*) were present in both the training and validation sets. Here, a fold within the  $k$ -fold cross-validation procedure is understood as the technical term for a partitioning of the dataset into complementary subsets (Wong, 2015). Each complementary partitioning (fold) was generated by random sampling of the sample scan data. Data elements (e.g. image patches, concatenated orthogonal patches or subvolume chunks) of a single scan sample were always kept in the same fold, i.e. our partitioning and fold-generating algorithm operated on the scan-level. This was done to avoid data leakage. Another safety measure against data leakage was an automated

check of unique dataset identifiers in the configuration of the training experiment to avoid accidental overlaps of training and validation sets. In the basic setup,  $N_{\text{fold}} = 3$  cross-validation for binary classification of *Pinus sylvestris* and *Acer pseudoplatanus* was used. Using the CV strategy, we report the model performance as the average of the three folds together with the standard deviation as a measure of variation. To estimate optimistic bias effects due to the implicit model selection effect in our CV strategy, we additionally performed training experiments for all pipeline formulations in a single-fold setting with two scan dataset instances as fully independent test set. In addition to the standard stratified  $k$ -folds strategy, experiments with  $N_{\text{fold}} = 2$  cross-validation with stratification across the orientation groups (axial-like and transversal, due to the sample mounting) were conducted. Experiments with these splits were performed to assess the transfer of model skill across the two orientation groups. In this setting, the training was only conducted with all axial-like oriented datasets in fold 1 (yielding a training set size of eight instances), while the validation set was comprised of all datasets of the transversal orientation group (validation set size of 12). For fold 2, the training and validation data corpus were switched entirely to assess the inverse transfer capabilities.

#### TRAINING PROCEDURE

The internal weight parameters of the model were learned by iterative minimisation of the loss function with an optimisation algorithm. We used mini-batched stochastic gradient descent (Nesterov 1983; Robbins & Monro 1951) and regularised AdamW (Loshchilov & Hutter 2019) for this purpose. The methods rely on the iterative updating of the internal weight parameters in the direction of the negative gradient of the loss function. The iterative step size was controlled via the user set learning rate  $\eta$ . The loss function was categorical cross-entropy, used in the special binary variant for the corresponding classification task. The class-wise scores produced by the model through the final nonlinearity specified in equation (2) were used to compute the cross-entropy loss value. The cross-entropy loss is given by equation (3).

$$\text{CE}(z, t) = - \sum_{i=1}^c t_i \cdot \log([f(z)]_i) \quad (4)$$

Here,  $z$  denotes the score vector output by the final classification FC layer (Fig. 3 right),  $f$  the logistic function from equation (2) and  $t$  the target ground-truth vector in binary encoding for the positive and negative class. The qualifier “mini-batched” denotes the method of approximating the loss gradient using a randomly sampled subset (i.e. a mini-batch) of the dataset. The batch size  $b_s$  was variably adjusted to computer hardware memory constraints for the 2D, 2.5D and 3D experiments. Specifically for the volumetric 3D approach, small batch sizes close to fully stochastic gradient descent  $b_s \leq 2$  were observed to lead to less stable training with intermittent oscillations in validation set performance measures. This undesirable behaviour could be mitigated partly by gradient clipping during training and learning rate scheduling. Scheduling was applied by warming up the numerical value for the learning rate  $\eta$  to the global plateau value of  $\eta_{\text{max}} = 1 \times 10^{-3}$  and annealing it again for the final part of the training iterations. However, larger batch sizes  $b_s \geq 8$  were observed to be effective as well. We did not conduct systematic hyperparameter optimisation with Bayesian optimisation or (grid) search methods. To generally reduce the memory footprint of both data and model weights and to expedite the training experiment process, an optimised mixed-precision implementation (Micikevicius *et al.* 2018) of the training and inference code was employed. The model was trained on a workstation with an AMD Epyc 7452 32-core CPU and an nVidia Ampere A100 40 GB GPU. Evaluations were done partly on a second workstation with an AMD Ryzen 7 3700X CPU and an nVidia RTX 2080Ti 11 GB GPU. The software stack was run on the Debian 10 “buster” distribution for the first workstation and Windows 10 for the second workstation. We used Python version 3.11 and PyTorch version 2.0.1.

To reduce overfitting and improve model performance on possibly distorted data samples, we synthetically increased the amount of training data via a data augmentation scheme. The protocol incorporated rotations, artificial

Gaussian and Poisson noise corruption and Gaussian blurring. The augmentations were performed on the data element level, e.g. individually for each 2D patch, 2.5D triaxial concatenation and 3D subvolume chunk. Each model training run was allocated a fixed budget of parameter update iterations  $N_{\text{iter, max}}$  and epochs  $N_{\text{epoch, max}}$ . An epoch denotes the full pass-through of all distinct data elements in the training dataset. The training procedure was stopped whenever the first of the two constraints was reached. Every 500-th parameter update iteration was deemed a log iteration where the diagnostic information values loss function, training performance metrics and model weights and gradients were recorded. A validation phase run was performed every 1000 parameter update iterations. The validation run comprised the prediction of all hold-out data items of the respective cross validation set. Loss function value and performance metrics were logged. Over the course of the full training run, a dynamic registry retained the model parameters of the top-5 performing models in the validation run context. The performance criterion was accuracy (ACC), or alternatively, correct classification rate (CCR). The top-5 dynamic registry approach means that for every training experiment, we obtained five model parameter samples along the trajectory through weight parameter space. This classifier ensemble collection method is approximately comparable to snapshot ensembling (Huang *et al.* 2017). Thus, for every training experiment with a CV fold count of  $N_{\text{fold}}$ , a model ensemble of the size  $N_{\text{instance}} = 5 \cdot N_{\text{fold}}$  was saved. Performance metrics are thusly reported aggregated both fold-wise and instance-wise with their respective variation measure. Generally, an ensemble of models can be used to obtain and combine the predictions of multiple models, improving overall performance and robustness (Opitz & Maclin 1999).

## ROBUSTNESS EVALUATION

In addition to the basic experiments about the ability of the deep learning-based model to classify volumetric  $\mu$ CT scans of *Pinus sylvestris* and *Acer pseudoplatanus* wood species in the clean-data regime, we investigated the trained model's robustness against various input data transformations and corruptions during the inference phase. For this, six different synthetic parametrised data transformations were applied with increasing strength to the input data patches or chunks. For every transform and effect strength, a full validation prediction and classification performance computation was conducted. As noise-like corruptions, Gaussian noise, Poisson noise and Gibbs noise were selected. The second class of corruptions broadly affected image resolution. For this class, Gaussian smoothing, simulated low resolution and a zoom transform were chosen. The Gaussian noise was sampled from the corresponding distribution with zero-mean and variable standard deviation  $\sigma \in \{0.5, 1.25, 2.5, 3.0, 3.5, 4.0\}$  and applied additively to the data. Similarly, the Poisson noise was sampled from the corresponding distribution with the variable rate parameter  $\lambda \in \{0.025, 0.05, 0.1, 0.2, 1.3, 2.0\}$  and applied additively to the data. The Gibbs noise (Morelli *et al.* 2011) or ringing artifacts were applied with increasing strength steered through the parameter  $\alpha \in \{0.7, 0.8, 0.85, 0.875, 0.9, 0.95\}$ . For the image-resolution corrupting transformations, the Gaussian smoothing was applied with increasing kernel size  $\sigma \in \{1.0, 1.5, 2.0, 2.5, 3.0, 4.0\}$ . The synthetic low-resolution corruption uses down-sampling to a lower resolution with subsequent up-sampling to the original resolution, resulting in image quality loss and aliasing artifacts. This transform was parametrised by the scale factor or residual voxel budget factor sequence  $f \in \{0.8, 0.7, 0.5, 0.3, 0.15, 0.10\}$ . The last transform uses a zoom-out approach that shrinks the salient information-bearing data region into the centre of the patch or chunk, effectively presenting the network the original data in a mismatched voxel or pixel size. The shrunk salient region is furthermore embedded in an uninformative, artefactual region generated by padding with the edge voxel or pixel values. The utilised zoom factors were  $z \in \{0.90, 0.85, 0.80, 0.75, 0.70, 0.65\}$ . An overview of the effect of the transforms on a representative image patch is displayed in Fig. 5. Each row corresponds to one of the chosen image transforms. The first row-element is the clean data element, whereas subsequent row elements show the transformed data element with the corresponding parameter values. In addition, we provide a plot of the effect-size of the noise-like and resolution-like transformations in terms of two quantitative, abstract distance measures between the uncorrupted original data and the transformed data in Figure 6. This is intended to support some intuition about the relative mutation magnitude of every transform

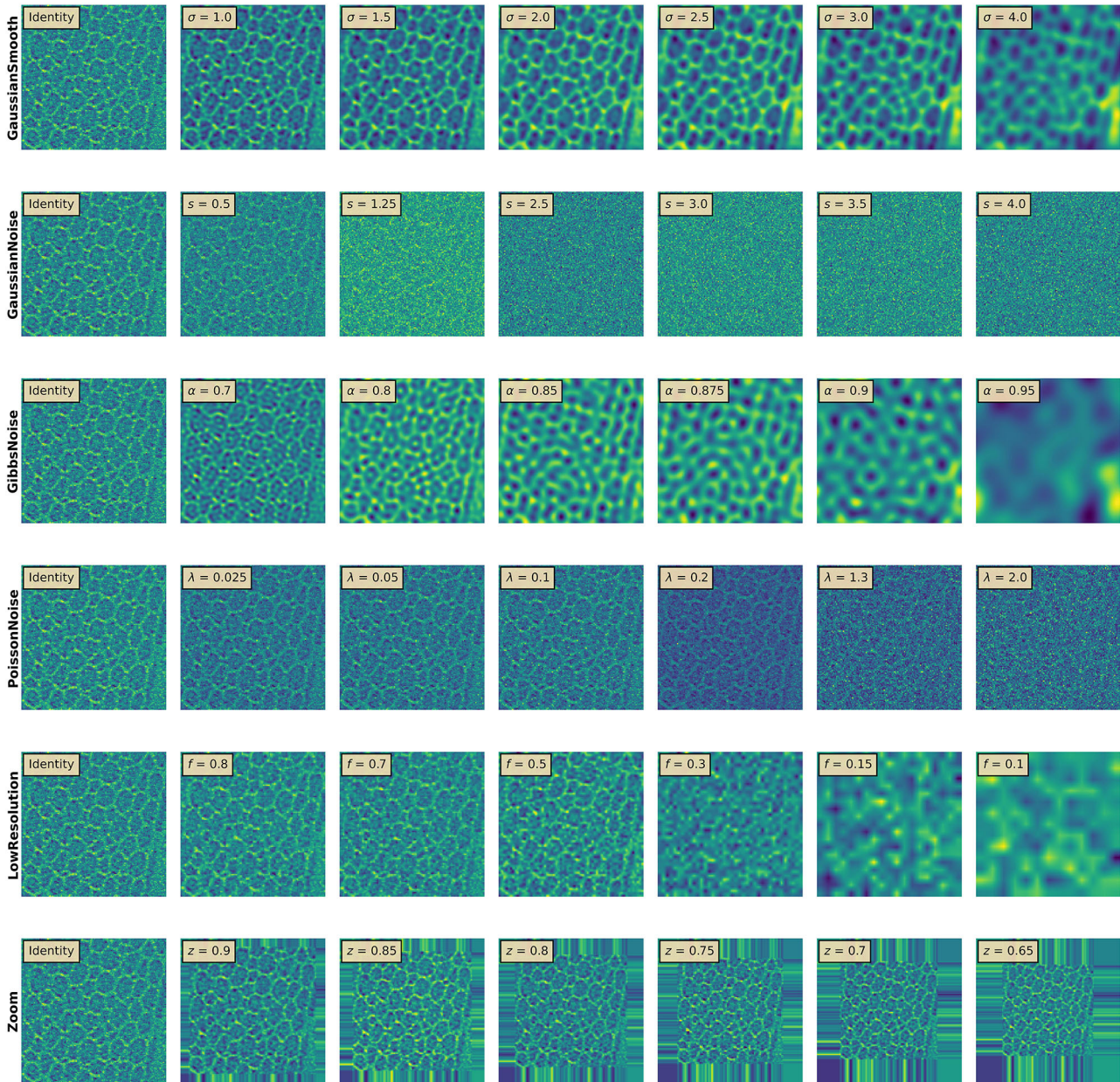
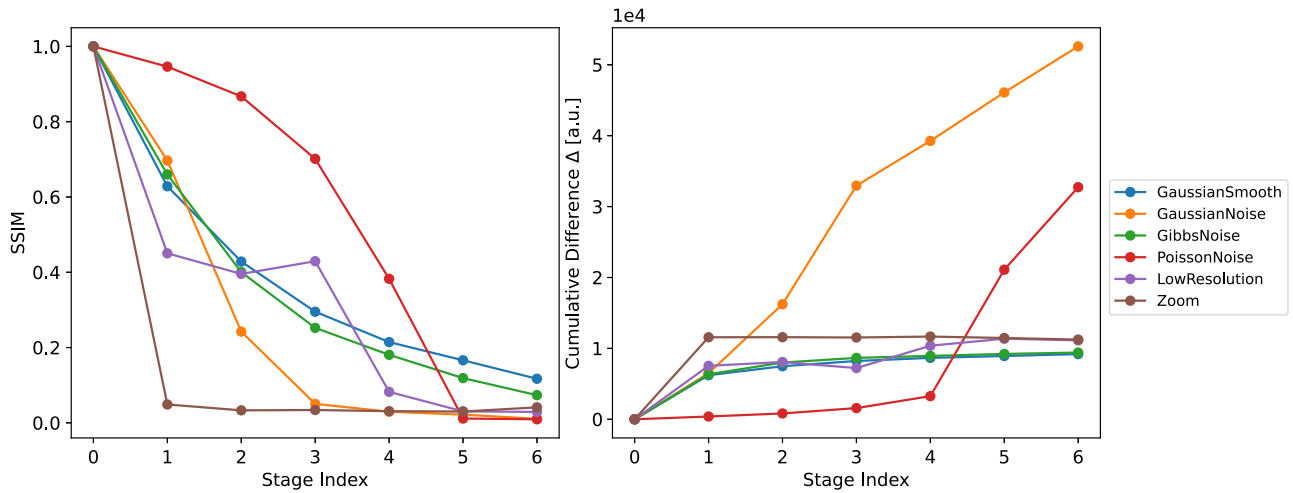


Fig. 5. Display of the effects of the six different transforms on a representative CT image slice. The first element of each row corresponds to the pristine CT data. Subsequent row elements show the slice mutated by the respective transform with the corresponding parameter value. The transform effect strength increases from left to right.

at every stage compared to the pristine data. First, the cumulative absolute difference  $\Delta$  between the original CT patch and the transformed patch, and second the structural similarity index measure (SSIM; Wang *et al.* 2004) are regarded.



**Fig. 6.** Quantification of the effect size of the various parametrized transforms on a representative image patch. On the left, the structural similarity index measure (SSIM), computed between the original and variably transformed image, is shown. On the right, the cumulative absolute difference  $\Delta$  of the original and variably transformed image is displayed.

## Results

### BASE TASK PERFORMANCE

In the basal setting of stratified three-fold cross-validation, all pipeline-model formulations could be trained to near-perfect accuracy or perfect accuracy. The values for the mean accuracy (ACC) and for the corresponding standard deviation  $\text{StDev}(\text{ACC})$  provided in Table 1 resulted from the aggregation over the  $N_{\text{instances}} = 15$  model instances, consisting of the three cross-validation folds and the  $N_s = 5$  sized snapshot ensembles from every fold. All models were trained with the AdamW optimiser and the specified default training procedure. The lowest classification accuracy of  $\text{ACC} = (0.988 \pm 0.010)$  was obtained in the 2D planar formulation with  $s = 64$ . All other formulations exceeded the accuracy threshold of 0.99. The pseudo-volumetric, triaxial and fully volumetric formulations commonly performed the binary classification task above the 0.9999 accuracy threshold. Perfect differentiation between *Acer* and *Pinus* was accomplished by the 3D formulation with chunks sizes  $s \in \{128, 256\}$ . For the experiments using two instances as test set, similarly high classification accuracy values between  $[0.98, 1.0]$  were observed (Supplementary Section 2).

### ROBUSTNESS EXPERIMENTS

The experiments to assess the resilience of the automated classifier to input corruptions were conducted using all dimensional formulations. Using the same model ensembles from the basal accuracy evaluation, we transformed the input data iteratively with increasing strength. In the following, the relative strength of the transformation is generically described by the stage index. Stage index 0 corresponded to no transformation, i.e., the basal result stated in the first section of the results. The following, strictly positive, stages 1 to 6 denote evaluations with actually transformed inputs. The transformation parameter value for the corresponding stage index can be read from the transformation parameter specifications in the methods section. Starting with one, the  $i$ -th parameter value corresponds to the stage number  $i$ . The resulting accuracy values were jointly aggregated over the folds and the instances of the snapshot ensemble. We report the devolution of the CNN classifier accuracy as line plots in Fig. 7. The standard deviation of the aggregation over instances and folds is displayed as the lightly shaded region around the

**Table 1.** Automated model performance measure as classification accuracy for the differing 2D, 2.5D and 3D dimensional approaches.

Dimensionality	Patch/chunk size $s$	ACC	SD (ACC)
2D planar	64	0.988	0.010
2D planar	128	0.99939	0.00057
2D planar	256	0.994	0.011
2.5D triaxial	64	0.99942	0.00027
2.5D triaxial	128	0.999992	0.000015
2.5D triaxial	256	0.999983	0.000048
3D volumetric	64	0.99996	0.00012
3D volumetric	128	1.0	0.0
3D volumetric	256	1.0	0.0

The chunk size  $s$  is given in pixel or voxel units. The mean accuracy is denoted as ACC, and SD denotes the standard deviation.

mean accuracy values. Shaded regions above the theoretical maximum accuracy value of 1.0 are caused by symmetric plotting of the standard deviation interval around the mean value.

Via the robustness evaluation experiments, the accuracy performance devolution to the point of usability breakdown or failure mode could be determined for every pipeline-model formulation. Generally, the volumetric formulations displayed in Fig. 7 exhibited the highest overall robustness against input data corruption. Both retained accuracy values above 0.9 for stages 1 through 3 for both transformation categories. Only Poisson noise posed a challenge that decreased classification accuracy below the 0.9 threshold value. Interestingly, Poisson noise proved to be the least challenging corrupting factor for the other pipeline formulations shown in the devolution plots for the planar and the triaxial formulation. All non-volumetric formulations however, presented a mean accuracy value below the 0.9 threshold for at least one corrupting transform as early as stage 2.

#### MODEL SKILL TRANSFER ACROSS ORIENTATION GROUPS

For the model skill transfer experiments across the two orientation groups of the dataset, we compared the 2D planar formulation with the 3D volumetric formulation on the patch/chunk size with edge length  $s = 256$ . This choice patch and chunk size provided the model with the greatest amount of information and context per sample to generalise across orientation groups. The results were generated with the default training procedure. The optimiser was AdamW with a weight decay value of 0.01. For this sub-experiment, we did not accumulate the performance over the two folds to better exhibit the directional aspect of the generalization performance of the formulations. For fold 1, the planar 2D approach yielded a validation set accuracy of  $ACC_{2D} = (0.608 \pm 0.067)$ . The volumetric formulation achieved an increased accuracy of  $ACC_{3D} = (0.774 \pm 0.032)$ . For fold 2, the classification accuracy advantage of the 3D approach with  $ACC_{3D} = (0.921 \pm 0.034)$  was even higher than the 2D approach accuracy of  $ACC_{2D} = (0.4995 \pm 0.0033)$ . Overall, a strictly higher orientation-wise generalisation skill was evident for the volumetric 3D approach in comparison to the planar 2D approach.

## Discussion

In this study, we described and implemented an automated pipeline for wood species classification with volumetric  $\mu$ CT data and deep neural networks and evaluated it in a conceptual hardwood-softwood classification context. The aim of the proof-of-concept study was to provide a first step in the development of a fully automated system for the benefit of wood scientists, timber trade entities and other stakeholders in need of automated wood species

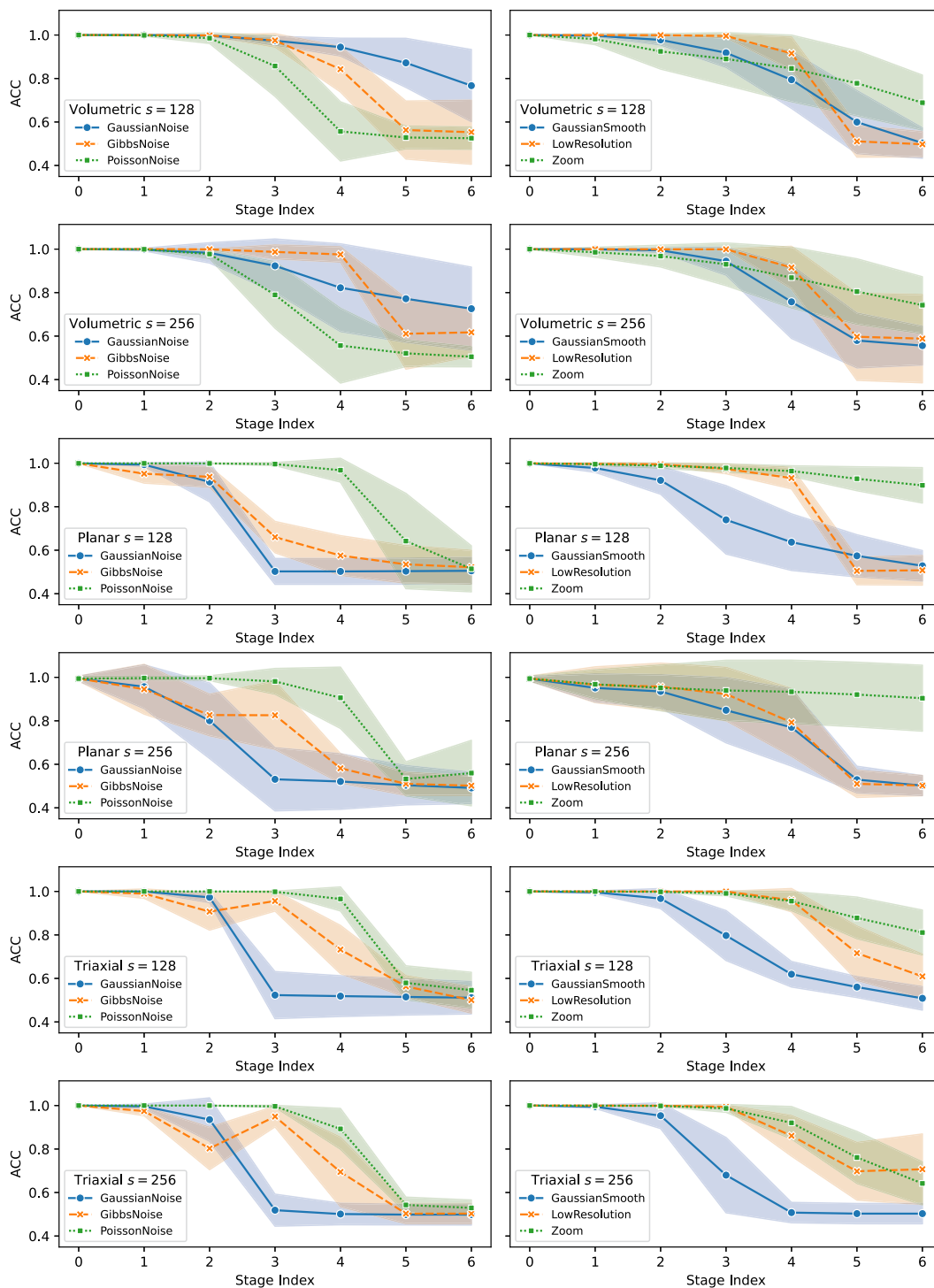


Fig. 7. Robustness evaluation of the classifier accuracy performance against variable intensity input data corruptions for volumetric, planar and triaxial formulations. For every formulation, results for larger FOV with  $s = 256$  and smaller FOV with  $s = 128$  are shown.

classification. We compared three different dimensional offspring approaches based on the ResNet-based deep learning model and pipeline: a planar 2D; a pseudo-volumetric, triaxial 2.5D; and a volumetric 3D formulation.

For the base case of softwood-hardwood binary classification without *in-silico* image corruption, and equal class-wise and orientation-group-wise dataset fractions in the training and validation set, all dimensional network formulations and patch/chunks sizes yielded near perfect accuracy, with minimal advantages for the volumetric 3D approach. This can be attributed to the major distinct cellular structures of the two wood species (Fig. 2) in conjunction with the high resolution and high contrast qualities of the sub- $\mu$ CT data that accurately images these differences (Fig. 1). Both factors lead to the presence of highly characteristic and discriminative structures in the input data, that are learned well by deep neural networks of all formulations, in turn effecting highly accurate classifications for all formulations. The reduced accuracy of 0.988 of the most downsampled planar formulation with  $s = 64 \text{ px} \approx 61.7 \text{ }\mu\text{m}$  is attributable to some image slices, where discriminative wood species features (detailed in Fig. 2) are less prominent or missing due to the small field-of-view (FOV). There, the informational substance about anatomical structures appears to be insufficient for the higher classification observed in the other formulations. The slight accuracy drop for the planar parsimonious formulation is contrasted by the optimal performance of the larger volumetric formulations. The larger FOV and spatial context of the volumetric  $\mu$ CT data could be harnessed particularly effectively by the 3D formulations of the pipeline and model. With subvolume edge sizes of  $s \in \{128, 256\}$ , these formulations achieved overall optimal performance with perfect classification accuracy. The volumetric model with  $s = 64$  achieved the best accuracy compared to the size-corresponding 2.5D and 2D formulations, displaying a high volumetric sample efficiency. The other approaches using the 2.5D and 3D formulation showed very high accuracy values of at least 0.999. Although it is not possible to deduce a general advantage of the volumetric formulation from these experiments with distinct hardwood and softwood species due to the minuscule accuracy differences, we surmise that the relative accuracy advantage of the volumetric formulation increases for more complex wood species datasets. There, distinctive features may only manifest in a three-dimensional and anisotropic manner.

Compared to existing formulations based on optical macroscopic surface imaging or microscopic cross-section imaging (Hwang & Sugiyama 2021), further advantages of the described pipeline entail the non-destructive and objectivized character of the  $\mu$ CT data acquisition process. Of the proposed formulations, the 2D planar approach closely approximates the methodology employed in optical microscopy imaging. Here, data dimensionality and possible machine learning pipelines align. Nonetheless, a direct equivalence cannot be asserted, as optical microscopy data includes spectral colour information, whereas basal CT images provide only absorption values as the singular per-pixel measurement. Here, phase-contrast CT may be useful to retrieve additional information about the physico-chemical parameters of the wood sample due to its improved soft-matter and density contrast for materials with low atomic numbers in comparison to pure absorption-transmission-based imaging (Diemoz *et al.* 2012). Beyond that, dual-energy or multispectral CT (Garnett 2020) can be harnessed to extract further information from wood samples. Generally,  $\mu$ CT data can readily be obtained from highly valued artifacts embedded into shells whose structural integrity is paramount. The proposed imaging setup utilises region-of-interest scanning for the described volumes, which does not necessitate physical cutting of the sample to the dimensions of the subvolume dimensionality utilised in the machine vision experiments. Considering possible damaging sample alterations due to X-rays, we identify radiolysis as a potential factor. However, the creation of oxidising or reducing molecule species via X-ray photons (Féron 2012) is not expected to play a role in samples with low residual water content, such as the samples used in our experiments. For waterlogged wood samples, magnetic resonance imaging can be a better fit (Capuani *et al.* 2020; Longo *et al.* 2023) due to reduced CT image contrast and problems induced by radiolysis. Another advantage is that the volumetric data in the  $\mu$ CT scanner setup is recorded in a controlled manner, where detector configuration, tube current, and potential (i.e., wood sample illumination) and reconstruction algorithms are tightly controlled and quantified parameter settings. This means that subsequent deep-learning-based classification algorithms can be calibrated well to the data source. Then, in-distribution and out-of-distribution input data to the DNN with respect to the image

data source are well-defined concepts. For mission-critical applications, this knowledge allows for targeted data augmentation or out-of-distribution detection (Liang *et al.* 2018) to avert high-confidence, low-correctness predictions by the classifier (Nguyen *et al.* 2015). Initial results for this line of work were generated with the performed robustness experiments. The observed performance degradation for increasingly strong input data transforms varies for the 2D, 2.5D and 3D formulation. In line with the findings concerning the accuracy on pristine data and the generalisation skill across orientation groups, we discovered an advantage of the 3D approach. However, for specific combinations of corrupting transforms (i.e., Poisson noise) and patch or chunk size  $s$ , non-volumetric formulations may exhibit better performance retention. Investigating the root cause of these performance differences is another interesting line of work. In realistic deployments, distorted input data resulting from suboptimal images is to be expected. The presented results (Fig. 7) provide valuable quantitative insights about the expected accuracy loss.

The volumetric  $\mu$ CT data resulting from the described imaging protocol straddles the traditional resolution gap between macroscopic images and micrographs. This enables subsequent learning-based algorithms to extract and combine global texture information, macroscopic fibre and canal structures and microscopic localised cell information (Hwang & Sugiyama 2021) for task-optimal purposes. Convolutional deep neural networks act as hierarchical feature detectors and may select task-optimal discriminative features during the training phase (LeCun *et al.* 2015). In the context of volumetric input data, this feature construction is not restricted to a single slice with a predetermined orientation fixed prior to training by the image acquisition setup. Instead, it can occur freely in the 3D volume during the training phase. This potentially constitutes another advantage for volumetric methods in contexts where data acquisition and processing is performed without orientation or registration, resulting in oblique fiber orientation (see Supplementary Figs S5 and S6 for examples). In volumetric imaging, no prior decision about the wood section aspect (transversal vs. tangential vs. radial) with corresponding domain knowledge of estimated classification accuracy has to be made manually (Barmpoutis *et al.* 2018). Imbuing network-level rotation equivariance into the naturally merely translation equivariant convolution operation (Fukushima 1980; Bronstein *et al.* 2021), either by data augmentation or architectural choices (Chidester *et al.* 2019), would be a very worthwhile extension of the presented experiments. We expect that the benefits of using the volumetric method in conjunction with rotation equivariance techniques extend to and become more pronounced when image data is captured from samples positioned with oblique angles in real-world scenarios. This hypothesis is supported by the advantage of the 3D formulation in the context of model skill transfer for orientation-group stratified training. The volumetric formulation achieved better classification accuracy than the planar formulation for the orientation-wise training. As the training procedure and data basis were fundamentally similar, we deduce that increased information per input data sample and spatial context allowed the model to learn features that better generalise across orientation groups. The observed gap towards the non-group-wise trained model may be closed by including the mentioned rotation-equivariance-imbuing modifications to the network architecture. We conjecture that this advantage of the 3D approach may also generalise to envisioned, more complex classification settings where differences between closely related wood species may manifest only in an anisotropic manner in highly localised subvolumes. Generally, it would be interesting in terms of deep learning to explore architectures or augmentation techniques that directly address the typical orientation groups of wood science. A balanced and volumetric dataset with all three typical aspects (i.e., transversal, radial, and tangential and additional oblique orientations) as scan directions in conjunction with an extended orientation group transfer evaluation could help disentangle aspect-related questions for the anticipated large-scale volumetric  $\mu$ CT dataset.

As stated, further interesting deep learning related investigations are expected to arise when applying the presented pipeline and model variations to more complex datasets that incorporate a larger number and more closely related wood species. For example, it would be interesting to know whether the utilized, lean model class of ResNet can retain the near-perfect classification accuracy and, if not, which available models and techniques, for example transformers, further augmentation and hyperparameter tuning and stochastic weight averaging, manifest as valuable to automated wood classification performance on volumetric  $\mu$ CT data.

## WEAKNESSES

The wood species utilised for the training and evaluation of the automated framework possess a substantially different cellular structure and are thus readily distinguishable. For a useful demonstration of the framework, wood identification with closely related timbers possessing similar wood structures would be required. However, the study does not aim to carry out complex wood species identification. The study is more concerned with presenting the integrated technological framework of combining deep learning with high-resolution volumetric  $\mu$ CT data and releasing the corresponding code infrastructure in the form of an open-source Python package.

Apart from the transfer to a more wood species-rich classification context, we identify a second necessity. To accurately assess the automated system capacity to replace a human expert in wood species identification, it must undergo evaluation using a variety of field-sampled specimens. Here, temporal variations in the wood microstructure due to differing age of the timber source or cellular alterations due to dehydration and cessation of accretion and transport processes may play an important role.

Generally, a larger data set is also required for future work. The smaller 20-scan data set motivated the usage of cross-validation without a fully independent test set, since it can produce realistic performance estimates in the small-data regime (Xu & Goodacre 2018). Although the data homogeneity (compare Fig. S1), quality and standard deviation evaluation of the CV-ensemble evaluation (see Table 1) suggest that the resulting model performance does not contain a substantive optimistic bias, a fully separate test set for all experiments is warranted for subsequent scenarios with direct field use. Cross-validation still can subtly lead to over-optimistic performance evaluation without a test set through implicit model selection effects. Since no explicit hyperparameter search for model selection was performed and we observed similarly high classification scores for initial single-fold-training experiments with a fully separate test set (Supplementary Section 2), we estimate that this effect is minimal for the hardwood-softwood binary classification context. Summarising, many interesting areas of research and validation experiments arise naturally when applying the presented framework in the context of more complex datasets of wood species.

Considering the near-term practical applicability of the method, we currently identify the high cost of equipment and longer image acquisition time as hindering factors for potential adoption. These considerations are minor in the context of envisioned applications such as bespoke forensic investigations, but more important in other applications such as low-cost and high-throughput environments, including customs and inspections. While the ClickCT system is currently expensive, it has the potential to be customised, particularly for customs applications. This specialisation could lead to cost reductions in hardware over time due to economies of scale, making the system more affordable and feasible for utilisation in customs contexts. With targeted modifications, scaled production, minimised personnel costs and improved throughput due to tailored machine vision support, the technology could become economically viable for routine inspections and other customs needs.

## PROSPECTS

Due to the demonstrated accuracy and robustness of the deep learning-based pipeline for the pilot hardwood-softwood  $\mu$ CT dataset, the establishment of a state-of-the-art method for fully automated non-destructive wood identification is imaginable. Especially in the context of forensic applications, the advantages of the method may come to fruition. For this, further in-depth validation and testing of the pipeline with a yet to be compiled extensive wood species dataset from high-resolution volumetric  $\mu$ CT data is worthwhile. Using insights from the pilot dataset results of this study, the per-sample scan time could potentially be reduced by 87.5 % from 2 hours down to 15 minutes. With the 20 $\times$  objective lens of the ClickCT and 4 $\times$  binning, a measurement region of 0.5 mm<sup>3</sup> can be recorded with similar image quality. The relevant chunk sizes with  $s \in \{128, 256\}$  shown to allow near-perfect classification accuracy can still be extracted from these optimised region-of-interest scan volumes. However, for other wood species with typical anatomical feature sizes, the measurement region may have to be enlarged to capture sufficient anatomical context. The insights generated by the presented experiments increase the production feasibility of an

extensive volumetric CT reference database. The joint advantage of objectivized imaging with volumetric data could mitigate some of the challenges identified in recent literature (He *et al.* 2024) concerning the automation of wood identification. Large-scale volumetric CT datasets could improve unsupervised feature extraction while averting dataset and feature incompatibility often present in macroscopic or microscopic section datasets due to differing data collection methods or section aspects.

Possible extensions of the presented study are then to further gauge the analytical classification prowess of the model and to conduct further in-depth robustness testing using the already established computational code infrastructure. Building a varied, high quality open-source benchmark dataset of volumetric  $\mu$ CT scans of wood species would be an unprecedented and highly worthwhile goal to enable the establishment of the method, since current complementing databases are of genuinely planar nature (Hwang & Sugiyama 2021) and consist of either micrographs (Lens *et al.* 2020), macroscopic images (Kwon *et al.* 2017) or stereograms (Ravindran & Wiedenhoef 2020).

## Conclusions

This study shows the potential of a wood species classification pipeline based volumetric sub- $\mu$ CT as the imaging technique and deep neuronal networks as the fully automated algorithmic backbone. A dataset of 20 instances of softwood (*Pinus sylvestris*) and hardwood (*Acer pseudoplatanus*) is used as a basis to produce sub- $\mu$ CT scan instances. The description and evaluation of the sub- $\mu$ CT modality and classification model with the purpose-built open-source pipeline constitute the contribution of the proof-of-concept study, which aimed to link high-resolution volumetric sub- $\mu$ CT imaging with machine learning in the context of forensic wood identification. Planar, triaxial and volumetric approaches to classification were generally similarly successful in classifying the distinct species in the basal case, whereas volumetric approaches showed improved skill transfer for contexts with variable sample orientation. The study evaluates opportunities and weaknesses of the technologies involved to illustrate avenues for the further development of this technology to help meet the existing and future challenges of anatomical wood identification for both wood science and the wider lumber industry. Finally, we observed excellent accuracy of all automated classifier formulations for our conceptual hardwood-softwood binary classification task, highlighting the potential of combining volumetric sub- $\mu$ CT imaging with machine learning to further develop vision-based automated wood identification.

## Supplementary materials

Data is available on <https://doi.org/10.1163/22941932-bja10196> under Supplementary Materials.

## Acknowledgements

The authors would like to thank Dr. Gerald Koch (Thünen Institut of Wood Research, Hamburg) for stimulating discussions and suggestions during the wood anatomical studies for the present investigation and Dinah Führmann for revising the English text.

## References

Barmpoutis P, Dimitropoulos K, Barboutis I, Grammalidis N, Lefakis P. 2018. Wood species recognition through multidimensional texture analysis. *Comput. Electron. Agric.* 144: 241–248. DOI: 10.1016/j.compag.2017.12.011.

- Beeckman H, Blanc-Jolivet C, Boeschoten LE, Jez Willian Batista Braga, Cabezas JA, *et al.* 2020. *Overview of current practices in data analysis for wood identification. A guide for the different timber tracking methods.* Global Timber Tracking Network, GTTN secretariat, European Forest Institute and Thünen Institute, Hamburg. DOI: 10.13140/RG.2.2.21518.79689.
- Bello I, Fedus W, Du X, Cubuk ED, Srinivas A, Lin T-Y, Shlens J, Zoph B. 2021. Revisiting ResNets: Improved training and scaling strategies. *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*. DOI: 10.48550/arXiv.2103.07579.
- Bello R-W. 2023. Wood species identification using mask RCNN-residual network. *Pro Ligno* 19(1): 41–51.
- Bird MI, Ascough PL, Young IM, Wood CV, Scott AC. 2008. X-ray microtomographic imaging of charcoal. *J. Archaeol. Sci.* 35(10): 2698–2706. DOI: 10.1016/j.jas.2008.04.018.
- Bronstein MM, Bruna J, Cohen T, Velicković P. 2021. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. *arXiv: arXiv:2104.13478*.
- Capuani S, Stagno V, Missori M, Sadori L, Longo S. 2020. High-resolution multiparametric MRI of contemporary and waterlogged archaeological wood. *Mag. Res. Chem.* 58(9): 860–869. DOI: 10.1002/mrc.5034.
- Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, *et al.* 2022. MONAI: An open-source framework for deep learning in healthcare. *arXiv: arXiv:2211.02701*. DOI: 10.48550/arXiv.2211.02701.
- Chidester B, Zhou T, Do MN, Ma J. 2019. Rotation equivariant and invariant neural networks for microscopy image analysis. *Bioinformatics* 35(14): i530–i537. DOI: 10.1093/bioinformatics/btz353.
- Diemoz PC, Bravin A, Coan P. 2012. Theoretical comparison of three X-ray phase-contrast imaging techniques: propagation-based imaging, analyser-based imaging and grating interferometry. *Opt. Express* 20(3): 2789. DOI: 10.1364/OE.20.002789.
- Dierickx S, Genbrugge S, Beeckman H, Hubau W, Kibleur P, Van Den Bulcke J. 2024. Non-destructive wood identification using X-ray  $\mu$ CT scanning: which resolution do we need? *Plant Methods* 20(1): 98. DOI: 10.1186/s13007-024-01216-0.
- Dormontt EE, Boner M, Braun B, Breulmann G, Degen B, *et al.* 2015. Forensic timber identification: It's time to integrate disciplines to combat illegal logging. *Biol. Conserv.* 191: 790–798. DOI: 10.1016/j.biocon.2015.06.038.
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, *et al.* 2021. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv: arXiv:2010.11929*. DOI: 10.48550/arXiv.2010.11929.
- Duncan KE, Czymbek KJ, Jiang N, Thies AC, Topp CN. 2022. X-ray microscopy enables multiscale high-resolution 3D imaging of plant cells, tissues, and organs. *Plant Physiol.* 188(2): 831–845. DOI: 10.1093/plphys/kiab405.
- Fabijańska A, Danek M, Barniak J. 2021. Wood species automatic identification from wood core images with a residual convolutional neural network. *Comp. Electron. Agric.* 181: 105941. DOI: 10.1016/j.compag.2020.105941.
- Féron D. 2012. Overview of nuclear materials and nuclear corrosion science and engineering. In: *Nuclear Corrosion Science and Engineering*: 31–56. Woodhead, Cambridge. DOI: 10.1533/9780857095343.1.31.
- Friedrichsdorf SP, Arana-Chavez VE, Cattaneo PM, Spin-Neto R, Dominguez GC. 2019. Effect of the software binning and averaging data during microcomputed tomography image acquisition. *Sci. Rep.* 9(1): 10562. DOI: 10.1038/s41598-019-46530-z.
- Fukushima K. 1969. Visual feature extraction by a multilayered network of analog threshold elements. *IEEE Trans. Syst. Sci. Cyber.* 5(4): 322–333. DOI: 10.1109/TSSC.1969.300225.
- Fukushima K. 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cyber.* 36(4): 193–202. DOI: 10.1007/BF00344251.
- Garnett R. 2020. A comprehensive review of dual-energy and multi-spectral computed tomography. *Clin. Imag.* 67: 160–169. DOI: 10.1016/j.clinimag.2020.07.030.
- Gejdoš M, Gergeš T, Michajlová K, Bucha T, Gracovský R. 2023. The accuracy of CT scanning in the assessment of the internal and external qualitative features of wood logs. *Sensors* 23(20): 8505. DOI: 10.3390/s23208505.
- Grabner M, Salaberger D, Okochi T. 2009. The need of high resolution  $\mu$ -X-ray CT in dendrochronology and in wood identification. In: *Proceedings of the 2009 6th International Symposium on Image and Signal Processing and Analysis. Salzburg*: 349–352. DOI: 10.1109/ISPA.2009.5297695.
- Haag V, Kirsch S, Koch G, Zemke V, Richter H-G, Kaschuro S. 2018. Wooden musical instruments – different forms of knowledge. In: Pérez MA, Marconi E (eds.), *Wooden musical instruments – different forms of knowledge. COST Action FP1302*: 143–156. Philharmonie de Paris, Paris.
- Haag V, Zemke VT, Lewandowski T, Zahnen J, Hirschberger P, Bick U, Koch G. 2020. The European charcoal trade. *IAWAJ.* 41(4): 463–477. DOI: 10.1163/22941932-bja10017.

- Haag V, Dremel K, Zabler S. 2022. Volumetric imaging by micro computed tomography: a suitable tool for wood identification of charcoal. *IAWA J.* 44(2): 210–224. DOI: 10.1163/22941932-bja10106.
- Haneca K, Deforce K, Boone MN, Van Loo D, Dierick M, Van Acker J, Van Den Bulcke J. 2012. X-ray sub-micron tomography as a tool for the study of archaeological wood preserved through the corrosion of metal objects. *Archaeometry* 54(5): 893–905. DOI: 10.1111/j.1475-4754.2011.00640.x.
- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, et al. 2020. Array programming with NumPy. *Nature* 585(7825): 357–362. DOI: 10.1038/s41586-020-2649-2.
- Hastie T, Tibshirani R, Friedman JH. 2001. *The elements of statistical learning: data mining, inference, and prediction*. Springer, New York, NY.
- He K, Zhang X, Ren S, Sun J. 2015. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile: 1026–1034. DOI: 10.1109/ICCV.2015.123.
- He K, Zhang X, Ren S, Sun J. 2016. Deep residual learning for image recognition. In: *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 770–778.
- He T, Lu Y, Jiao L, Zhang Y, Jiang X, Yin Y. 2020. Developing deep learning models to automate rosewood tree species identification for CITES designation and implementation. *Holzforschung* 74(12): 1123–1133. DOI: 10.1515/hf-2020-0006.
- He X, Pelt DM, Gao J, Gravendeel B, Zhu P, Chen S, Qiu J, Lens F. 2024. Machine learning-based wood anatomy identification: towards anatomical feature recognition. *IAWA J.* 45(4): 457–475. DOI: 10.1163/22941932-bja10157.
- Helmling S, Olbrich A, Tepe L, Koch G. 2016. Qualitative and quantitative characteristics of macerated vessels of 23 mixed tropical hardwood (MTH) species: a data collection for the identification of wood species in pulp and paper. *Holzforschung* 70(9): 839–844. DOI: 10.1515/hf-2015-0195.
- Helmling S, Olbrich A, Heinz I, Koch G. 2018. Atlas of vessel elements: Identification of Asian timbers. *IAWA J.* 39(3): 249–352. DOI: 10.1163/22941932-20180202.
- Huang G, Li Y, Pleiss G, Liu Z, Hopcroft JE, Weinberger KQ. 2017. Snapshot ensembles: train 1, get M for free. *arXiv*: arXiv:1704.00109.
- Hubau W, Den Bulcke JV, Kitin P, Brabant L, Van Acker J, Beeckman H. 2013. Complementary imaging techniques for charcoal examination and identification. *IAWA J.* 34(2): 147–168. DOI: 10.1163/22941932-00000013.
- Hüllermeier E, Waegeman W. 2021. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Mach. Learn.* 110(3): 457–506. DOI: 10.1007/s10994-021-05946-3.
- Hwang S-W, Sugiyama J. 2021. Computer vision-based wood identification and its expansion and contribution potentials in wood science: A review. *Plant Methods* 17(1): 47. DOI: 10.1186/s13007-021-00746-1.
- Hwang S-W, Tazuru S, Sugiyama J. 2020. Wood identification of historical architecture in Korea by synchrotron x-ray microtomography-based three-dimensional microstructural imaging. *J. Kor. Wood Sci. Technol.* 48(3): 283–290. DOI: 10.5658/WOOD.2020.48.3.283.
- Ioffe S, Szegedy C. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv*: arXiv:1502.03167.
- Irie MS, Spin-Neto R, Borges JS, Wenzel A, Soares PBF. 2022. Effect of data binning and frame averaging for micro-CT image acquisition on the morphometric outcome of bone repair assessment. *Sci. Rep.* 12(1): 1424. DOI: 10.1038/s41598-022-05459-6.
- Kırbaş İ, 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.
- Kobayashi K, Hwang S-W, Okochi T, Lee W-H, Sugiyama J. 2019. Non-destructive method for wood identification using conventional X-ray computed tomography data. *J. Cult. Herit.* 38: 88–93. DOI: 10.1016/j.culher.2019.02.001.
- Koch G, Haag V. 2015. Control of internationally traded timber – The role of macroscopic and microscopic wood identification against illegal logging. *J. Forensic Res.* 6(6). DOI: 10.4172/2157-7145.1000317.
- Koddenberg T, Greving I, Hagemann J, Flenner S, Krause A, et al. 2021. Three-dimensional imaging of xylem at cell wall level through near field nano holotomography. *Sci. Rep.* 11(1): 4574. DOI: 10.1038/s41598-021-83885-8.
- Krizhevsky A, Sutskever I, Hinton GE. 2012. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems 25 (NIPS 2012)*: 1097–1105.

- Kwon O, Lee HG, Lee M-R, Jang S, Yang S-Y, Park S-Y, Choi I-G, Yeo H. 2017. Automatic wood species identification of Korean softwood based on convolutional neural networks. *J. Kor. Wood Sci. Technol.* 45(6): 797–808. DOI: 10.5658/WOOD.2017.45.6.797.
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521(7553): 436–444. DOI: 10.1038/nature14539.
- Lens F, Liang C, Guo Y, Tang X, Jahanbanifard M, Da Silva FSC, Ceccantini G, Verbeek FJ. 2020. Computer-assisted timber identification based on features extracted from microscopic wood sections. *IAWA J.* 41(4): 660–680. DOI: 10.1163/22941932-bja10029.
- Lewandrowski T, Koch G, Haag V. 2022. Anatomical and topochemical features of the genera *Acacia*, *Acaciella*, *Senegalia* and *Vachellia*. *IAWA J.* 43(4): 353–371. DOI: 10.1163/22941932-bja10093.
- Liang S, Li Y, Srikant R. 2018. Enhancing the reliability of out-of-distribution image detection in neural networks. *arXiv*: arXiv:1706.02690.
- Longo S, Egizi F, Stagno V, Di Trani MG, Marchelletta G, Gili T, Fazio E, Favero G, Capuani S. 2023. A multi-parametric investigation on waterlogged wood using a magnetic resonance imaging clinical scanner. *Forests* 14(2): 276. DOI: 10.3390/f14020276.
- Loshchilov I, Hutter F. 2019. Decoupled weight decay regularization. *arXiv*: arXiv:1711.05101.
- Micikevicius P, Narang S, Alben J, Diamos G, Elsen E, et al. 2018. Mixed precision training. *arXiv*: arXiv:1710.03740.
- Morelli JN, Runge VM, Ai F, Attenberger U, Vu L, Schmeets SH, Nitz WR, Kirsch JE. 2011. An image-based approach to understanding the physics of MR artifacts. *RadioGraphics* 31(3): 849–866. DOI: 10.1148/rg.313105115.
- Nesterov YE. 1983. A method of solving a convex programming problem with convergence rate  $O(k^{-2})$ . *Dokl. Akad. Nauk SSSR*: 543–547.
- Nguyen A, Yosinski J, Clune J. 2015. Deep neural networks are easily fooled: High confidence predictions for unrecognisable images. In: *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*: 427–436. DOI: 10.1109/CVPR.2015.7298640.
- Opitz D, Maclin R. 1999. Popular ensemble methods: an empirical study. *J. Artif. Intell. Res.* 11: 169–198. DOI: 10.1613/jair.614.
- Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L. 2019. Pytorch: An imperative style, high-performance deep learning library. In: *33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada*: 8026–8037. *arXiv*: arXiv:1912.01703v1.
- Patera A, Bonnin A, Mokso R. 2021. Micro- and nano-scales three-dimensional characterisation of softwood. *J. Imag.* 7(12): 263. DOI: 10.3390/jimaging7120263.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. 2011. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.* 12(85): 2825–2830.
- Ravindran P, Costa A, Soares R, Wiedenhoeft AC. 2018. Classification of CITES-listed and other neotropical Meliaceae wood images using convolutional neural networks. *Plant Methods* 14(1): 25. DOI: 10.1186/s13007-018-0292-9.
- Ravindran P, Wiedenhoeft AC. 2020. Comparison of two forensic wood identification technologies for ten Meliaceae woods: computer vision versus mass spectrometry. *Wood Sci. Technol.* 54(5): 1139–1150. DOI: 10.1007/s00226-020-01178-1.
- Richter HG, Dallwitz MJ. 2015. Commercial timbers: descriptions, illustrations, identification, and information retrieval (2002 onwards). Available online at <https://www.delta-intkey.com/wood/index.htm>.
- Robbins H, Monro S. 1951. A stochastic approximation method. *Ann. Math. Statist.* 22(3): 400–407. DOI: 10.1214/aoms/1177729586.
- Sarker IH. 2021. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *Sn. Comput. Sci.* 2(6): 420. DOI: 10.1007/s42979-021-00815-1.
- Silva JL, Bordalo R, Pissarra J, De Palacios P. 2022. Computer vision-based wood identification: a review. *Forests* 13(12): 2041. DOI: 10.3390/f13122041.
- Stebani J, Lewandrowski T, Dremel K, Zabler S, Haag V. 2024. *Companion data artifacts: Technical framework demonstration for deep learning-based wood species classification with advanced sub- $\mu$ -CT imaging*. DOI: 10.5281/ZENODO.13905948.
- Stelzner J, Million S, Stelzner I, Nelle O, Banck-Burgess J. 2023. Micro-computed tomography for the identification and characterization of archaeological lime bark. *Sci. Rep.* 13(1): 6458. DOI: 10.1038/s41598-023-33633-x.
- Talaei Khoei T, Ould Slimane H, Kaabouch N. 2023. Deep learning: systematic review, models, challenges, and research directions. *Neural Comput. Appl.* 35(31): 23103–23124. DOI: 10.1007/s00521-023-08957-4.
- Vacek O, Gergeř T, Bucha T, Gracovský R, Gejdoř M. 2024. Automatic wood species classification and pith detection in log CT images. *Forests* 15(12): 2207. DOI: 10.3390/f15122207.

- Van Den Bulcke J, Boone M, Van Acker J, Stevens M, Van Hoorebeke L. 2009. X-ray tomography as a tool for detailed anatomical analysis. *Ann. For. Sci.* 66(5): 508–508. DOI: 10.1051/forest/2009033.
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13(4): 600–612. DOI: 10.1109/TIP.2003.819861.
- Willemink MJ, Noël PB. 2019. The evolution of image reconstruction for CT – from filtered back projection to artificial intelligence. *Eur. Radiol.* 29(5): 2185–2195. DOI: 10.1007/s00330-018-5810-7.
- Withers PJ, Bouman C, Carmignato S, Cnudde V, Grimaldi D, Hagen CK, Maire E, Manley M, Du Plessis A, Stock SR. 2021. X-ray computed tomography. *Nature Rev. Methods Primers* 1(1): 18. DOI: 10.1038/s43586-021-00015-4.
- Xu Y, Goodacre R. 2018. On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalisation performance of supervised learning. *J. Anal. Test* 2(3): 249–262. DOI: 10.1007/s41664-018-0068-2.
- Zabler S, Ullherr M, Fella C, Schielein R, Focke O, Zeller-Plumhoff B, Lhuissier P, DeBoever W, Hanke R. 2020. Comparing image quality in phase contrast sub  $\mu$  X-ray tomography—A round-robin study. *Nuclear Instrum. Methods Phys. Res. A* 951: 162992. DOI: 10.1016/j.nima.2019.162992.
- Zahnen J, Hirschberger P, Haag V, Lewandrowski T. 2020. *Analysis of the EU charcoal market*. WWF Germany, Berlin.

*Edited by Yafang Yin*