റ

Original article

Yusuke Kita and Junji Sugiyama*

Wood identification of two anatomically similar Cupressaceae species based on two-dimensional microfibril angle mapping

https://doi.org/10.1515/hf-2020-0079 Received March 26, 2020; accepted October 14, 2020; published online November 13, 2020

Abstract: Identifying two anatomically similar species of Cupressaceae, Chamaecyparis obtusa and Thujopsis spp., is important to better understand the culture of wood use in Japan. However, the conventional method, which involves observing their cross-field pitting, cannot identify them in many cases. This study solves the above problem by introducing an anatomical criterion based on the micro fibril angle (MFA). MFA values were obtained through twodimensional MFA images using the uniaxial optical anisotropy of cellulose microfibrils. A combination of the preprocessed MFA images and a convolutional neural network (CNN) yielded an accuracy nearly of 90% in classifying these species in cases of present and old wood specimens. Our feature extraction and classification techniques provide a new way for describing the anatomical features of wood and identifying featureless softwoods. Using the model interpretation-related methodologies of the CNN, distinct features of the two wood species were partly explained by MFA anisotropy in the S2 wall induced by the existence of pits.

Keywords: convolutional neural network (CNN); Cupressaceae; microfibril angle (MFA); wood anatomy; wood identification.

Present address: Junji Sugiyama, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan

E-mail: sugiyama.junji.6m@kyoto-u.ac.jp. https://orcid.org/0000-0002-5388-4925

Yusuke Kita, Research Institute for Sustainable Humanosphere, Kyoto University, Uji, Kyoto, 611-0011, Japan,

E-mail: kita.yusuke.46x@outlook.jp

1 Introduction

The culture of wood use in Japan has long flourished because the country has an abundance of forested land. Knowledge of the use of different species of wood in historical sites can help us learn more about our cultural background. An important example of this is two species of Cupressaceae, Chamaecyparis obtusa (hinoki in Japanese) and Thujopsis spp., including Thujopsis dolabrata (asunaro) and Thujopsis dolabrata var. hondae (hinokiasunaro). These species have been used in structural materials in traditional buildings such as temples because of their excellent mechanical properties and durability (Ido 2018). However, there is marked difference in their use despite the similarities between them, grounded in preferences that ancient populations had for using each species (Hirai 1996). Hence, the accurate identification of C. obtusa and Thujopsis spp. can help us better understand the use of wood in Japanese culture, and this can in turn help the historical and architec/tural sciences.

These species are similar not only in terms of mechanical properties, but also in their anatomies. Conventionally, they have been identified by observing only the types of cross-field pitting (Shimaji and Itoh 1982), which is widely recognized as the most important anatomical features for the identification of conifers among wood anatomists (Richter et al. 2004). Noshiro (2011), following extensive work, concluded that the types of cross-field pitting are piceoid and cupressoid in C. obtusa, and cupressoid and taxodioid in Thujopsis spp. However, identifying them using cross-field pitting remains challenging owing to the absence of a quantitative definition of each type of cross-field pitting. Thus, methods based on unambiguous criteria should be introduced for their identification.

The microfibril angle (MFA) is a reliable means of describing the characteristics of wood because this quantitative parameter can explain a number of important features, such as mechanical properties (Bendtsen and Senft

^{*}Corresponding author: Junji Sugiyama, Research Institute for Sustainable Humanosphere, Kyoto University, Uji, Kyoto, 611-0011, Japan; and College of Materials Science and Engineering, Nanjing Forestry University, Nanjing, China,

1986; Cave and Walker 1994; Cowdrey and Preston 1966) and shrinkage behaviors of wood (Abe and Yamamoto 2006; Barber and Meylan 1964; Meylan 1972). By contrast, features of the MFA specific to certain species of softwoods. such as MFA value itself and their transition from earlywood to latewood, remain undiscussed except for a few studies (Hori et al. 2002; Lichtenegger et al. 1999). This is because most of the conventional methods for MFA measurement suffer from their low spatial resolution and vulnerability to various kinds of variations in the MFA, inside and between the tracheids (Donaldson 2008). However, techniques of synchrotron X-ray (Lichtenegger et al. 1999), polarization (Abraham and Elbaum 2012; Mannan et al. 2016; Savić et al. 2016) and polarized Raman microscopy (Gielringer et al. 2010) have a potential to overcome these shortcomings and gives us many insights through the visualization of MFA spatial distributions in a manner of a two-dimensional (2D) image. These methods enable us to directly quantify many anatomical features, such as multi-layered and pit structures, and provide a new basis for analyzing species-specific features. Especially, plant anatomies based on 2D MFA images can be more accessible at the laboratory scale within a reasonable experimental time.

The purpose of the present research was to develop a technique for quantifying the anatomical features of wood using the MFA, and to apply it to identify the anatomically similar Japanese conifers *C. obtusa* and *Thujopsis* spp. Until now, this is the first attempt to apply the MFA to the identification of wood. MFA measurements were performed using a technique similar to the above-mentioned MFA imaging with polarization. A wood identification system was established based on a powerful image recognition technique, a convolutional neural network (CNN), to make the best use of 2D spatial distributions of the MFA. The performance of the model was evaluated through metrics obtained from the results of classification of present and old wood specimens collected from Manpuki-ji, Uji, Kyoto. Finally, the species-specific features of wood to discriminate between C. obtusa and Thujopsis spp. are discussed by examining the relationship between their anatomical features and the MFA.

2 Materials and methods

2.1 Sample information and preprocessing

Present and old specimens of *C. obtusa* and *Thujopsis* spp. were prepared (Table 1). All samples in this experiment were assigned the KYOw ID registered by the xylarium at the Research Institute for

Table 1: Present and old wood specimens, and information about their database.

Scientific name	State	Number	Sample number (KYOw)
Chamaecyparis	Present	12	00638, 04432, 04434,
obtusa			04435, 04436, 04439,
			08004, 08005, 08007,
			13832, 14105, 17824
	Old	2	20533, 20534
Thujopsis dolabrata	Present	10	00034, 00579, 00580,
			11401, 11402, 11403,
			13581, 20536, 20537,
			20538
Thujopsis dolabrata var. hondae	Present	2	08017, 08018
Thujopsis spp.	Old	2	20485, 20539

Sustainable Humanosphere at Kyoto University (https://database. rish.kyoto-u.ac.jp/cgi-bin/bmi/en/namazu.cgi). Small wood blocks in only the heartwood regions, distant from the pith, were selected for use as present wood specimens. The old wood specimens were collected from the Manpuku-ji temple built in the early Edo period (mid-17th century).

All wood blocks were cut into small pieces, approximately $4 \times 1.5 \times 5$ mm along the radial, tangential, and longitudinal directions, respectively. They were embedded into 100% PEG 1500. 10 µm-thick cross-sections of the embedded samples were cut using a sliding microtome and sealed by aqueous gum-chloral mounting media.

2.2 Acquisition of hyperspectral images

The sections were placed on the revolving stage of a polarized optical microscope (BX53-P, Olympus, Tokyo, Japan) equipped with a 40 × objective lens (UPLFLN40XP, NA=0.75, Olympus, Japan) and a full wave plate (U-TP530, λ = 530 nm, Olympus, Japan) for emphasizing interference color transitions between MFA values. Directions of a polarizer and an analyzer were set to a crossed Nicol condition. The radial and tangential walls in the earlywood zone were roughly parallel to the direction of high and low refractive index of the plate, generally called slow and fast axes, respectively. In our experiment, MFA imaging becomes possible via retardation values of cell walls under the crossed Nicol observed as interference colors. In this situation, the relationship between observed intensities of interference colors at the specific wavelength, $I(\lambda)$, and retardation values R can be expressed as below equation.

$$I(\lambda) = A^2 \sin^2 \theta_{\text{sample}} \sin^2 \frac{\pi R}{\lambda}$$
 (1)

A means an amplitude of an incident light; θ_{sample} is an angle of directions of a polarizer and a slow axis of a specimen; λ corresponds to a light wavelength.

From the above equation, an interference color is observed as a color spectrum under a normal illumination (not a laser). Therefore, we can extract retardation values from interference colors through the quantification of color spectra. For achieving this quantification, hyperspectral images (1536 × 1024 pixels) within the visible light region from 461 to 602 nm at intervals of 3 nm were taken by a CCD camera with a liquid crystal tunable filter, LCTF (VariSpecTM, Cambridge Research & Instrumentation, Inc, Massachusetts, USA). The exposure time at each wavelength was set to 3 s. Approximately 5-10 images per specimen labelled specific KYOw ID were taken and saved in 16 bit TIFF format. Following acquisition, the original hyperspectral images (1536×1024 pixels) were cut into two images (768×768 pixels) without any overlap. Finally, 500 sets of hyperspectral images were obtained in total (10–20 sets of hyperspectral images per specimen).

2.3 Conversion of hyperspectral images into twodimensional MFA images

The conversion procedure is shown in Figure 1. First, the spectra of the interference colors of the hyperspectral images were converted into retardation values pixel by pixel using the following equation after corrections of light intensity:

$$I(\lambda) = a \sin^2 \frac{\pi R_{\text{obs}}}{\lambda} + b \tag{2}$$

In the above, $I(\lambda)$ is the intensity of light at specific wavelengths quantized by the CCD camera with LCTF, such as 461, 464, ..., 601 nm; a and b are scaling parameters determined through the fitting procedure. Parameter a expresses an angle dependent term shown in the first half of Eq. (1). Parameter *b* is a term for correcting the imperfection of the polarizer, analyzer and CCD camera; $R_{\rm obs}$ is the net observed retardation values obtained by fitting.

All image pixels were fitted using Eq. (2) through a nonlinear least-squares fitting algorithm called the Levenberg-Marquardt method (Levenberg 1944; Marquardt 1963). 2D retardation images $(768 \times 768 \text{ pixels})$ were thus obtained. The below equations, that of an index ellipsoid (Figure 2) and the definition of the retardation were used to transform the observed retardation values into the MFA:

$$\frac{1}{n_e^2(\theta_{\rm MFA})} = \frac{\sin^2\theta_{\rm MFA}}{n_e^2} + \frac{\cos^2\theta_{\rm MFA}}{n_o^2}$$
 (3)

$$R_{\rm obs} = R_{\rm FWP} \pm R_{\rm cell}$$
 (4)

$$R_{\text{cell}} = d_{\text{net}} |n_{\text{e}} (\theta_{\text{MFA}}) - n_{\text{o}}| \tag{5}$$

In Eqs. (3)–(5), n_0 and n_e are the principle refractive indices of ordinary and extraordinary rays, respectively ($n_0 = 1.529$, $n_e = 1.599$; excerpt from Iyer et al. (1968)); $n_{\rm e}(\theta_{\rm MFA})$ is the net refractive index of extraordinary rays when the MFA is θ_{MFA} ; R_{FWP} and R_{cell} correspond to the retardation values of full wave plate and cell walls, respectively.

The parameter d_{net} represents the thickness of the section when cellulose content in the cell walls is considered, and is defined below:

$$d_{\text{net}} = c_{\text{cellulose}} d_{\text{section.}} \tag{6}$$

In the above, d_{section} is the thickness of section (10 µm); $c_{\text{cellulose}}$ represents the cellulose ratio in the cell walls, and ranges from zero to one.

The parameter d_{net} should be introduced to avoid overestimating the MFA values. Following Panshin et al. (1964), $c_{\text{cellulose}}$ was set to 0.5, which was the general value in the S_2 wall of the conifers. Using Eqs. (3)-(6), the formula for converting retardation to the MFA can be

$$\theta_{\text{MFA}} = \frac{180}{\pi} \times \arccos\left(\sqrt{\frac{n_{\text{o}}^{2}}{n_{\text{e}}^{2} - n_{\text{o}}^{2}} \times \left\{\frac{n_{\text{e}}^{2}}{(|R_{\text{obs}} - R_{\text{FWP}}|/d_{\text{net}} + n_{\text{o}})^{2}} - 1\right\}}\right)$$
(7)

Finally, 500 2D MFA images were obtained after the conversion using Eq. (7).

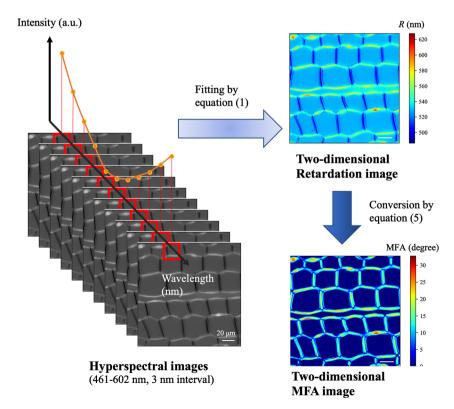


Figure 1: The experimental procedure for converting hyperspectral images into a twodimensional retardation and an MFA image.

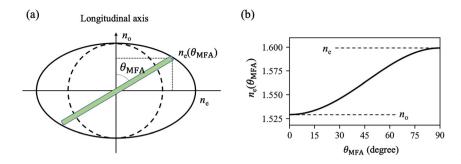


Figure 2: (a) An index ellipsoid of cellulose microfibrils in the cell walls, and (b) the relationship between the refractive index of extraordinary rays and the MFA.

2.4 Stratification of MFA images

Stratification processing is the conversion of one-channel 2D MFA images into three-channel images, such as in RGB. The details of this procedure are shown in Table 2 and Figure 3. First, three ranges of the MFA $-x_1 \le \theta_{MFA} \le x_2$, $x_3 \le \theta_{MFA} \le x_4$, and $x_5 \le \theta_{MFA} \le x_6$ —were determined within $0^{\circ} \le \theta_{MFA} \le 40^{\circ}$. Second, pixel values corresponding to the first range of the MFA were extracted and set as Layer 1. Irrelevant pixels $(0 \le \theta_{MFA} \le x_1 \text{ and } x_2 \le \theta_{MFA} \le 40)$ were set to zero in this channel. The same processing was also applied to the second and third ranges of the MFA, and Layers 2 and 3 were respectively created. For example, Layer 1-3 correspond to MFA ranges, 0°-10°, 10°-20°, 20°-40°, respectively in the case of the upper row of Table 2. Third, images of each channel were stacked as one three-channel image called the stratified MFA image in the following. Finally, three kinds of conditions for ranges of the MFA and three preprocessing methods (Preprocesses 1, 2, and 3; see Figure 3) were prepared. In Preprocess 1, all layers were divided by x_6 . In Preprocess 2, Layers 1, 2, and 3 were divided by x_2 , x_4 , and x_6 , respectively. In Preprocess 3, normalization (zero-to-one scaling) was applied after Preprocess 2. For example, all layers in Stratified 1 were divided by 40; Layer 1-3 in Stratified 2 were divided by 10, 20 and 40, respectively; Layer 1-3 in Stratified 3 were divided by 10, 20 and 40, respectively, and normalized to 0 to 1 range (See the bottom histograms in Figure 3).

When the original one-channel 2D MFA images were used, the same image was copied two times, and these images were stacked as a three-channel image because the number of input channels was fixed to three in the CNN used in this time. In total, 10 kinds of stratified

Table 2: List of MFA ranges of each layer and preprocess conditions in each stratified MFA image.

Name	Layer 1		Layer 2		Layer 3		Preprocess	
	X ₁	X ₂	<i>X</i> ₃	<i>X</i> ₄	X ₅	X ₆	condition	
Stratified 1	0	10	10	20	20	40	1	
Stratified 2	0	10	10	20	20	40	2	
Stratified 3	0	10	10	20	20	40	3	
Stratified 4	0	15	15	30	30	40	1	
Stratified 5	0	15	15	30	30	40	2	
Stratified 6	0	15	15	30	30	40	3	
Stratified 7	0	15	10	25	20	40	1	
Stratified 8	0	15	10	25	20	40	2	
Stratified 9	0	15	10	25	20	40	3	
Original	0	40	0	40	0	40	3	

(Stratified 1, Stratified 2, ..., Stratified 10) and three-channel original images were created (Table 2) and utilized as inputs of a CNN model.

2.5 Classification by CNN

The two Cupressaceae species were classified using the original and stratified MFA images in a CNN. A CNN model consists of mainly two sections, convolutional layers for feature extracting from images and fully connected layers for classification. Convolutional layers of our model reused a VGG16 model (Simonyan and Zisserman 2015) trained on the ILSVRC-2012 dataset (Russakovsky et al. 2015). Fully connected layers were two; the first had 64 nodes with the ReLU activation unit (Glorot et al. 2011) and the L2 regularizer (l = 0.0000001) for avoiding overfitting, and the second had two nodes for classification using the softmax function. In addition, two familiar methodologies, dropout (Srivastava et al. 2014) and batch normalization (Ioffe and Szegedy 2015) were applied to achieve better model generalization performance. The dropout layers (dropout rate = 0.2) was inserted between the global averaging polling layer (Min et al. 2014) and the first fully connected layer. Batch normalization layers were placed behind the first fully connected layer and between the two fully connected layers.

Nearly 80% of all images were randomly selected as the training set, which did not contain MFA images obtained from the old specimens. A total of 15% of the training images were used for model validation. During training, minibatch learning was performed which means that training images are splitted into many small batches and these batches are used as inputs until all batches are exhausted for model training. The minibatch size was set to 14 images in the experiment. The parameters of the model were updated by the ADAM algorithm (learning rate = 0.0001, and the other parameters were the default values) (Kingma and Ba 2015) to minimize the cross-entropy function in the validation dataset. This optimization was continued for 100 times (epochs). The model with the lowest cross-entropy value during the validation step was used to assess generalization performance by calculating test accuracy and cross-entropy. To consider biases originating from dataset splitting, five patterns of train-validate-test sets were randomly prepared. Five models were thus prepared in each image preprocessing condition, and the mean and standard deviation of all metrics obtained from each model were used to assess their generalization performance.

2.6 Interpretation of results

Two methods were used to assess the discriminating features of the two species identified by the CNN during parameter optimization. The first was the degree of decay in accuracy accompanied by

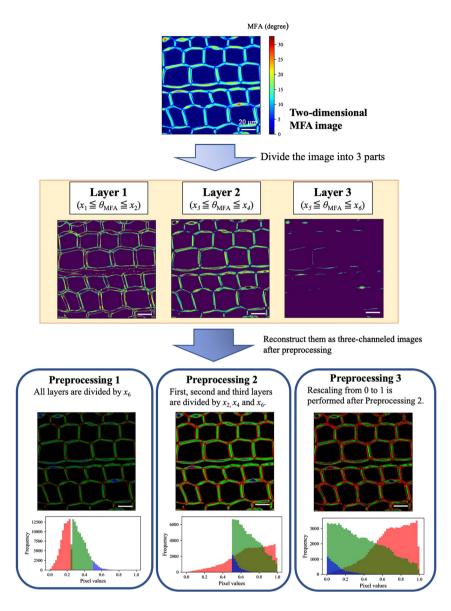


Figure 3: Experimental flow of the stratification preprocessing of 2D MFA images. First, an original 2D MFA map is divided into three images depending on the three ranges of the MFA, $x_1 \le \theta_{MFA} \le x_2$, $x_3 \le \theta_{\mathsf{MFA}} \le x_4$, and $x_5 \le \theta_{\mathsf{MFA}} \le x_6$. Second, these images are stacked as one threechannel image (stratified MFA image). Finally, the stratified image is preprocessed in one of three ways: Preprocess 1, Preprocess 2, and Preprocess 3.

channel-wise information erasure. The idea is that a large decay in accuracy is observed when a certain channel containing important discriminating features is erased. In cases of stratified MFA images, we could easily interpret the contribution of ranges of the MFA to the classification because each channel directly corresponded to a certain range. The second method was a reliable method of visual explanation for the CNN called the gradient-weighted class activation map (Grad-CAM) (Selvaraju et al. 2019). The mapping was calculated using the keras-vis library (https://github.com/raghakot/keras-vis). These techniques were applied to the three most accurate classification models in terms of scores on present and old wood specimens.

2.7 Raman imaging

Raman imaging was only applied to one T. dolabrata specimen (KYOw20538). Its cross-section at a thickness of 15 µm was sandwiched between a quartz coverslip and a microscope slide in wet condition, and was sealed by a manicure to avoid evaporation and spectral deterioration by photoluminescence (Tuschel, 2016). An XploRA confocal Raman microscope (HORIBA, Ltd., Kyoto, Japan) equipped with a 785 nm laser source (90-100 mW) and motorized XY mapping stage was used at 50 × magnification of the objective lens (MPLN 50 \times , NA = 0.75, Olympus, Japan). The spectral range was set to 250-1800 cm⁻¹, where almost all main bands of cellulose, lignin, and hemicellulose appeared (Agarwal 2014). The Raman spectra were acquired with a 10 s exposure time and four scans. Spectral mapping was conducted to approximate 12 \times 4.5 μ m in the X and Y directions in steps of $0.5 \, \mu m$ in each direction. The obtained spectra were preprocessed by the air-PLS method (Zhang et al. 2010) for baseline correction. Two band intensities, at 380 cm⁻¹ from crystalline cellulose and 1600 cm⁻¹ from lignin were used for heatmapping. Each peak intensity was corrected by subtracting the baseline intensities determined by two basepoints located at the periphery around the peak of interest. For example, the corrected intensity at 380 cm⁻¹ was obtained by subtracting baselines drawn by 360 cm⁻¹ and 400 cm⁻¹ from the raw peak intensity at 380 cm⁻¹ (Figure 4). In the same manner, the peak at 1600 cm⁻¹ was corrected by baselines

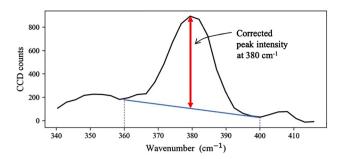


Figure 4: Peak intensity correction of Raman spectra at 380 cm $^{-1}$. The baseline is drawn connecting two basepoints, at 360 cm $^{-1}$ and 400 cm $^{-1}$.

connecting 1550 cm⁻¹ and 1700 cm⁻¹. Finally, Raman heatmap was created based on the intensity ratios of the band, 380 cm⁻¹/1600 cm⁻¹. Pixels whose intensities at 1096 cm⁻¹ from cellulose and polysaccharides were below 100 were considered the background regions and set to zero in all heatmaps.

3 Results and discussion

3.1 Close view of two-dimensional MFA image

Figure 5a shows a representative result of the 2D MFA images obtained from T. dolabrata, and Figures 5b–5f shows the 2D MFA distributions of different ranges of the MFA: 0° – 10° , 10° – 15° , 15° – 20° , 20° – 30° , and 30° – 40° . The abundant component in the tracheids, the S_2 wall, was assigned at approximately 0° – 15° in the MFA, as shown in Figures 5b and c. Gradual increments in the MFA were observed in the transitional zones between S_2 and S_1 , and S_2 and the bordered pits. A thin S_1 wall was designated to approximate

15°-30° in the MFA, as shown in Figures 5d and e. Finally, the bordered pit regions had the longest range of the MFA from 20°-40°, corresponding to Figures 5e and f.

3.2 Raman mapping

Raman heatmaps indicating the peak intensity ratios of $380 \text{ cm}^{-1}/1600 \text{ cm}^{-1}$ is shown in Figure 6. This Raman heatmap mirrored the qualitative proportions of crystalline cellulose versus lignin in the cell walls. As shown in Figure 6, the Raman mapping indicated that the crystalline cellulose ratio compared with lignin of S_1 was lower than that of S_2 in a qualitative manner. Moreover, these peak intensity ratios were nearly uniform in each section of the wall, and radical changes in them occurred only in the transitional zones between walls of S_1 and S_2 .

3.3 Classification results

The results of the classification of present and old wood specimens in the test datasets with and without various

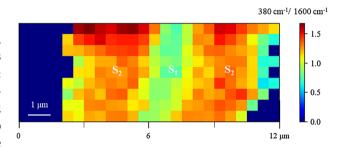
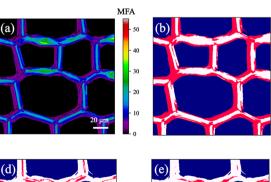
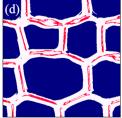
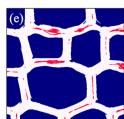
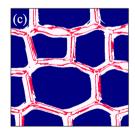


Figure 6: Raman heatmap based on peak intensity ratios of $380 \text{ cm}^{-1}/1600 \text{ cm}^{-1}$.









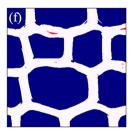


Figure 5: (a) An expanded 2D MFA image of *Thujopsis dolabrata* (KYOw20538) and (b-f) 2D MFA images specifying certain ranges of the MFA highlighted by the red pixels. Each panel corresponds to (b) 0°–10°, (c) 10°–15°, (d) 15°–20°, (e) 20°–30°, and (f) 30°–40°.

Table 3: List of classification results.

Name		Cross-validation		New wood	Old wood	
	Accuracy	Log loss	Accuracy	Log loss	Accuracy	Log loss
Stratified 1	0.963 (0.015)	0.095 (0.024)	0.823 (0.075)	0.341 (0.142)	0.955 (0.024)	0.143 (0.026)
Stratified 2	0.977 (0.033)	0.079 (0.030)	0.878 (0.031)	0.311 (0.107)	0.895 (0.037)	0.223 (0.070)
Stratified 3	0.977 (0.012)	0.082 (0.024)	0.802 (0.146)	0.475 (0.377)	0.960 (0.020)	0.156 (0.036)
Stratified 4	0.957 (0.026)	0.122 (0.020)	0.843 (0.031)	0.364 (0.039)	0.860 (0.041)	0.254 (0.035)
Stratified 5	0.974 (0.028)	0.100 (0.039)	0.837 (0.043)	0.315 (0.102)	0.859 (0.035)	0.317 (0.106)
Stratified 6	0.977 (0.012)	0.092 (0.034)	0.884 (0.060)	0.268 (0.104)	0.905 (0.037)	0.247 (0.070)
Stratified 7	0.971 (0.009)	0.095 (0.019)	0.847 (0.093)	0.323 (0.131)	0.940 (0.034)	0.198 (0.062)
Stratified 8	0.966 (0.037)	0.075 (0.044)	0.874 (0.053)	0.307 (0.123)	0.880 (0.048)	0.294 (0.110)
Stratified 9	0.969 (0.019)	0.093 (0.019)	0.844 (0.047)	0.380 (0.139)	0.840 (0.064)	0.450 (0.157)
Original	0.968 (0.017)	0.094 (0.023)	0.774 (0.088)	0.506 (0.233)	0.800 (0.052)	0.450 (0.091)

The numbers with and without parentheses indicate the standard deviation and the mean value, respectively, of each metric.

stratification and normalization conditions are listed in Table 3. An accuracy of nearly 90% was achieved in the classification of specimens of present wood under stratification conditions 2, 6, and 8. The accuracy values correspond to improvements in accuracy by approximately 10% in comparison with classification based on the original 2D MFA images. Moreover, the combination of stratification preprocessing and the CNN yielded an accuracy of over 90% in the classification of old specimens of wood. The results of only stratification conditions 2, 6, and 8 are subjected to interpretation below.

3.4 Interpretation based on two methods

The decays in the accuracies of Stratified 2, 6, and 8 when channel-wise erasure was implemented in each channel are listed in Table 4. As a general tendency, small decays in accuracy, below 10%, occurred mainly in Layer 3 corresponding to regions of high MFA around 20°-40°. Conversely, large decays in accuracy occurred when smallor middle-sized regions of the MFA were eliminated. In Stratified 2 (10°-20° in the MFA), accuracy decreased by 0.298 and 0.368 from the original test accuracy, 0.878, in Layers 1 and 2, respectively. In the same manner, Layer 1 of Stratified 6 (0°-15° in the MFA) also showed a large deterioration in accuracy, by 0.359 from 0.884.

Figures 7a and 7b are representative Grad-CAM heatmaps obtained from C. obtusa and Thujopsis spp, respectively. These heatmaps show areas important for the decision making by CNN in each image. For example, we can qualitatively access characteristics of the species in the image of C. obtusa using this method. Figure 7 shows distinct differences between them at positions where the CNN focused: tangential walls in C. obtusa, and radial walls

Table 4: List of decays in accuracy through channel-wise erasure.

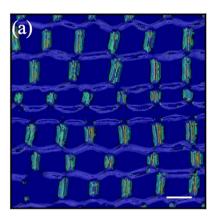
Name			Accuracy decay
	Layer 1	Layer 2	Layer 3
Stratified 2	-0.298 (0.118)	-0.368 (0.034)	-0.030 (0.031)
Stratified 6	-0.359 (0.052)	-0.177 (0.092)	0.000 (0.060)
Stratified 8	-0.091 (0.063)	-0.216 (0.130)	-0.112 (0.088)

The numbers with and without parentheses indicate the standard deviation and the mean value, respectively, of each metric.

in Thujopsis spp. According to these heatmaps, the anisotropy of the MFA in transverse distributions depended on the existence of each species. Figure 8 shows the mean MFA distributions of the radial and tangential walls of *C. obtusa* and *Thujopsis* spp. The anisotropy of the MFA was observable in each species although large natural variations covered all MFA ranges. As a general trend, differences in the MFA appeared around regions of low, middle, and high MFA between them. In particular, the largest gap occurred at their peak top positions located within 0°-20° in each direction, where this is consistent with the results of the decays in accuracy.

3.5 Description of anatomical features via **MFA**

The tracheid cell walls of the coniferous species consisted of multiple layers that are called the primary and the secondary walls $(S_1, S_2, \text{ and } S_3)$. In addition, the tracheid cells had circular pit structures penetrating each wall for the transportation of water and chemicals. Figures 5b-5f shows the relationship between the anatomical structures and the MFA of S₁ and S₂, and their bordered pits. The MFA



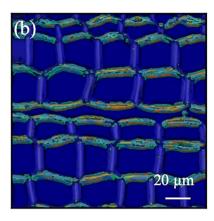


Figure 7: Representative Grad-CAM heatmaps of (a) Chamaecyparis obtusa (KYOw04435) and (b) Thujopsis dolabrata (KYOw11402). Orange and blueish parts correspond to relevant and irrelevant areas for decision makings by the CNN model, respectively. These maps were obtained from the first convolutional layer in the third block of VGG16.

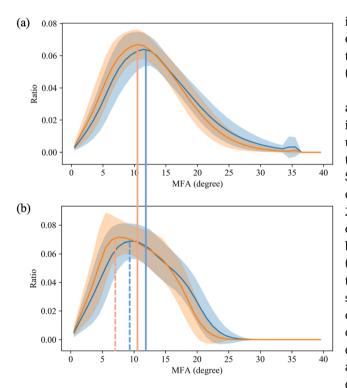


Figure 8: Mean MFA distributions from present wood specimens of the sides of the (a) radial and (b) tangential walls obtained from 2D MFA maps of Chamaecyparis obtusa (blue) and Thujopsis spp. (orange) present wood specimens. Blue and Orange solid and broken lines correspond to peak top positions of MFA distributions of each walls in each wood species, respectively. Blue and orange bands indicate the standard deviation ranges.

of the S₂ wall was affected by various factors (Donaldson 1992: Donaldson 1997), but generally had smaller values than the other layers (Donaldson 2008). In case of Thujopsis spp., Hori et al. (2002) reported that the average MFA of T. dolabrata was 9.2° in earlywood zones as measured by X-ray diffraction, where this value agreed well with that obtained by our experiment (Figures 5b and 5c; Figure 8b) in the prefixed $c_{\rm cellulose}$ condition. The MFA of $C.\ obtusa$ in earlywood fluctuated around 10° (Ohta et al., 1968), and this reported value coincided well that obtained here (Figure 8a).

There are three important points for evaluating anatomical features from 2D MFA images. First, the MFA images clearly show the features of cell walls like nonuniformities within the MFA of the S2 wall in the transitional zone and pits structure. The MFA transition between S₂ and the other structures has also been reported in other coniferous species (Abe et al. 1991; Donaldson and Xu 2005). About the latter, the net largest MFA shown in pits corresponded to large circumventions in the cellulose bundles occurring in the periphery of the pit structures (Hirakawa and Ishida 1981). These results indicated that the proposed methodology was valid for uncovering the spatial arrangement of the MFA at a high resolution. Second, the MFA of the S₁ wall was largely underestimated compared with generally reported MFA values, mainly by direct observations-50°-90° (Abe et al. 1991; Donaldson and Xu 2005; Panshin et al. 1964). This kind of MFA underestimation of S₁ wall is reported in the method using polarized Raman microscopy (Gierlinger et al. 2010). The MFA underestimation of S₁ wall in microscopic techniques partly comes from their net resolution. However, the primary reason in polarization techniques including this study was that the calculated MFA values were significantly affected by the cellulose content and crystallinity in each cell wall (Abraham and Elbaum 2012). Therefore, the over- or underestimation of the MFA was anticipated by discrepancies in cellulose characteristics between the S₁ and the S_2 walls because $c_{\text{cellulose}}$ was fixed to 0.5, referring to the general value in the S2 wall of the coniferous tracheids (Panshin et al. 1964). The differences in cellulosic components between S₁ and S₂ are supported by the Raman heatmap shown in Figure 6, and other reports (Agarwal 2006; Müller et al. 2002; Panshin et al. 1964). Finally, MFA

values estimated by this methodology is influenced by experimental imperfections. Generally, ideal section thickness set by microtome and actual thickness have some discrepancy. Hence, this experimental imperfection should be considered as shown in Eq. (6). For example, MFA errors calculated by Eqs. (2) and (7) are 8% at most in each MFA value from 0° to 50° if section thickness errors are within 10% (9–11 µm). These errors gradually increase from lower to higher MFA regions. In addition, cellulose microfibril orientation is inevitably disturbed to some degree by sectioning stress (Gierlinger et al. 2010; Konnerth et al. 2009). This kind of disturbance was locally observed in our experimental results (tangential walls in Figure 5). To avoid these effects, sample preparation should be performed as carefully as possible. In our experiment, these imperfections didn't have a critical influence on the classification by CNN listed in Table 3 because CNN recognized randomly occurred imperfections as irrelevant features for the classification.

From the above results, the methodology utilizing polarization is the simple and convenient technique to assess the MFA distribution of cell walls with high resolution. Especially, this technique has a great advantage in analyzing multi-layered structure and heterogeneity which have been widely observed in living organisms such as plants (Abraham and Elbaum 2012; Niklas 1992) or animals (for example, in bones. See Atesok et al. 2016). They serve as a crucial factor for these organisms in adapting to surrounding environment. 2D MFA visualization has the potential to shed light on novel ways of biomechanical adaptation.

3.6 Robust classification based on MFA and CNN

A combination of 2D MFA images and the CNN provides us with a reliable technique for distinguishing between anatomically similar Cupressaceae species regardless of the effects of aging for up to 400 years of cutting. This successful result was obtained for three reasons. The first is the maturity of the wood specimens. Biomechanically, there was a large discrepancy in terms of MFA values between juvenile and mature specimens. Juvenile parts close to the pith generally had larger MFA values, ruled by trends of growth irrelevant to any wood species, because a higher MFA enabled them to endure wind loads from the lateral sides (Lichtenegger et al. 1999). However, the mature part was characterized by relatively small and stable MFA values, which are preferable for extracting the speciesspecific features of wood. This led to a slightly better

classification of mature old specimens than present ones. The second reason for the positive result is stratification preprocessing, as it an improvement over low-contrast 2D MFA images by creating artificial cliffs of pixel values. Moreover, it was assumed that different information regarding the MFA in each channel facilitated feature extraction by the CNN, even from similar anatomical structures of C. obtusa and Thujopsis spp.—those that humans cannot recognize intuitively. This is identical to when an RGB image shows better performance than a grayscale image in terms of feature detection (Adbel-Hakim and Farag 2006). Worth mentioning as well is the robustness of the crystalline cellulose to the process of aging of the wood. Regions of crystalline cellulose did not significantly deteriorate for wood samples cut more than 1,300 years ago, even though large degradations had occurred in the other components (Tsuchikawa et al. 2005). As a result, our classification framework based on only crystalline cellulose performed well, whereas methods based on other chemical components, such as nearinfrared spectroscopy, failed in the classification of old woods (Horikawa et al. 2015).

3.7 Wood species specific MFA anisotropy

Results in terms of accuracy decay and Grad-CAM implied that the feature discriminating between C. obtusa and Thujopsis spp. comes from MFA anisotropy of the S₂ wall in the transverse plane. The origin of this species-dependent anisotropy remains unclear only from 2D MFA maps and classification results because MFA distributions in cell walls encode various kinds of anatomical features (tracehid length, cell wall thickness and so on) affected by surrounding environment which cannot be deciphered by a simple way (Donaldson, 2008). However, a part of this problem can be explained by the existence of pits in the earlywood zone (Bailey and Vestal 1937). In general, the bordered pits were concentrated only along the radial walls in the earlywood (Ilvessalo-Pfäffli 1995), and the net MFA values on the radial side became larger than those of the tangential walls owing to the presence of the pits (Sarén et al. 2004). The degree of MFA disturbance in S2 wall was naturally determined by the anatomical features of the pits, including their slit shape, size, arrangement (uniseriate or more) (Richter et al. 2004), and frequency. Considering the results obtained from the two explainers, our methodology suggests that low and middle MFA regions affected by anatomical features including pits geometry work as the discriminative features of the two anatomically similar coniferous species. This hypothesis was partly supported

by the general trend of mean MFA distributions of each species as the positions of peak intensity maxima of each walls (Figure 8).

4 Conclusion

A combination of 2D MFA images with stratification preprocessing and the CNN led to an accuracy of nearly 90% in the classification of present and old wood specimens of two anatomically similar species, C. obtusa and Thujopsis spp. The results here verified the good generalization performance of the proposed method, unless crystalline parts of cellulose microfibrils were heavily degraded. The results in terms of the decay in accuracy and Grad-CAM suggested that distinctive species-specific features of wood based on the MFA appeared in its anisotropy in regions of the S₂ wall in transverse planes. The existence of pits and their characteristics might have led to this anisotropy, and could be indirectly detected in the form of the MFA modulation induced by them.

This work first showed the possibility of the MFA for identifying wood using an image recognition technique. Technically and theoretically, this technique can be further simplified, with a combination of RGB color images of the interference colors and the CNN. This framework is applicable to any other species of wood and has a potential to give us many insights for reviewing wood anatomies. In addition, the MFA information in a 2D manner provides us a sound basis for accelerating researches on the relationship between MFA and cell geometry, and the anisotropic behavior of wood like mechanics, shrinkage from a new perspective.

Acknowledgments: We are indebted to Prof. Katsuhiko Takata (Akita Prefectural University) for providing some of the T. dolabrata specimens, and Hironobu Takeshita (Cultural Properties Division, Kyoto Prefectural Board of Education) for offering old wood specimens collected from Manpuku-ji. We thank Saad Anis, PhD, from Edanz Group (https://en-author-services.edanzgroup.com/) for editing a draft of this manuscript.

Author contributions: YK performed the experiments and analyses, and was a major contributor to the writing of the manuscript. JS designed and supervised the work. All authors read and approved the final manuscript.

Research funding: This study was supported by Grantsin-Aid for Scientific Research (grant numbers 17F17402, 18H05485) from the Japan Society for the Promotion of Science to JS, and RISH Cooperative Research (2019ZAIKAN-06) and RISH Mission Research V to YK and IS.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

- Abe, K., and Yamamoto, H. (2006). Behavior of the cellulose microfibril in shrinking woods. J. Wood Sci. 52: 15-19.
- Abe, H., Ohtani, J., and Fukazawa, K. (1991). FE-SEM observations on the microfibrillar orientation in the secondary wall of tracheids. IAWA (Int. Assoc. Wood Anat.) Bull. 12: 431-438.
- Abraham, Y., and Elbaum, R. (2012). Quantification of microfibril angle in secondary cell walls at subcellular resolution by means of polarized light microscopy. New Phytol. 197: 1012-1019.
- Adbel-Hakim, A.E. and Farag, A.A. (2006). CSIFT: A SIFT descriptor with color invariant characteristics. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition, 17-22 June 2006. 2. IEEE, New York, USA, pp. 1978-1983. https:// en.wikipedia.org/wiki/Conference_on_Computer_Vision_and_ Pattern_Recognition.
- Agarwal, U.P. (2006). Raman imaging to investigate ultrastructure and composition of plant cell walls: distribution of lignin and cellulose in black spruce wood (Picea mariana). Planta 224: 1141-1153.
- Agarwal, U.P. (2014). 1064 nm FT-Raman spectroscopy for investigations of plant cell walls and other biomass materials. Front. Plant Sci. 5, https://doi.org/10.3389/fpls. 2014.00490.
- Atesok, K., Doral, M.N., Karlson, J., Egol, K.A., Jazrawi, L.M., Coelho, P.G., Martinez, A., Matsumoto, T., Owens, B.D., and Ochi, M., et al. (2016). Multilayered scaffolds in orthopaedic tissue engineering. Knee Surg. Sports Traumatol. Arthrosc. 24: 2365-2373.
- Bailey, I.W., and Vestal, M.R. (1937). The orientation of cellulose in the secondary wall of tracheary cells. J. Arnord Arbor 18: 185-195.
- Barber, N.F., and Meylan, B.A. (1964). The anisotropic shrinkage of wood. A theoretical model. Holzforschung 18: 146-156.
- Bendtsen, B.A. and Senft, J. (1986). Mechanical growth and anatomical properties in individual growth rings of plantationgrown eastern cottonwood and loblolly pine. Wood Fiber Sci. 1: 23-38.
- Cave, I.D. and Walker, J.C.F. (1994). Stiffness of wood in fast-grown plantation softwoods: the influence of microfibril angle. For. Prod. I 44: 43-48.
- Cowdrey, D.R. and Preston, R.D. (1966). Elasticity and microfibrillar angle in the wood of Sitka spruce. Proceedings of the Royal Society B 166: 245-272.
- Donaldson, L.A. (1992). Within- and between-tree variation in microfibril angle in Pinus radiata. N. Z. J. For. 22: 77-86.
- Donaldson, L.A. (1997). Between-tracheid variation in microfibril angles in radiata pine. In: Proceedings of international workshop on microfibril angle, Westport, New Zealand, November. Canterbury University Press, Christchurch, NZ, pp. 206–224, https://www.worldcat.org/title/microfibril-angle-in-wood-theproceedings-of-theiawaiufro-international-workshop-on-thesignificance-of-microfibril-angle-to-woodquality-westport-new-

- zealand-november-1997/oclc/248096262/editions? referer=di&editionsView=true.
- Donaldson, L., and Xu, P. (2005). Microfibril orientation across the secondary cell wall of Radiata pine tracheids. Trees (Berl.) 19: 644-653.
- Donaldson, L. (2008). Microfibril angle: Measurement, variation and relationships-A review. IAWA J. 29: 345-386.
- Gielringer, N., Luss, S., König, C., Konnerth, J., Eder, M., and Fratzl, P. (2010). Cellulose microfibril orientation of Picea abies and its variability at the micron-level determined by Raman imaging. J. Exp. Bot. 61: 587-595.
- Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In: Gordon, G., Dunson, D., and Dudik, M. (Eds.), Proceedings of the fourteenth international conference on artifici al intelligence and statistics, Ft Lauderdale, FL, USA, 11-13 April 2011. 15. JMLR, Cambridge MA, pp. 315-323, https://portal. issn.org/resource/ISSN/2640-3498#.
- Hirai, S. (1996). Ki no Daihyakka (Encyclopedia of wood). Tokyo, Japan: Asakura Publishing.
- Hirakawa, Y. and Ishida, S. (1981). A scanning and transmission microscopic study of layered structure of wall in pit border region between earlywood tracheids in conifer. Res. Bull. Hokkaido Univ. For. 38: 249-263.
- Hori, R., Müller, M., Watanabe, U., Lichtenegger, H.C., Fratzl, P., and Sugiyama, J. (2002). The importance of seasonal differences in the cellulose microfibril angle in softwoods in determining acoustic properties. J. Mater. Sci. 37: 4279-4284.
- Horikawa, Y., Mizuno-Tazuru, S., and Sugiyama, J. (2015). Nearinfrared spectroscopy as a potential method for identification of anatomically similar Japanese diploxylons. J. Wood Sci. 61: 251-261.
- Ido, Y. (2018). History of sawn lumber standards and allowable stresses in Japan. Bull. For. For. Prod. Res. Inst. 18: 1-33.
- Ilvessalo-Pfäffli, M-S. (1995). Fiber atlas. Heidelberg, Germany: Springer.
- loffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd international conference on machine learning, 37. JMLR, Cambridge MA, pp. 448-456, https://portal. issn.org/resource/ISSN/2640-3498#.
- lyer, K.R.K., Neelakantan, P., and Radhakrishnan, T. (1968). Birefringence of native cellulose fibers. I. Untreated cotton and ramie. J. Polymer Sci. 2 Polymer Phys. 6: 1747-1758.
- Kingma, D.P. and Ba, L.J. (2015). Adam: A method for stochastic optimization. In: Paper represented at the 3rd international conference on learning representations, San Diego, CA, USA, 7–9 May 2015. International Conference on Learning Representation (ICLR), https://research.monash.edu/en/publications/learningfactorized-representationsfor-open-set-domain-adaptatio.
- Konnerth, J., Gielringer, N., Keckes, J., and Gindl, W. (2009). Actual versus apparent within cell wall variability of nanoindentation results from wood cell walls related to cellulose microfibril angle. J. Mater. Sci. 44: 4399-4406.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. 2: 164-168.
- Lichtenegger, H., Reiterer, A., Stanzl-Tschegg, S.E., and Fratzl, P. (1999). Variation of cellulose microfibril angles in softwoods and

- hardwoods-A possible strategy of mechanical optimization. J. Struct. Biol. 128: 257-269.
- Mannan, S., Zaffar, M., Pradhan, A., and Basu, S. (2016). Measurement of microfibril angles in bamboo using Mueller matrix imaging. Appl. Optic. 55: 8971-8978.
- Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Ind. Appl. Math. 11: 431-441.
- Meylan, B.A. (1972). The influence of microfibril angle on the longitudinal shrinkage-moisture content relationship. Wood Sci. Technol. 6: 293-301.
- Min, L., Chen, Q., and Yan, S. (2014). Network in network. In: Paper presented at the 2nd International Conference on Learning Representations, Banff, AB, Canada, 14-16 April 2014. International Conference on Learning Representation (ICLR).
- Müller, M., Hori, R., Itoh, T., and Sugiyama, J. (2002). X-ray microbeam and electron diffraction experiments on developing xylem cell walls. Biomacromolecules 3: 182-186.
- Niklas, K.J (1992). Plant biomechanics: An engineering approach to plant form and function. Chicago, USA: University of Chicago Press.
- Noshiro, S. (2011). Identification of Japanese species of Cupressaceae from wood structure. Jpn. J. Histor. Bot. 19: 125-132.
- Ohta, S., Watanabe, H., Matsumoto, T., and Tsutsumi, J. (1968). Studies on mechanical properties of juvenile wood. II Variation of fundamental structural factors and mechanical properties of hinoki trees (Chamaecyparis obtusa Sieb. et Zucc.). Mokuzai Gakkaishi 14: 261-268.
- Panshin, A.J., De Zeeuw, C., and Brown, H.P. (1964). Textbook of wood technology Volume I-Structure, identification, uses, and properties of the commercial woods of the United States, 2nd ed. New York: McGraw-Hill.
- Richter, H.G., Grosser, D., Heinz, I., and Gasson, P.E. (2004). IAWA list of microscopic features for softwood identification. IAWA J.
- Russakovsky, O., Deng, J., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., and Berg, A.C., et al. (2015). ImageNet large scale visual recognition challenges. Int. J. Comput. Vis. 115: 211-252.
- Sarén, M-P., Serimaa, R., Anderson, S., Saranpaa, P., Kecks, J., and Fratzl, P. (2004). Effect of growth rate on mean microfibril angle and cross-sectional shape of tracheids of Norway spruce. Trees (Berl.) 18: 354-362.
- Savić, A., Mitrović, A., Donaldson, L., Simonović Radosavljević, J., Bogdanović Pristov, J., Steinbach, G., Garab, G., and Radotić, K. (2016). Fluorescence detected linear dichroism of wood cell walls in juvenile Serbian spruce: estimation of compression wood severity. Microsc. Microanal. 22: 361-367.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2019). Grad-CAM: Visual explanations from deep networks via gradient-based localization. Int. J. Comput. Vis. 128: 336-359.
- Shimaji, K., and Itoh, T. (1982). Zusetsu Mokuzai-Soshiki (Illustrated wood anatomy). Tokyo: Chikyusha.
- Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In: Paper presented at the 3rd International Conference on Learning Representations,

- San Diego, CA, USA, 7-9 May 2015. International Conference on Learning Representation (ICLR).
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15: 1929-1958.
- Tsuchikawa, S., Yonenobu, H., and Siesler, H.W. (2005). Near-Infrared spectroscopic observation of the ageing process in
- archaeological wood using a deuterium exchange method. Analyst 130: 379-384.
- Tuschel, D. (2016). Selecting an excitation wavelength for Raman spectroscopy. Spectroscopy 4: 11-17.
- Zhang, Z.M., Chen, S., and Liang, Y.Z. (2010). Baseline correction using adaptive iteratively reweighted penalized least squares. Analyst 135: 1138-1146.