Convolutional Neural Network Performance and the Factors Affecting Performance for Classification of Seven *Quercus* Species using Sclereid Characteristics in the Bark

Jong Ho Kim,[a] Byantara Darsan Purusatama,[b] Alvin Muhammad Savero,[a] Denni Prasetia,[a] Jae Hyuk Jang,[c] Se Yeong Park,[a] Seung Hwan Lee,[a] and Nam Hun Kim [a,]*

Based on the sclereids in the bark of oak species, a convolutional neural network (CNN) was employed to validate species classification performance and its influencing factors. Three optimizers including stochastic gradient descent (SGD), adaptive moment estimation (Adam), root mean square propagation (RMSProp), and dataset augmentation were adopted. The accuracy and loss stabilized at approximately 15 to 20 and 70 to 80 epochs for the augmented and non-augmented condition, respectively. In the last five epochs, the RMSProp-augmented condition achieved the highest accuracy of 89.8%, whereas the Adam-augmented condition achieved the lowest accuracy of 73.8%. Regarding the loss, SGD-non-augmented condition was the lowest at 0.498, whereas Adam-augmented condition was the highest at 2.740. The highest accuracy was influenced by RMSProp at 0.194. Dataset augmentation had a significant influence on accuracy at 0.456. Homogeneous subsets among the validation conditions indicated that the accuracy and loss were classified into the same subset using an augmented dataset during the training, regardless of the optimizer. Only Adam and RMSProp with non-augmented datasets were categorized into the same subset during the test. Hence, species classification using CNN and sclereid characteristics in the bark was feasible, and RMSProp with augmented datasets showed optimal performance for species classification.

DOI: 10.15376/biores.19.1.510-524

Keywords: Bark; Convolutional neural networks (CNNs); Oak; Sclereids; Species classification

*Contact information: a: Department of Forest Biomaterials Engineering, College of Forest and Environmental Sciences, Kangwon National University, Chuncheon 24341, Republic of Korea; b: Institute of Forest Science, Kangwon National University, Chuncheon 24341, Republic of Korea; c: FC Korea Land Co., Ltd., Seoul 07271, Republic of Korea; * Corresponding author: kimnh@kangwon.ac.kr*

INTRODUCTION

Globally, there is excess demand for wood (UNECE 2007), resulting in a gradual appreciation of the value of wood as a resource in various sectors. Species identification is increasingly recognized as a significant process for enhancing value of wood resources and is required in numerous fields such as optimizing the utilization of conventional wood resources, protecting endangered species, and facilitating practical customs clearance operations.

Recently, there has been a surge in research aimed at automating the classification of wood species and enhancing their precision of automated classification. These efforts have been driven by the objectives of streamlining the process and reducing subjectivity. Ilic (1993) proposed the potential for computer-based automated species identification, while Wheeler and Baas (1998) emphasized the importance of computer-aid specification identification. Since the 2000s, studies applying computer vision to identify wood species have been conducted in earnest. Tou *et al*. (2007) proposed a system to identify wood species in real time using macroscopic images of wood cross sections. Bremananth *et al*. (2009) attempted to computerize a wood species recognition system using computer vision technology. In the 2010s, research on artificial intelligence-based wood species identification increased explosively, owing to the rapid development of computing performance and machine learning technology. As a representative example, Hermanson and Wiedenhoeft (2011) introduced an overview and advantages of species identification using machine vision in their review paper. Since the first meeting of the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) in 2010, many high-performance convolutional neural network (CNN) models, such as VGGNet (Simonyan and Zisserman 2015), GoogLeNet (Szegedy *et al*. 2015), and ResNet (He *et al*. 2016) have been developed. These inventions have led to rapid developments in image classification using CNN.

Recently, CNNs have been employed to classify wood species using datasets containing wood images. Kwon *et al*. (2017) reported the possibility and performance of softwood species classification using CNN-based models, such as LeNet and miniVGGNet. Kwon *et al*. (2019) improved the performance of automated species classification by using a CNN with ensemble methods. Yang *et al*. (2019) also employed a CNN with ensemble methods for classifying Korean softwood species using datasets from near-infrared spectra (NIR) analysis results and macroscopic images of radial sections. Hwang *et al*. (2020) attempted to visualize anatomical features using machine learning technology for quantitative analysis. Zhao *et al*. (2021) demonstrated the efficacy of a CNN for accurately classifying various wood species based on microscopic image datasets. Huang *et al*. (2021) demonstrated the efficacy of a transfer-learning model with enhanced pooling layers for enhanced wood identification using cross-sectional wood images. Hwang and Sugiyama (2021) conducted a methodological examination of computer-vision-based species identification. Their research anticipated the potential of computer vision to enhance the accessibility of wood identification to the public and make significant contributions to the field of wood science. Cao *et al*. (2022) utilized an artificial neural network trained using species-specific thermal conductivity trends to classify different species.

Most studies on wood species identification have focused on the xylem anatomical characteristics (Kim *et al*. 2021; Savero *et al*. 2022; Savero *et al*. 2023). However, bark also displays distinct characteristics in each species, rendering it a helpful tool for species identification (IAWA Committee 2016). Species identification using bark has a notable benefit because bark can be obtained from standing trees without timber harvesting. A few studies (Fiel and Sablatnig 2010; Bertrand *et al*. 2017; Carpentier *et al*. 2018; Kim *et al*. 2022) have attempted to utilize bark for automated species identification using artificial intelligence. They focused only on the outer appearance of the bark from standing trees for dataset preparation. However, the anatomical features of the bark could be more systematic than the surface features, showing a higher efficiency in feature selection and extraction, which affects species classification performance using computer vision. Sclereids are sclerenchyma cells that are variable in form and size, but typically not much elongated, with thick, often polylamellate, lignified secondary walls with many pits. They develop mainly in the nonconducting phloem, cortex, and periderm by modification of parenchyma cells. However, there are some examples of earlier development directly from cambial derivatives. Sclereids are extremely variable in size and shape, and this diversity has been classified into different sclereid types for the plant body, including brachysclereids, columnar sclereids, osteosclereids, astrosclereids, and filiform sclereids. Thus, sclereids can be used as keys for species differentiation (IAWA committee 2016).

Therefore, in the present study, the performance and performance-influencing factors of CNNs using sclereid characteristics in the bark to identify seven oak species were investigated. Three optimizers, stochastic gradient descent (SGD), adaptive moment estimation (Adam), and root mean square propagation (RMSProp), were used to analyze the feasibility and efficacy of the classification performance of the wood species.

EXPERIMENTAL

Materials

The barks of six domestic oak species obtained from the research forest of Kangwon National University, and *Quercus suber*, donated by FC Korea Land Co., Ltd. (Seoul, Korea), were used in this study. Three stems of each of the six domestic oak species were harvested, and bark samples were collected from the breast height of the stems. Several *Q. suber* bark samples with dimensions of 60 × 100 cm were also used. Comprehensive information about the samples is presented in Table 1.

Table 1. Sample Information

Scientific Name	D.B.H. (cm)*	Origin
Quercus dentata Thunb.	22.7 ± 1.8	Research Forest of Kangwon National University (Chuncheon, Korea N37.7748857, E127.8134654)
Quercus serrata Murray	27.7 ± 3.7	
Quercus mongolica Fisch. ex Ledeb.	25.4 ± 2.3	
Quercus variabilis Blume	25.5 ± 3.1	
Quercus aliena Blume	20.7 ± 4.9	
Quercus acutissima Carruth.	22.1 ± 6.2	
Quercus suber L.	** Planks of 60 × 100 cm	Portugal cork provided by FC Korea Land Co., Ltd. (Seoul, Korea)

*D.B.H.: Diameter at breast height

Methods

Sample preparation for the dataset

Wood discs (3 cm thick) were prepared using a chainsaw from the breast height of each oak trees. The barks were carefully detached from the discs. After separating the bark from the disc, a table saw (Professional Cabinet Saw model with 100-teeth saw blade, SawStop, Oregon, USA) was used to cut the top and bottom of the bark specimen to a uniform thickness of about 1 cm for easier microscopic observation. Transverse sections were sanded using a series of progressively finer sandpapers: #80, #120, #220, and #400. After the sanding treatment, the surface was cleaned using an air compressor. Twenty specimens of each species were examined using a visual microscope (MM-40, Nikon, Tokyo, Japan) equipped with a 2.5X objective lens (CF Plan 2.5X, Nikon, Tokyo, Japan).

Digital micrographs of the inner phloem of the bark specimens were obtained using a microscope camera (IMTCam; IMT, British Columbia, Canada) and subsequently used as a dataset. The collection contained 1,400 (200 for each species) microscopic images.

General macroscopic features

Sclereid characteristics, which are representative macroscopic anatomical features of the seven oak species, were analyzed using micrographs obtained to prepare a dataset. Five micrographs of each species were analyzed to determine the quantitative characteristics of the sclereids. The number of sclereids within one micrograph (3.75 × 2.50 mm), the area of individual sclereids, and the minimum and maximum area of sclereids for each species were measured.

Dataset pretreatment

The training dataset was augmented to examine the influence of the dataset size on the training procedure for the neural network architecture. Several parameters were applied to quantitatively pretreat and augment the dataset images. The parameters included rescaling the data using a ratio of 1/255 for normalization, 10° rotation, shifting the data horizontally and vertically by 10%, zooming the data by 20%, and performing horizontal and vertical flipping. Figure 1 illustrates the examples of data augmentation.

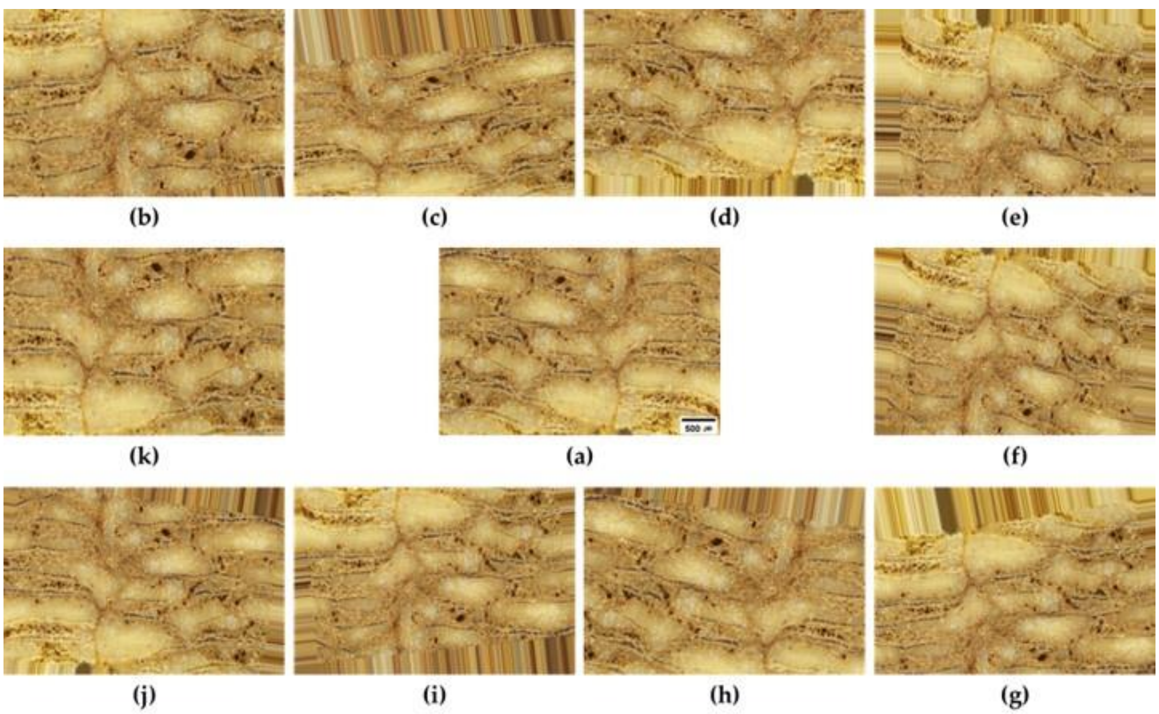

Fig. 1. Example micrographs of *Quercus aliena*: (a) original image; (b-k) augmented images

The dataset consisted of 1,400 images, with 200 images per species. The dataset was divided into 80% for training and 20% for testing. The test dataset was created by using a file random extraction program to extract 40 out of 200 images of each species dataset. The selected images were then saved in the separated paths with the training dataset. Data augmentation was exclusively implemented on the training dataset. Table 2 shows the disparity in quantity between the pre- and post-augmentation data.

Table 2. Composition of Training and Test Dataset in Augmentation

Scientific Name	Non-augmented			Augmented		
	Train (80%)	Test (20%)	Sum	Train	Test	Sum
Quercus dentata	160	40	200	1,773	40	1,813
Quercus serrata	160	40	200	1,761	40	1,801
Quercus mongolica	160	40	200	1,768	40	1,808
Quercus variabilis	160	40	200	1,751	40	1,791
Quercus aliena	160	40	200	1,784	40	1,824
Quercus acutissima	160	40	200	1,764	40	1,804
Quercus suber	160	40	200	1,778	40	1,818
Total	1,120	280	1,400	12,379	280	12,659

Verification factors influencing CNN

Classification performance and the factors influencing performance were analyzed using basic CNN architecture, as depicted in Fig. 2. The architecture of the CNN was designed by referring to the general structure reported in the past (Elgendy 2021; Loy 2020). The CNN design comprised four convolutional layers, four MaxPooling layers, and two fully connected layers. Furthermore, the model was enhanced by incorporating two dropout layers and a flattened layer. Finally, the Softmax activation function was applied.

Fig. 2. Convolutional neural network architecture for species classification in the present study

The CNN architecture was trained using a dataset comprising seven oak species. Categorical cross-entropy was used as the loss function for the multiclass classification. Optimizers are tools used to optimize the resources used during weight updates in each training phase (Cho 2018). In this study, three specific optimizers (i.e., SGD, Adam, and RMSProp) were used and compared to assess the influence of different optimizers on the efficiency of a CNN. The learning rates for each optimizer were set to SGD 0.0001, Adam 0.001 (default), and RMSProp 0.00001, respectively.

Statistical analysis of factors influencing CNN

Bivariate correlation analysis (SPSS 26.0; IBM, New York, USA) was used to examine the Pearson correlation coefficients between the variables. The optimizer type and augmentation for the analysis were used as nominal variables, whereas the accuracy and loss rate were used as scale variables. Additionally, a one-way ANOVA and Duncan's post-hoc analysis were used to examine the homogeneous subsets of the results.

RESULTS AND DISCUSSION

Figure 3 shows cross-sectional micrographs of the bark of the seven oak species in the dataset. Except for *Q. suber* bark, sclereids were found in the bark of all species. The sclereids' size, shape, and frequency varied among the species. Table 3 shows a comparison of the number of sclereids per unit area and the size of individual sclereids by species. The sclereids in *Q. variabilis* and *Q. aliena* (181.9×10^3 µm^2 and 175.4×10^3 µm^2, respectively) were larger than those in the other species, while *Q. serrata* and *Q. acutissima* showed smaller sizes (52.7×10^3 µm^2 and 65.4×10^3 µm^2, respectively) compared to other species. This represented 2- to 4-fold differences in the sclereids size between species. *Q. variabilis* and *Q. aliena* showed lower sclereid frequencies than other species. As shown in Fig. 3f, sclereids in *Q. acutissima*, in particular, had a large variation in size. Prasetia *et al*. (2022) previously reported that sclereids were frequently found in the bark of *Q. variabilis* and were absent in *Q. suber*. Kim (1993) conducted a similar study on the anatomical characteristics of *Q. variabilis* and *Q. suber* and reported that sclereids were rarely present in *Q. suber*.

Table 3. Quantitative Characteristics of Sclereids in the Six Oak Species

Quantitative Factors		*Q. dentata*	*Q. serrata*	*Q. mongolica*	*Q. variabilis*	*Q. aliena*	*Q. acutissima*
Sclereids number in 9.4 mm^2		35.0 ± 2.6^{bc}	10.0 ± 2.6^{a}	26.0 ± 3.6^{abc}	19.3 ± 10.2^{ab}	20.7 ± 4.0^{ab}	39.3 ± 16.9^{c}
Area of sclereids (1,000 µm^2)	Area range (Min–Max)	11.2–314.4	7.2–226.2	12.2–434.2	7.2–892.4	14.2–845.8	2.8–400.1
	Average area	72.7^{a}	52.7^{a}	125.3^{b}	181.9^{c}	175.4^{bc}	65.4^{a}
Note: Sclereids were not observed in the bark of *Q. suber.* The same superscript lowercase letters beside the mean values in the same row denote non-significant outcomes at the 5% significance level for comparisons between species.							

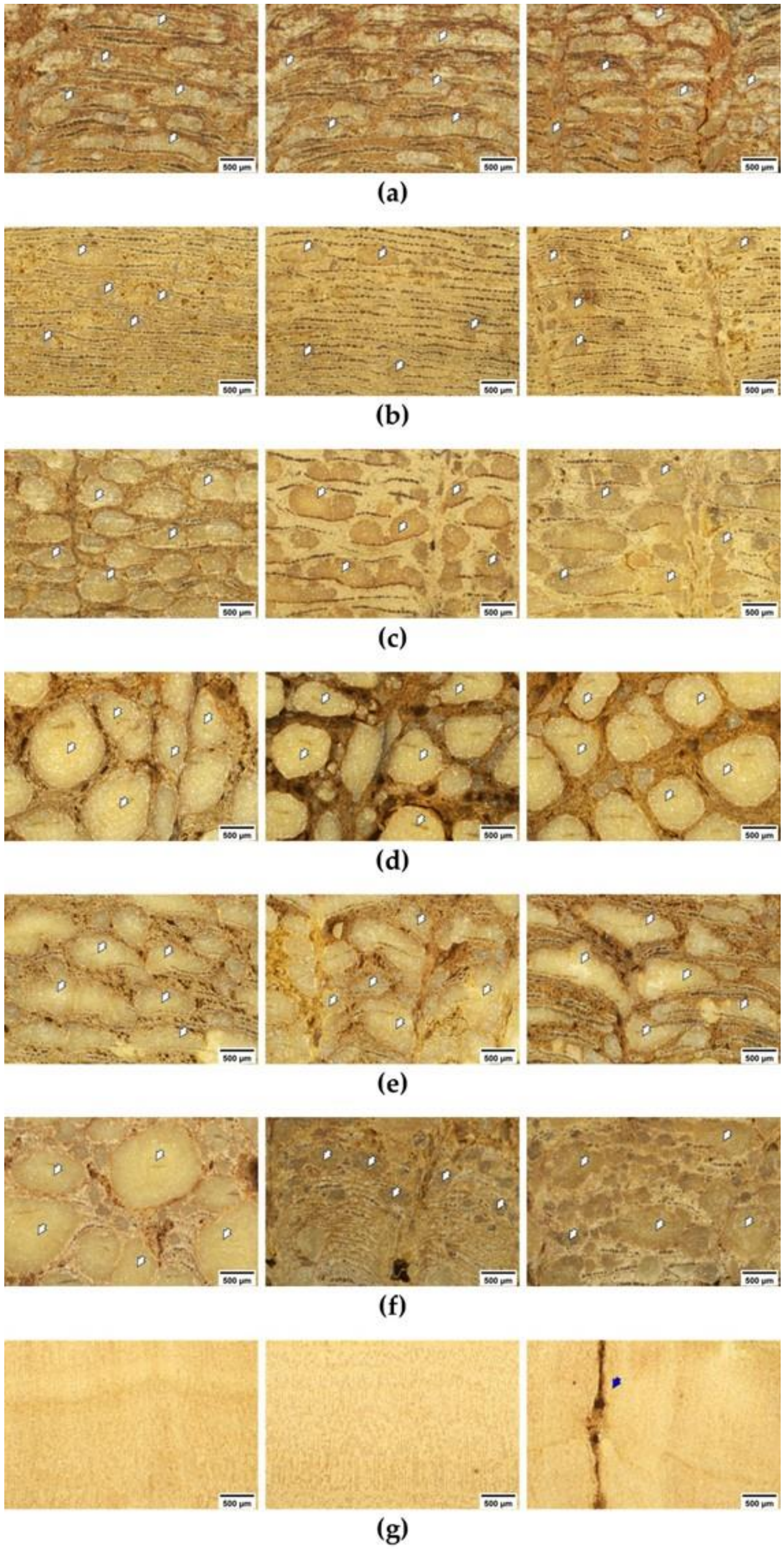

Fig. 3. Cross-section micrographs of the barks from seven *Quercus* species: *Quercus dentata* (a), *Quercus serrata* (b), *Quercus mongolica* (c), *Quercus variabilis* (d), *Quercus aliena* (e), *Quercus acutissima* (f), and *Quercus suber* (g). White arrows indicate sclereids. Scale bars: 1,000 µm

Table 4. Comparison of Average Loss and Accuracy on the Last Five Steps Per Optimizer

		SGD		Adam		RMSProp	
		Non-aug.	Aug-mented	Non-aug.	Aug-mented	Non-aug.	Aug-mented
Training Process	Loss	0.430[d]	0.001[a]	0.020[ab]	0.032[b]	0.365[c]	0.031[b]
	Accuracy	0.833[a]	1.000[c]	0.996[c]	0.995[c]	0.860[b]	0.996[c]
Test Process	Loss	0.498[a]	0.855[ab]	0.843[ab]	2.740[c]	0.623[ab]	1.021[b]
	Accuracy	0.820[ab]	0.830[ab]	0.798[a]	0.738[a]	0.795[a]	0.898[b]

Note: The same superscript lowercase letters beside the mean values in the same row denote non-significant outcomes at the 5% significance level for comparisons between species.

Table 4 lists the average losses and classification accuracies during the last five stages for each validation condition. During the training phase, Adam produced the lowest loss and highest accuracy, suggesting superior performance irrespective of the implementation of data augmentation. However, Adam's performance was relatively poor compared to the other optimizers during the test phase. SGD demonstrated a classification accuracy of approximately 83%, whereas RMSProp achieved a classification accuracy ranging from 79% to 90%. SGD is a function related to the random extraction and calculation of partial data at each learning stage and updating them more quickly and frequently. RMSProp is a function of applying an exponential moving average to prevent gradient loss problems, which can be obtained by lowering the reflection weight of the initial training and increasing the reflection weight of the recent gradient, thereby preventing the vanishing gradient problem during early training (Géron 2020). Therefore, in this study, SGD demonstrated superior performance in high-resolution microscopic image analysis required for species classification using bark, while RMSProp showed excellent performance in microscopic image analysis of the bark despite potential variations.

The classification accuracy and loss of bark for the seven oak species using a CNN are shown in Figs. 4 and 5. Under most training and test conditions, there was a notable increase in accuracy and a decrease in loss as the number of epochs progressed. Nevertheless, a noteworthy anomaly was detected in the training phase when the Adam optimizer was employed with the augmented datasets. In this particular instance, the observed loss exhibited an increasing pattern as the number of epochs increased, in contrast to the outcomes observed under other conditions. The observed pattern indicates overfitting, a phenomenon in which the model demonstrates high performance on data that closely resemble the training set, but exhibits poor performance on test or validation data (Oh 2021). This was because the noise generated during the augmentation process of the dataset negatively affected the learning results and only caused overfitting when the Adam optimizer with an augmented dataset was used for training. In this study, all conditions except the Adam-augmented dataset condition exhibited a notable trend of reduced loss and enhanced accuracy as the number of epochs increased, particularly when utilizing the augmented dataset as opposed to the non-augmented dataset. The utilization of the augmented dataset led to the stability of both the loss and classification accuracy after roughly 15 to 20 epochs, whereas the non-augmented dataset needed approximately 70 to 80 epochs for stabilization.

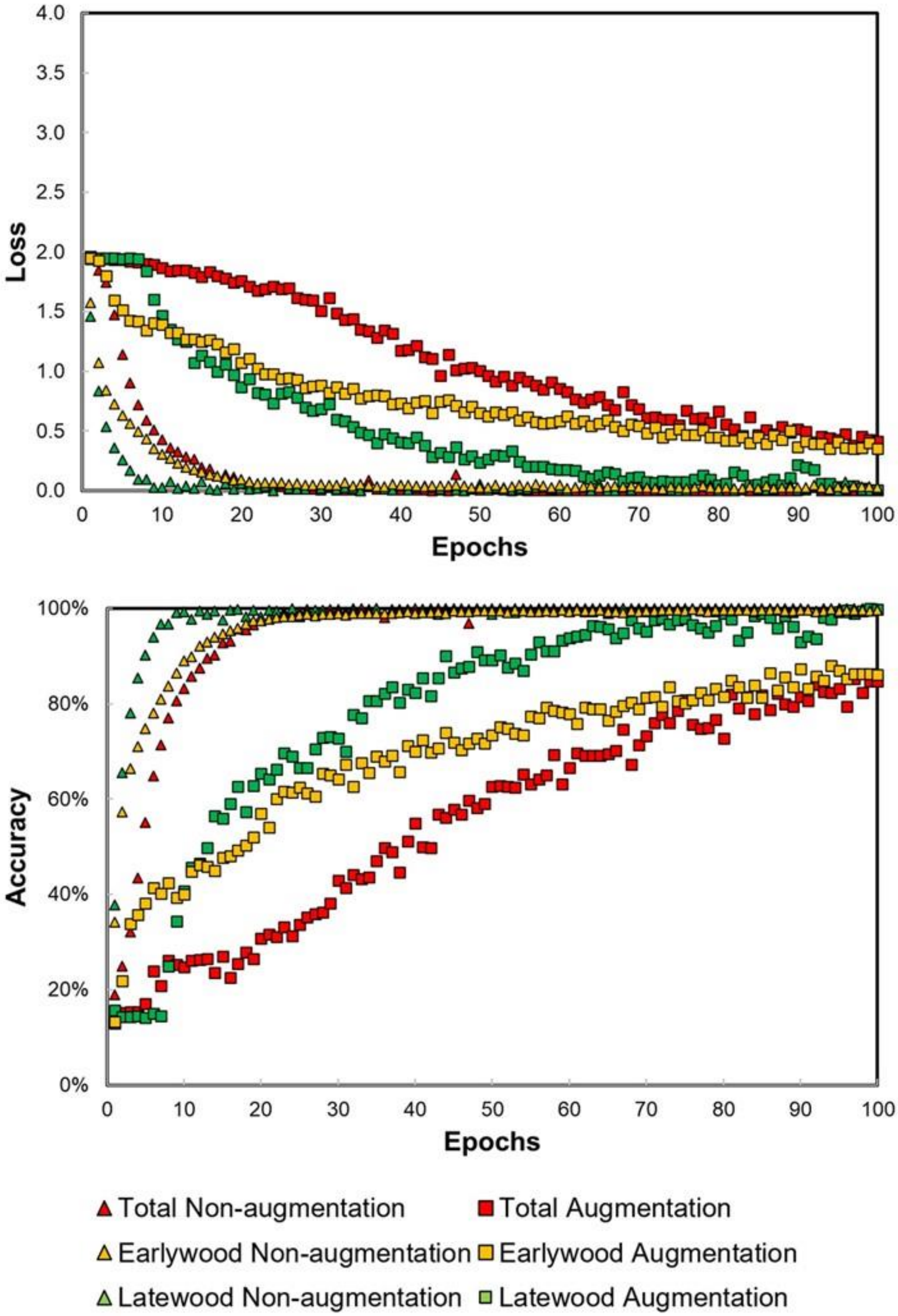

Fig. 4. Verification results of convolutional neural networks (CNNs) architecture for the seven oak species classification using the barks in training phase

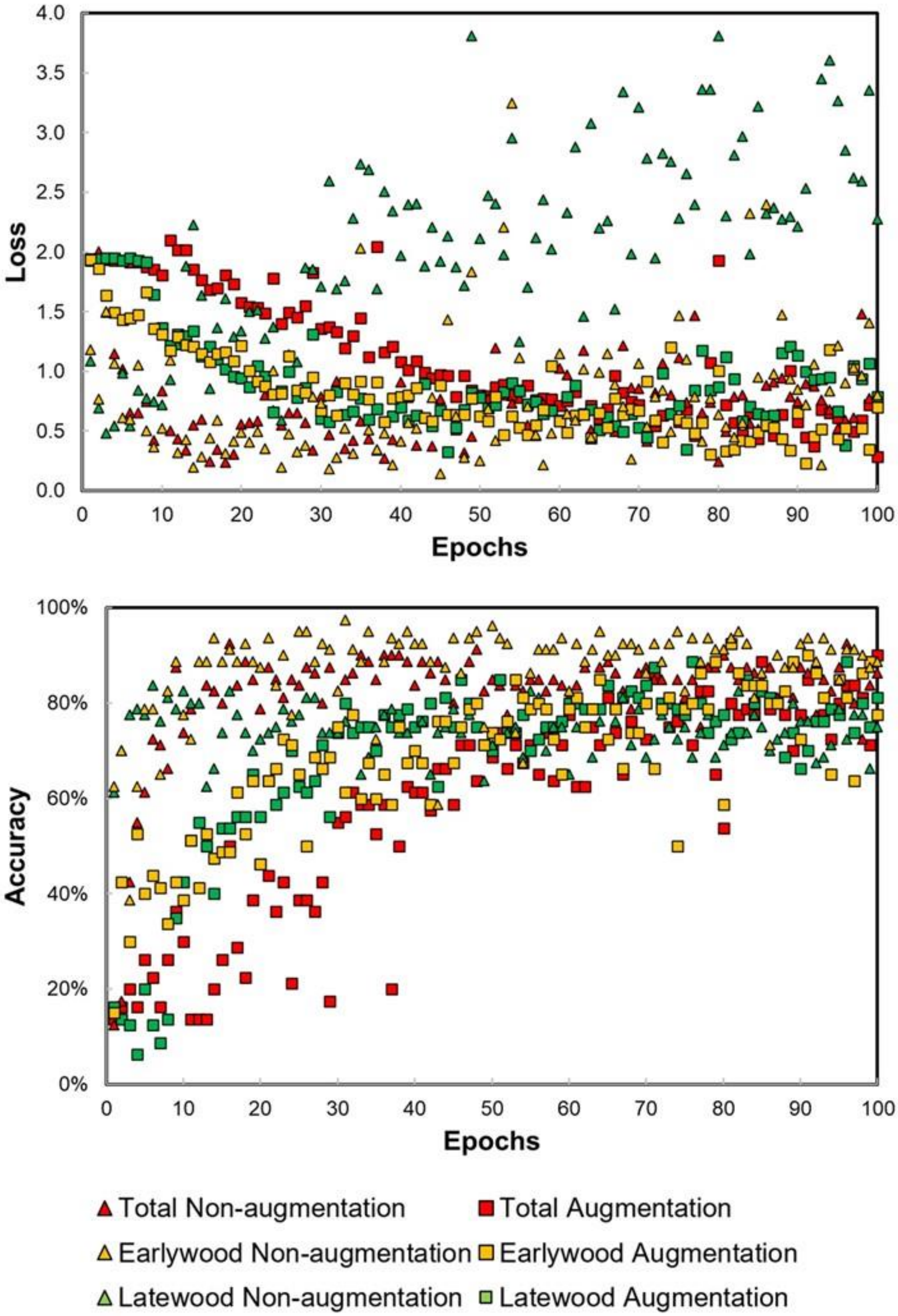

Fig. 5. Verification results of convolutional neural networks (CNNs) architecture for the seven oak species classification using the barks in test phase

Table 5. Correlation of the Factors Influencing Convolutional Neural Networks

N = 600	Epochs	Loss (Train)	Accuracy (Train)	Loss (Test)	Accuracy (Test)	Optimizer			Augmentation	
						(SGD)	(Adam)	(RMS Prop)	(No)	(Yes)
Epochs	1	−0.525** p = 0.000	0.530** p = 0.000	−0.031 p = 0.443	0.493** p = 0.000	0.000 p = 1.000	0.000 p = 1.000	0.000 p = 1.000	0.000 p = 1.000	0.000 p = 1.000
Loss (Train)	−0.525** p = 0.000	1	−0.998** p = 0.000	0.110** p = 0.007	−0.854** p = 0.000	0.218** p = 0.000	−0.208** p = 0.000	−0.011 p = 0.791	0.624** p = 0.000	−0.624** p = 0.000
Accuracy (Train)	0.530** p = 0.000	−0.998** p = 0.000	1	−0.121** p = 0.003	0.864** p = 0.000	−0.212** p = 0.000	0.196** p = 0.000	0.015 p = 0.706	−0.614** p = 0.000	0.614** p = 0.000
Loss (Test)	−0.031 p = 0.443	0.110** p = 0.007	−0.121** p = 0.003	1	−0.458** p = 0.000	−0.167** p = 0.000	0.428** p = 0.000	−0.260** p = 0.000	−0.189** p = 0.000	0.189** p = 0.000
Accuracy (Test)	0.493** p = 0.000	−0.854** p = 0.000	0.864** p = 0.000	−0.457** p = 0.000	1	−0.118** p = 0.004	−0.076 p = 0.063	0.194** p = 0.000	−0.456** p = 0.000	0.456** p = 0.000
Optimizer (SGD)	0.000 p = 1.000	0.218** p = 0.000	−0.212** p = 0.000	−0.167** p = 0.000	−0.118** p = 0.004	1	−0.500** p = 0.000	−0.500** p = 0.000	0.000 p = 1.000	0.000 p = 1.000
Optimizer (Adam)	0.000 p = 1.000	−0.208** p = 0.000	0.196** p = 0.000	0.428** p = 0.000	−0.076 p = 0.063	−0.500** p = 0.000	1	−0.500** p = 0.000	0.000 p = 1.000	0.000 p = 1.000
Optimizer (RMSProp)	0.000 p = 1.000	−0.011 p = 0.791	0.015 p = 0.706	−0.260** p = 0.000	0.194** p = 0.000	−0.500** p = 0.000	−0.500** p = 0.000	1	0.000 p = 1.000	0.000 p = 1.000
Augmentation (No)	0.000 p = 1.000	0.624** p = 0.000	−0.614** p = 0.000	−0.189** p = 0.000	−0.456** p = 0.000	0.000 p = 1.000	0.000 p = 1.000	0.000 p = 1.000	1	−1.000** p = 0.000
Augmentation (Yes)	0.000 p = 1.000	−0.624** p = 0.000	0.614** p = 0.000	0.189** p = 0.000	0.456** p = 0.000	0.000 p = 1.000	0.000 p = 1.000	0.000 p = 1.000	−1.000** p = 0.000	1
**The correlation is significant at the 0.01 level (2-tailed).										

Table 6. Homogeneous Subset Output of the Basic CNN Model

Process	Output	SGD		Adam		RMSProp	
		Non-aug.	Augmented	Non-aug.	Augmented	Non-aug.	Augmented
Training	**Loss**	1.106^{d}	0.142^{a}	0.519^{b}	0.059^{a}	0.772^{c}	0.115^{a}
	Accuracy	0.559^{a}	0.946^{d}	0.790^{c}	0.979^{d}	0.691^{b}	0.961^{d}
Test	**Loss**	1.099^{c}	0.708^{a}	0.908^{b}	2.128^{d}	0.807^{ab}	0.807^{ab}
	Accuracy	0.583^{a}	0.818^{d}	0.678^{b}	0.744^{c}	0.687^{b}	0.871^{e}
Note: The same superscript lowercase letters beside the mean values in the same row denote non-significant outcomes at the 5% significance level for comparisons between species.							

During the test phase, when the augmented dataset was applied during the training process, the SGD and RMSProp optimizer conditions reached a stable state after 15 to 20 epochs without overfitting. In contrast, the non-augmented dataset took approximately 40 to 50 epochs for the optimizer conditions to reach a stable state. The results obtained in this study are in line with previous results from the authors (Kim *et al*. 2023), which demonstrated that the utilization of an augmented dataset resulted in a more rapid stabilization of loss and classification accuracy compared to the non-augmented dataset. Rapid stabilization can also be ascribed to the advantages of using data augmentation, which includes mitigating the occurrence of overfitting and enhancing the classification accuracy (Wong *et al*. 2016; Fujita and Takahara 2017; Shorten and Khoshgoftaar 2019). However, it should be noted that there are possibilities of noise influences, as observed in the Adam optimizer.

Table 5 presents the correlation between the conditions applied to the training and testing of the CNN. In the training phase, the loss decreased as the number of epochs increased; however, no clear relationship was observed during the test phase. However, in both the training and test phases, the accuracy showed a proportional increase with the increasing number of epochs. Notably, in both the training and test phases, the accuracy displayed a negative trend with the epochs when SGD was used. However, the accuracy increased with the application of RMSProp and Adam during the training and test phases. During the test phase, data augmentation improved the performance in terms of both loss and accuracy. In contrast, it tended to increase both loss and the accuracy, during the training phase. Accordingly, the results showed that the number of epochs and optimizer selection affected the classification accuracy.

Table 6 presents the results of verifying the homogeneous subsets for each validation condition during CNN training. During the training phase, the augmented dataset was classified into the same subset for loss and accuracy regardless of the type of optimizer used. During the testing phase, the augmented dataset was classified into the same subset of losses only when SGD and RMSProp were used. The non-augmented dataset revealed independent subsets for all validation conditions during the training phase. Only the loss and accuracy from the non-augmented Adam and RMSProp methods were classified into the same subset during the test phase.

CONCLUSIONS

1. The bark of the seven oak species exhibited distinct variations in the size, shape, and frequency of sclereids.
2. The classification accuracy of species based on the bark was significantly improved when using data augmentation (0.456**) and the RMSProp optimizer (0.194**). The loss reduction was pronounced when using RMSProp (−0.260**), data augmentation (-0.189**), and SGD (−0.167**). Consequently, the augmented dataset with the RMSProp optimizer exhibited optimal performance, reaching 89.8%.
3. Homogeneous subsets using the augmented dataset among the validation conditions were classified into the same subsets for accuracy and loss during the training phase, regardless of the optimizer used. Only the conditions with the Adam and RMSProp optimizers for the non-augmented dataset were classified into the same subsets for accuracy and loss during the testing phase.

4. In conclusion, CNNs used for species classification using sclereid characteristics in the bark of seven oak species showed classification accuracies ranging between 74% and 90%. Sclereid characteristics are expected to serve as useful indicators for facilitating species classification. Especially, the relatively simple preprocessing and inspection procedures in the trade and quarantine procedures of commercial cork resources from *Quercus suber* and *Quercus variabilis* are expected.

ACKNOWLEDGMENTS

This research was supported by the Science and Technology Support Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT) (No. 2022R1A2C1006470), the Basic Science Research Program through the NRF funded by the Ministry of Education (No. 2018R1A6A1A03025582), and the R&D Program for Forest Science Technology (Project Nos. 2021350C10-2323-AC03 and 2021311A00-2122-AA03) provided by the Korea Forest Service (Korea Forestry Promotion Institute).

REFERENCES CITED

Bertrand, S., Cerutti, G., and Tougne, L. (2017). "Bark recognition to improve leaf-based classification in didactic tree species identification," in: *Proceedings of the 12th International Joint Conference* on *Computer Vision, Imaging* and *Computer Graphics Theory* and *A*pplications (VISIGRAPP 2017), Porto, Portugal, pp. 435-442.

Bremananth, R., Nithya, B., and Saipriya, R. (2009). "Wood species recognition system," *World Academy of Science, Engineering and Technology* 52, 873-879.

Cao, S., Zhou, S., Liu, J., Liu, X., and Zhou, Y. (2022). "Wood classification study based on thermal physical parameters with intelligent method of artificial neural networks," *BioResources* 17(1), 1187-1204. DOI: 10.15376/biores.17.1.1187-1204

Carpentier, M., Giguere, P., and Gaudreault, J. (2018). "Tree species identification from bark images using convolutional neural networks," in: *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, pp. 1075-1081.

Cho, T. H. (2018). *Deep-Learning for Everybody*, Gilbut Publishing, Seoul, Republic of Korea.

Elgendy, M. (2021). *Deep Learning for Vision Systems*, Hanbit Media, Seoul, Republic of Korea.

Fiel, S., and Sablatnig, R. (2010). "Automated identification of tree species from images of the bark, leaves and needles," in: *Proceedings of the 16th Computer Vision Winter Workshop*, Mitterberg, Austria, pp. 67-74.

Fujita, K., and Takahara, A. (2017). *Deep Learning Bootcamp with Keras*, Gilbut Publishing, Seoul, Republic of Korea.

Géron, A. (2020). *Hands-on machine learning with Scikit-Learn, Keras, and Tensorflow*, Hanbit Media, Seoul, Republic of Korea.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in: *Proceeding of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, pp. 770-778.

Hermanson, J. C., and Wiedenhoeft, A. C. (2011). “A brief review of machine vision in the context of automated wood identification systems,” *IAWA Journal* 32(2), 233-250. DOI: 10.1163/22941932-90000054

Huang, P., Zhao, F., Zhu, Z., Zhang, Y., Li, X., and Wu, Z. (2021). “Application of variant transfer learning in wood recognition,” *BioResources* 16(2), 2557-2569. DOI: 10.15376/biores.16.2.2557-2569

Hwang, S. W., and Sugiyama, J. (2021). “Computer vision-based wood identification and its expansion and contribution potentials in wood science: A review,” *Plant Methods* 17(1), 47. DOI: 10.1186/s13007-021-00746-1

Hwang, S. W., Kobayashi, K., and Sugiyama, J. (2020). “Detection and visualization of encoded local features as anatomical predictors in cross-sectional images of Lauraceae,” *Journal of Wood Science* 66(1), 1-16. DOI: 10.1186/s10086-020-01864-5

IAWA Committee. (2016). “IAWA list of microscopic bark features,” *IAWA Journal* 37(4), 517-615. DOI: 10.1163/22941932-20160151

Ilic, J. (1993). “Computer aided wood identification using Csiroid,” *IAWA Journal* 14(4), 333-340. DOI: 10.1163/22941932-90000587

Kim, B. R. (1993). “Studies on the physical and mechanical properties of imported and domestic corks,” *Journal of the Korean Wood Science and Technology* 21(4), 45-54.

Kim, J. H., Kim, D. H., Kim, S. H., Suri, I. F., Purusatama, B. D., Jo, J. I., Lee, H. S., Hidayat, W., Febrianto, F., Lee, S. H., and Kim, N. H. (2021). “Comparison of anatomical features in the three *Syzygium* species,” *BioResources* 16(2), 3631-3642. DOI: 10.15376/biores.16.2.3631-3642

Kim, J. H., Purusatama, B. D., Savero, A. M., Prasetia, D., Yang, G. U., Han, S. Y., Lee, S. H., and Kim, N. H. (2023). “Performance influencing factors of convolutional neural network models for classifying certain softwood species,” *Forests* 14(6), article 1249. DOI: 10.3390/f14061249

Kim, T. K., Hong, J. H., Ryu, D., Kim, S. K., Byeon, S. Y., Huh, W. J., Kim, K. H., Baek, G. H., and Kim, H. S. (2022). “Identifying and extracting bark key features of 42 tree species using convolutional neural networks and class activation mapping,” *Scientific Report* 12(1), article 4772. DOI: 10.1038/s41598-022-08571-9

Kwon, O. K., Lee, H. G., Lee, M. R., Jang, S. J., Yang, S. Y., Park, S. Y., Choi, I. G., and Yeo, H. M. (2017). “Automatic wood species identification of korean softwood based on convolutional neural networks,” *Journal of the Korean Wood Science and Technology* 45(6), 797-808. DOI: 10.5658/WOOD.2017.45.6.797.

Kwon, O. K., Lee, H. G., Yang, S. Y., Kim, H. B., Park, S. Y., Choi, I. G., and Yeo, H. M. (2019). “Performance enhancement of automatic wood classification of korean softwood by ensembles of convolutional neural networks,” *Journal of the Korean Wood Science and Technology* 47(3), 265-276. DOI: 10.5658/WOOD.2019.47.3.265

Loy, J. (2020). *Neural Network Projects with Python*, Gilbut Publishing, Seoul, Republic of Korea.

Oh, S. H. (2021). *Python Deep Learning Machine Learning*, Information Publishing Group, Seoul, Republic of Korea.

Prasetia, D., Purusatama, B. D., Kim, J. H., Jang, J. H., Park, S. Y., and Kim, N. H. (2022). “Qualitative anatomical characteristics of the virgin cork in *Quercus variabilis* Grown in Korea,” *BioResources* 18(1), 884-898. DOI: 10.15376/biores.18.1.884-898

Savero, A. M., Kim, J. H., Purusatama, B. D., Prasetia, D., Park, S. H., and Kim, N. H. (2022). "A comparative study on the anatomical characteristics of *Acacia mangium* and *Acacia hybrid* grown in Vietnam," *Forests* 13(10), 1700. DOI: 10.3390/f13101700

Savero, A. M., Kim, J. H., Purusatama, B. D., Prasetia, D., Park, S. H., Van Duong, D., and Kim, N. H. (2023). "Characterization of anatomical and non-anatomical properties for the identification of six commercial wood species from Vietnamese plantation forests," *Forests* 14(3), article 496. DOI: 10.3390/f14030496

Shorten, C., and Khoshgoftaar, T. M. (2019). "A survey on image data augmentation for deep learning," *Journal of Big Data* 6(1), 1-48. DOI: 10.1186/s40537-019-0197-0

Simonyan, K., and Zisserman, A. (2015). "Very deep convolutional for large-scale image recognition," in: *Proceeding of the International Conference on Learning Representations (ICLR) 2015*, San Diego, USA, pp. 1-14.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). "Going deeper with convolutions," in: *Proceeding 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, USA, pp. 1-9.

Tou, J. Y., Lau, P. Y., and Tay, Y. H. (2007). "Computer Vision-Based Wood Recognition System," in: *Proceedings of the international workshop on advanced image technology*, Bangkok, Thailand, pp. 1-6.

United Nations Economic Commission for Europe (2007). *Wood Resources Availability and Demands - Implications of Renewable Energy Policies*, United Nations Economic Commission for Europe, Geneva, Switzerland.

Wheeler, E. A., and Baas, P. (1998). "Wood identification – A review," *IAWA Journal* 19(3), 241-264. DOI: 10.1163/22941932-90001528

Wong, S. C., Gatt, A., Stamatescu, V., and McDonnell, M. D. (2016). "Understanding data augmentation for classification: When to warp?," in: *2016 International Conference on Digital Image Computing*, Gold Coast, Australia, pp. 1-6.

Yang, S. Y., Lee, H. G., Park, Y. G., Chung, H. W., Kim, H. B., Park, S. Y., Choi, I. G., Kwon, O. K., and Yeo, H. M. (2019). "Wood species classification utilizing ensembles of convolutional neural networks established by near-infrared spectra and images acquired from Korean softwood lumber," *Journal of the Korean Wood Science and Technology* 47(4), 385-392. DOI: 10.5658/WOOD.2019.47.4.385

Zhao, Z., Yang, X., Ge, Z., Guo, H., and Zhou, Y. (2021). "Wood microscopic image identification method based on convolution neural network," *BioResources* 16(3), 4986-4999. DOI: 10.15376/biores.16.3.4986-4999

Article submitted: September 12, 2023; Peer review completed: November 11, 2023; Revised version received and accepted: November 14, 2023; Published: November 27, 2023.
DOI: 10.15376/biores.19.1.510-524