

Wood Species Classification Utilizing Ensembles of Convolutional Neural Networks Established by Near-Infrared Spectra and Images Acquired from Korean Softwood Lumber¹

Sang-Yun Yang^{2,3} · Hyung Gu Lee⁴ · Yonggun Park^{2,3} · Hyunwoo Chung² · Hyunbin Kim² · Se-Yeong Park⁵ · In-Gyu Choi^{2,3,6} · Ohkyung Kwon^{4,†} · Hwanmyeong Yeo^{2,3,†}

ABSTRACT

In our previous study, we investigated the use of ensemble models based on LeNet and MiniVGGNet to classify the images of transverse and longitudinal surfaces of five Korean softwoods (cedar, cypress, Korean pine, Korean red pine, and larch). It had accomplished an average F1 score of more than 98%; the classification performance of the longitudinal surface image was still less than that of the transverse surface image. In this study, ensemble methods of two different convolutional neural network models (LeNet3 for smartphone camera images and NIRNet for NIR spectra) were applied to lumber species classification. Experimentally, the best classification performance was obtained by the averaging ensemble method of LeNet3 and NIRNet. The average F1 scores of the individual LeNet3 model and the individual NIRNet model were 91.98% and 85.94%, respectively. By the averaging ensemble method of LeNet3 and NIRNet, an average F1 score was increased to 95.31%.

Keywords: lumber species classification, convolutional neural networks, ensemble methods, LeNet, NIRNet, NIR spectroscopy

1. INTRODUCTION

Species identification of lumber is important for various purpose (Kim and Choi, 2016; Eom and Park,

2018; Lee *et al.*, 2018; Park *et al.*, 2018). In our previous study (Kwon *et al.*, 2019), we showed that an ensemble model combining the LeNet2, LeNet3, and MiniVGGNet4 models improved the classification

¹ Date Received May 1, 2019, Date Accepted July 2, 2019

² Department of Forest Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

³ Research Institute of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

⁴ National Instrumentation Center for Environmental Management (NICEM), Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea

⁵ Department of Forest Biomaterials Engineering, Kangwon National University, 1 Gangwondaehakgil, Chuncheon 24341, Republic of Korea

⁶ Institutes of Green Bio Science and Technology, Seoul National University, 1447 Pyeongchang-daero, Daehwa-myeon, Pyeongchang 25354, Republic of Korea

† Corresponding author: Hwanmyeong Yeo (e-mail: hyeo@snu.ac.kr, ORCID: 0000-0002-1779-069X),

† Corresponding author: Ohkyung Kwon (e-mail: zoom@snu.ac.kr, ORCID: 0000-0002-6307-0060)

performance of five Korean softwood species. We confirmed that the ensemble model improved the classification performance for the ten classes of species-surface combinations (e.g., cedar-transverse, cedar-longitudinal, Korean pine-transverse, Korean pine-longitudinal, and so on). However, the performance improvement was limited when only images were used. This limitation originated from the similarity between convolutional neural network (CNN) architectures and a high degree of similarity between the characteristic patterns in the images. If no significant difference was observed between the image patterns, CNN models with a similar structure or an ensemble model of such CNN models might have a little chance to increase the classification performance.

Ensemble models showed better performance when they were composed of different characteristics (Schapire, 1990; Samat *et al.*, 2014). As the images from different wood surfaces might reveal very similar patterns, we need to obtain different types of information related to the characteristics of wood species. A near-infrared (NIR) spectrum acquired from a surface of wood is a good example. The NIR region contains information about the chemical characteristics of the wood (Cho *et al.*, 2016; Yang *et al.*, 2017). Therefore, it has potential for the classification of wood species. Thus far, several studies on the automatic classification of wood species by utilizing the NIR spectra have been conducted (Adedipe *et al.*, 2008; Russ *et al.*, 2009; Nisgoski *et al.*, 2017; Park *et al.*, 2017; Yang *et al.*, 2019). However, the models for NIR spectra also have a limitation to the improvement of the classification performance for certain species. In our previous study (Yang *et al.*, 2019), there were a lot of misclassifications between cedar and cypress, and between Korean red pine and Korean pine.

An image and an NIR spectrum from the same surface of wood contain different types of information. Thus, models characterizing patterns on wood surfaces by

using images as well as models characterizing chemical information on wood by using NIR spectra are good candidates for an ensemble model. When two types of models show better classification performance for different wood species, the ensemble of these two models can compensate each other for better classification performance.

In this study, we developed an ensemble model combining an one-dimensional CNN (1D CNN) model trained with the NIR spectra and a two-dimensional CNN (2D CNN) model trained with smartphone images from the radial surfaces of lumbers for the wood species classification of five Korean softwood species (cedar, cypress, Korean pine, Korean red pine, and larch). We used precision, recall, and F1 score as well as the confusion matrix to describe the classification performance of the model selected as the ensemble model.

2. MATERIALS and METHODS

2.1. Sample preparation

Five Korean softwood species (cedar [*Cryptomeria japonica*], cypress [*Chamaecyparis obtusa*], Korean pine [*Pinus koraiensis*], Korean red pine [*Pinus densiflora*], and larch [*Larix kaempferi*]) were used for the wood species classification utilizing ensembles of CNN models established by different features (image and NIR). We collected 50 lumber samples of each species, measuring 50 mm × 100 mm × 600 mm (thickness × width × length), from several mills participating in the National Forestry Cooperative Federation of Korea. The lumber samples in each species were from different regions of Korea. The lumber samples were air-dried (10 – 15% MC) at 25°C and 65 ±10% relative humidity for three months. After air-drying, wood blocks measuring 20 mm × 2 mm × 40 mm (R × T × L) were prepared from each lumber

(50 wood samples per species) and images of the radial surface were taken from the blocks. The same radial surface was used for acquiring the NIR spectrum.

2.2. Image acquisition and dataset preparation

We used a smartphone camera (Galaxy Note 9, Samsung, Korea) to obtain macroscopic images of the radial section of the wood samples. The smartphone was placed on a simple frame for stable support during image acquisition. The camera in the smartphone had an $f/1.5$ lens and produced a color picture of 4032×3024 pixels. Totally, we prepared 475 images from the samples. After the background removal of the images, the images were cropped to a size of approximately 2000×1000 pixels. According to Brunner *et al.* (1990), there was no distinct relationship between moisture content and discoloration of wood below fiber saturation point. In this study, sample discoloration was negligible because air-dried samples were employed for image acquisition.

The NIR spectra were taken from the same surface as that used by the smartphone to acquire images of the radial surfaces of woods (in all, 475 spectra). All the NIR spectra were acquired at 1-nm intervals of wavelength ranging from 780 nm to 2500 nm by a NIR spectrometer (SpectraStar 2600XT-R Unity Scientific, USA). The scans-to-average value was 12.

A random k-fold cross-validation ($k = 5$) was applied

to the dataset. Table 1 lists the number of samples for each species, and the sample distributions of the cross-validation. Thus, an image and its corresponding NIR spectrum were paired in the k-fold cross-validation.

2.3. Model training

For training the CNN models and the ensemble model, we utilized a workstation with XEON CPU (14 cores) with 64 GB of memory as well as GPU with 24 GB (NVIDIA Quadro M6000). The operating system was Ubuntu 16.04 LTS with CUDA 8.0, Python 3.5, Tensorflow 1.2, and Keras 2.0.

The structures of CNN models are shown in Table 2. LeNet3 is a 2D CNN model for the species classification model using image data, as already discussed in our previous study (Kwon *et al.*, 2017). The background-removed images were reduced to $128 \times 128 \times 3$ as input images for training purposes. The pixel values of the input images were normalized by 255. Image augmentation was performed before model training with the following parameters: rotation range = $\pm 30^\circ$, shift range in width and height = $\pm 10\%$, shear range = $\pm 20\%$, zoom range = $\pm 20\%$, and horizontal flip. We utilized the adaptive moment estimation (Adam) algorithm (Kingma and Ba, 2014) as an optimizer with learning rate = 0.0001. The number of epochs was 500.

NIRNet is a 1D CNN model using NIR spectrum data. The spectra having a dimension of 1721×1 were

Table 1. Sample class index and number of samples of each fold in k-fold cross-validation ($k = 5$)

Species	Class index	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Number of samples
Cedar	0	20	21	16	20	29	106
Cypress	1	18	26	21	20	17	102
Korean Pine	2	19	22	19	19	26	105
Korean Red Pine	3	18	17	25	20	15	95
Larch	4	18	10	14	16	9	67
Total		93	96	95	95	96	475

Table 2. Architecture of LeNet3 and NIRNet models

Layer set	Layer	LeNet3	NIRNet
Layer set 1	Convolution 2D	20, 5×5 , same, input	32, 3×1 , input
	Activation	ReLU	LeakyReLU
	Max pooling	2×2 , stride = 2×2	None
Layer set 2	Convolution 2D	50, 5×5 , same	16, 9×1
	Activation	ReLU	LeakyReLU
	Max pooling	2×2 , stride = 2×2	None
Layer set 3	Convolution 2D	50, 5×5 , same	8, 27×1
	Activation	ReLU	LeakyReLU
	Max pooling	2×2 , stride = 2×2	None
Layer set 4	Convolution 2D	50, 5×5 , same	1, 1×1
	Activation	ReLU	LeakyReLU
	Max pooling	2×2 , stride = 2×2	None
Fully connected layer	Dense	500	None
	Activation	ReLU	None
Output layer	Dense	5	5
	Activation	Softmax	Softmax

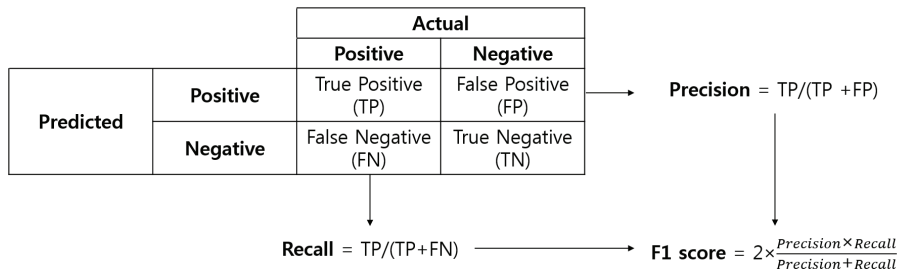


Fig. 1. Confusion matrix (Yang *et al.*, 2019).

used as the input for training. The absorbance values of the input spectra were standard normalized by subtracting the mean and then scaled to unit variance. Standard normalization was applied to each wavelength, not sample. We utilized the Adam algorithm as an optimizer with learning rate = 0.0001. The number of epochs was 1500.

The classification performances of the species classification for each CNN model and their ensembles were evaluated by constructing a confusion matrix. The confusion matrix is a table that consists of the

classification results. It shows a matrix with the actual and the predicted classes (Fig. 1). From the confusion matrix, some indices including precision, recall, and F1 score could be evaluated to describe the accuracy of a model (Goutte and Gaussier, 2005). The details of reliability indices are described in our previous paper (Yang *et al.*, 2019). In this study, F1 score is used for the evaluation of the classification performance of the models. The F1 score, which is the harmonic mean of precision and recall, can be calculated as described in Fig. 1.

2.4. Ensemble models and methods

An ensemble of CNN models requires several operational classification models such as LeNet (Lecun *et al.*, 1998) and VGGNet (Simonyan and Zisserman, 2014). We utilized two CNN models to construct ensemble models for the classification of two types of information from wood surfaces: (1) images of the radial surface and (2) NIR spectra for the same area in the radial surface. On the basis of the CNN classification results, we applied two types of ensemble models for prediction results: (1) averaging and (2) max voting.

3. RESULTS and DISCUSSION

3.1. Classification performance of LeNet3 and NIRNet models for radial surfaces

Table 3 shows the classification performance indices of the LeNet3 model based on the image data. The classification performance of the LeNet3 model was obtained by averaging the classification results of each fold in the k-fold classification (k=5). The average precision, recall, and F1 score of the LeNet3 model were 93.28%, 90.96%, and 91.98%, respectively. The LeNet3 model based on images from the radial surface had poorer classification performance than that from

Table 3. Performance measures of LeNet3 model by k-fold cross-validation (k = 5)

Class name	Precision (%)	Recall (%)	F1 score (%)
Cedar	94.93	92.75	93.83
Cypress	100.00	85.23	92.03
Korean Pine	92.07	94.74	93.38
Korean Red Pine	88.27	93.42	90.77
Larch	91.14	88.64	89.88
Average	93.28	90.96	91.98

the transverse surface in our previous study (Kwon *et al.*, 2017). From the classification results, we inferred that the radial surface of wood has less information than the transverse surface of wood. The precision ranged from 88.27% (Korean red pine) to 100.00% (cypress); however, the recall ranged from 85.23% (cypress) to 94.74% (Korean pine). F1 score, which is the harmonic average of precision and recall, ranged from 89.88% (larch) to 93.83% (cedar). Thus, the LeNet3 model based on images from the radial surface had low classification performance for larch.

Table 4 shows the classification performance indices of the NIRNet model based on the NIR spectral data. The k-fold cross-validation (k=5) was also applied. The classification performance of the NIRNet model was obtained by the same averaging in the case of the LeNet3 model. The average precision, recall, and F1 score of the NIRNet model were 86.22%, 85.69%, and 85.94%, respectively. The NIRNet model had lower classification performance than the LeNet3 model in this study. The precision ranged from 76.21% (Korean red pine) to 96.22% (cedar), and the recall ranged from 72.19% (Korean red pine) to 96.31% (cedar). F1 score ranged from 74.15% (Korean red pine) to 96.27% (cedar). Thus, the NIRNet model based on the NIR spectra faced difficulties in classifying the Korean red pine.

Table 4. Performance measures of NIRNet model by k-fold cross-validation (k = 5)

Class name	Precision (%)	Recall (%)	F1 score (%)
Cedar	96.22	96.31	96.27
Cypress	80.16	81.70	80.92
Korean Pine	82.84	84.04	83.44
Korean Red Pine	76.21	72.19	74.15
Larch	95.65	94.21	94.92
Average	86.22	85.69	85.94

Table 5. Performance measures of LeNet3-NIRNet ensemble model by averaging method.

Class name	Precision	Recall	F1 score
Cedar	100.00	100.00	100.00
Cypress	100.00	86.00	92.47
Korean Pine	93.00	95.79	94.37
Korean Red Pine	89.71	93.11	91.38
Larch	98.67	98.00	98.33
Average	96.28	94.58	95.31

Table 6. Performance measures of LeNet3-NIRNet ensemble model by max-voting method.

Class name	Precision	Recall	F1 score
Cedar	98.82	100.00	99.41
Cypress	100.00	86.00	92.47
Korean Pine	93.00	95.79	94.37
Korean Red Pine	89.71	92.31	90.99
Larch	98.67	98.00	98.33
Average	96.04	94.42	95.12

3.2. Performance of ensemble models

Tables 5 and 6 show the classification performance indices of the LeNet3-NIRNet ensemble models by using the averaging and the max-voting method, respectively. The k-fold cross-validation (k=5) was also applied. There were slight differences in the precision and the F1 score for cedar between the two methods. Both ensemble methods improved the classification indices compared with the individual LeNet3 model and the NIRNet model. The average precision, recall, and F1 score of the averaging ensemble model were 96.28%, 94.58%, and 95.31%, respectively. For the max-voting ensemble model, they were 96.04%, 94.42%, and 95.12%, respectively. For both the ensemble models, the F1 scores of cedar and larch were considerably improved; however, those of Korean red pine and cypress were similar to those calculated by the LeNet3 model. In our results, the LeNet3-NIRNet ensemble model using the averaging method was more accurate for the softwood species classification than individual CNN model.

The improvement of the classification performance of the ensemble model was interesting because two independent models with lower indices enhanced the classification result. It can be inferred that the ensemble method using different features (smartphone images and NIR spectra) compensated the classification performance. However, note that this study was a preliminary

application, so it was carried out with a small dataset (475 samples). The results should be clearly confirmed with database expansion in a further study.

4. CONCLUSION

In this study, we investigated the use of ensembles of the LeNet3 and NIRNet models to classify the species of five Korean softwoods (cedar, cypress, Korean pine, Korean red pine, and larch) using smartphone images and NIR spectra acquired from radial surfaces.

The experimental results showed that the ensemble model using the averaging method of LeNet3 and NIRNet was the most accurate. The max-voting ensemble model had slightly lower classification performance than the averaging ensemble model. The F1 score of the ensemble model using the averaging method was 95.31%. The classification performance for the radial surface of cedar and larch was considerably improved by the other individual models.

ACKNOWLEDGMENT

This study was carried out with the support of ‘R&D Program for Forest Science Technology (Project No. 2016009B10-1819-AB01)’ provided by Korea Forest Service (Korea Forestry Promotion Institute) and ‘R&D Program for Forest Science Technology (Project No. 2016009D10-1819-AB01)’

provided by Korea Forest Service (Korea Forestry Promotion Institute).

REFERENCES

- Adedipe, O.E., Dawson-Andoh, B., Slahor, J., Osborn, L. 2008. Classification of red oak (*Quercus rubra*) and white oak (*Quercus alba*) wood using a near infrared spectrometer and soft independent modelling of class analogies. *Journal of Near Infrared Spectroscopy* 16(1): 49-57.
- Brunner, C.C., Shaw, G.B., Butler, D.A., Funck, J.W. 2007. Using color in machine vision systems for wood processing. *Wood and Fiber Science* 22(4): 413-428.
- Cho, B.K., Lohoumi, S., Choi, C., Yang, S.M., Kang, S.G. 2016. Study on Rapid Measurement of Wood Powder Concentration of Wood-Plastic Composites using FT-NIR and FT-IR Spectroscopy Techniques. *Journal of the Korean Wood Science and Technology* 44(6): 852-863.
- Eom, Y.J., Park, B.D. 2018. Wood Species Identification of Documentary Woodblocks of Songkok Clan of the Milseong Park, Gyeongju, Korea. *Journal of the Korean Wood Science and Technology* 46(3): 270-277.
- Goutte, C., Gaussier, E. 2005. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *European Conference on Information Retrieval*. pp. 345-359. Springer, Berlin, Heidelberg.
- Kim, S.C., Choi, J. 2016. Study on Wood Species Identification for Daeungjeon Hall of Jeonghyesa Temple, Suncheon. *Journal of the Korean Wood Science and Technology* 44(6): 897-902.
- Kingma, D.P., Ba, J. 2014. Adam: A method for stochastic optimization. arXiv:1412.6980.
- Kwon, O., Lee, H.G., 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. *Journal of the Korean Wood Science and Technology* 45(6): 797-808.
- Kwon, O., Lee, H.G., Yang, S.Y., Kim, H., Park, S.Y., Choi, I.G., Yeo, H. 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.
- Lecun, Y., Bottou, L., Bengio, Y., Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11): 2278-2324.
- Lee, K.H., Seo, J.W., Han, G.S. 2018. Dating Wooden Artifacts Excavated at Imdang-dong Site, Gyeongsan, Korea and Interpreting the Paleoenvironment according to the Wood Identification. *Journal of the Korean Wood Science and Technology* 46(3): 241-252.
- Nisgoski, S., de Oliveira, A.A., de Muñiz, G.I.B. 2017. Artificial neural network and SIMCA classification in some wood discrimination based on near-infrared spectra. *Wood Science and Technology* 51(4): 929-942.
- Park, J., Oh, J., Hwang, I., Jang, H., Choi, J., Kim, S. 2018. Study on species identification for Pungnammun Gate (Treasure 308) in Jeonju, Korea. *Journal of the Korean Wood Science and Technology* 46(3): 278-284.
- Park, S.Y., Kim, J.C., Kim, J.H., Yang, S.Y., Kwon, O., Yeo, H., Cho, K.C., Choi, I.G. 2017. Possibility of Wood Classification in Korean Softwood Species Using Near-infrared Spectroscopy Based on Their Chemical Compositions. *Journal of the Korean Wood Science and Technology* 45(2): 202-212.
- Russ, A., Fišerová, M., Gigac, J. 2009. Preliminary study of wood species identification by NIR spectroscopy. *Wood Research* 54(4): 23-32.

- Samat, A., Du, P., Baig, M.H.A., Chakravarty, S., Cheng, L. 2014. Ensemble learning with multiple classifiers and polarimetric features for polarized SAR image classification. *Photogrammetric Engineering & Remote Sensing* 80(3): 239-251.
- Schapire, R.E. 1990. The strength of weak learnability. *Machine learning* 5(2): 197-227.
- Simonyan, K., and Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556.
- Yang, S.Y., Park, Y., Chung, H., Kim, H., Park, S.Y., Choi, I.G., Kwon, O., Cho, K.C., Yeo, H. 2017. Partial least squares analysis on near-infrared absorbance spectra by air-dried specific gravity of major domestic softwood species. *Journal of the Korean Wood Science and Technology* 45(4): 399-408.
- Yang, S.Y., Park, Y., Chung, H., Kim, H., Park, S.Y., Choi, I.G., Kwon, O., Yeo, H. 2019. Soft Independent Modeling of Class Analogy for Classifying Lumber Species Using Their Near-infrared Spectra. *Journal of the Korean Wood Science and Technology* 47(1): 101-109.