Original Article

Sung-Wook Hwang, Taekyeong Lee, Hyunbin Kim, Hyunwoo Chung, Jong Gyu Choi and Hwanmyeong Yeo*

Classification of wood knots using artificial neural networks with texture and local feature-based image descriptors

<https://doi.org/10.1515/hf-2021-0051>

Received March 17, 2021; accepted August 26, 2021;
published online November 9, 2021

Abstract: This paper describes feature-based techniques for wood knot classification. For automated classification of macroscopic wood knot images, models were established using artificial neural networks with texture and local feature descriptors, and the performances of feature extraction algorithms were compared. Classification models trained with texture descriptors, gray-level co-occurrence matrix and local binary pattern, achieved better performance than those trained with local feature descriptors, scale-invariant feature transform and dense scale-invariant feature transform. Hence, it was confirmed that wood knot classification was more appropriate for texture classification rather than an approach based on morphological classification. The gray-level co-occurrence matrix produced the highest F1 score despite representing images with relatively

low-dimensional feature vectors. The scale-invariant feature transform algorithm could not detect a sufficient number of features from the knot images; hence, the histogram of oriented gradients and dense scale-invariant feature transform algorithms that describe the entire image were better for wood knot classification. The artificial neural network model provided better classification performance than the support vector machine and *k*-nearest neighbor models, which suggests the suitability of the nonlinear classification model for wood knot classification.

Keywords: artificial neural network; local feature; machine learning; texture feature; wood knot classification.

1 Introduction

Accurate grading of lumber is a very important process not only for production quality control, but also for securing structural stability and enhancing consumer confidence (Feio and Machado 2015). In the wood industry, lumber grading is mostly performed manually, and visual inspection by humans is subjective and time-consuming. In addition, it has a remarkable disadvantage of high accuracy over 70–80% not being guaranteed due to the fatigue of the inspector caused by repetitive work (Gu et al. 2008; Lampinen et al. 1998).

Automated wood surface inspection techniques using computer vision and machine learning have been studied to aid or replace manual inspection. Early studies used color and shape-based approaches to detect wood knots (Alapuranen and Westman 1992; Lampinen et al. 1994; Kauppinen and Silvén 1996), and after which the knot classification was treated as a matter of texture classification (Kamal et al. 2017; Mahram et al. 2012; Xie and Wang 2015). Gray-level co-occurrence matrix (GLCM)-based Haralick texture features (hereinafter referred to as GLCM features) was preferred in studies for automated wood defect classification, and promising results were produced from models trained with GLCM features (Kamal et al. 2017;

***Corresponding author: Hwanmyeong Yeo**, Research Institute of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea; Department of Agriculture, Forestry and Bioresources, College of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea; and Department of Forest Sciences, College of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea, E-mail: hyeo@snu.ac.kr

Sung-Wook Hwang and Taekyeong Lee, Research Institute of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea, E-mail: dryellow@snu.ac.kr (S.-W. Hwang), 19850622@snu.ac.kr (T. Lee). <https://orcid.org/0000-0001-8265-8700> (S.-W. Hwang)

Hyunbin Kim and Hyunwoo Chung, Department of Agriculture, Forestry and Bioresources, College of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea, E-mail: yoururse@snu.ac.kr (H. Kim), totkfkarjwl@snu.ac.kr (H. Chung)

Jong Gyu Choi, Department of Forest Sciences, College of Agriculture and Life Sciences, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea, E-mail: chlwhdrb722@snu.ac.kr

Qayyum et al. 2016; Ruz et al. 2009; Xie and Wang 2015). Mahram et al. (2012) reported that models trained with multi-feature sets combined with Haralick and local binary pattern (LBP) features achieved better performance than those trained with each single feature set.

Recently, artificial intelligence algorithms, such as convolutional neural network (CNN) techniques, have become mainstream in the detection and classification of wood defects (Affonso et al. 2017; Chen et al. 2020; He et al. 2020; Kim et al. 2019; Ren et al. 2017). Several studies using CNN models have reported excellent performance close to 100% accuracy for wood defect classification (Chen et al. 2020; He et al. 2019, 2020; Jung et al. 2018). In addition, CNN techniques have been implemented as object detection models that localize and classify defects from wood surfaces (Chen et al. 2020; He et al. 2019; Urbonas et al. 2019).

Despite the latest developments in wood defect classification, manual feature extraction techniques are still useful tools for dealing with biological image data such as wood (Kobayashi et al. 2015; Lens et al. 2020; Souza et al. 2020). The usefulness of manual extraction techniques is that they are easier to access based on specific domain knowledge through operator manipulation or intervention (Hwang et al. 2020). While studies on manually extracted feature-based knot classification mainly focused on texture features, local features, another major category in general image recognition technology (Bay et al. 2008; Csurka et al. 2004; Dalal and Triggs 2005; Lowe 2004), were not often selected for wood defect classification (Hittawe et al. 2015; Kim et al. 2019). Hence, studies on the comparison between texture features and local features for the classification of wood defects are insufficient. The era of deep learning has arrived without a comparative study of manually extracted features for wood defect classification. Therefore, it is necessary to compare well-known manual techniques such as texture and local features, even though they showed relatively lower classification performance than CNNs.

In this study, features were extracted from wood knot images using the texture feature descriptors, gray-level co-occurrence matrix and local binary pattern, and the local feature descriptors, histogram of oriented gradients, scale-invariant feature transform, and dense scale-invariant feature transform, for the classification of wood knots. The extracted features were learned using an artificial neural network classifier to build classification models, and their performances were compared with those of the models based on the support vector machine and k -nearest neighbor classifiers. This paper provides a comparison between texture features and local features, and among classifiers from the analysis of the classification performances produced by the established models.

2 Materials and methods

2.1 Dataset

To build wood knot image dataset, surface images were acquired from the lumbers of *Larix kaempferi*, *Pinus densiflora*, *Pinus koraiensis*, *Pinus radiata*, *Cryptomeria japonica*, *Chamaecyparis obtusa*, and *Pseudotsuga menziesii*, the major commercial coniferous species in Korea. Both surfaces of each lumber were photographed using a digital single-lens reflex camera with EF-S 18–55 mm f/3.5–5.6 USM lens (Canon Inc., Tokyo, Japan) at a distance of 120 cm from the lumber. Afterwards, 1172 wood knots were cropped from the surface images and used as a dataset for classification (Table 1). Each image size was dependent on the actual size of the knot, and the pixel resolution was 0.187 mm/pixel. The knot images were annotated into four categories: decayed, encased, sound, and spike knots (Figure 1).

The dataset was divided into training and test sets at a four-to-one ratio, and used for classification model construction and evaluation. With stratified random sampling, the training and test sets preserved

Table 1: List of the knot defects and stratified division of the dataset into training and test sets.

Knot type	Dataset		Total
	Train	Test	
Decayed	197	50	247
Encased	219	55	274
Sound	449	112	561
Spike	72	18	90
Total	937	235	1172

Figure 1: Four kinds of wood knots in the dataset: (a) decayed knot, (b) encased knot, (c) sound knot, and (d) spike knot.

the percentage of samples for each class. As presented in Table 1, the dataset has an imbalanced composition ratio, with sound and spike knots accounting for 47.9% and 7.7% of the total, respectively.

2.2 Classification procedure

Figure 2 shows the experimental flow of the automated wood knot classification. First, because the color of the knots is an unstable characteristic that changes due to various biological and environmental factors, the original RGB images were converted to 8-bit gray scale. Next, the texture and local features were extracted from the training set images. Then, the extracted image features were learned using artificial neural network (ANN), k -nearest neighbor (k-NN), and support vector machine (SVM) classifiers with stratified k -fold cross-validation, respectively, to construct classification models. Subsequently, image features were extracted from the test set in the same manner as the training set. The features extracted from the test set were input to the classification model established by the training set, and the wood knot classification was terminated as the model returned the predicted classes of the test images.

2.3 Feature extraction

Texture feature extraction methods, GLCM (Haralick et al. 1973) and LBP (Ojala et al. 1996, 2002), and local feature extraction methods, histograms of oriented gradients (HOG) (Dalal and Triggs 2005), scale-invariant feature transform (SIFT) (Lowe 2004), and dense SIFT (DSIFT) (Liu et al. 2010) were employed to extract features from the wood knot images (Figure 3). Because the gray level and pattern of the pixels are different depending on the type of the wood knot, texture features such as GLCM and LBP, may be suitable for the representation of the knot images. On the other hand, local features such as HOG, SIFT, and DSIFT can be used as an approach based on morphological differences for each type of wood knot.

2.3.1 Gray level co-occurrence matrix: GLCM is a statistical method that represents the texture of an image based on the spatial relationship of adjacent pixels, that is, the difference in gray levels between

the pixels (Figure 3a). This method creates a matrix by calculating the occurrence frequency of pixel pairs having specific values in a specific spatial relationship (distance and angle between two pixels) in an image. Then, the values of the matrix are input into statistical equations to describe the image as texture, and the GLCM features of the image are produced. GLCM textures have a higher computational cost than other textures. Therefore, five texture properties, contrast, dissimilarity, homogeneity, energy, and correlation, were used among the texture properties proposed by Haralick et al. (1973), as follows:

$$\text{Contrast} = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2,$$

$$\text{Dissimilarity} = \sum_{i,j=0}^{N-1} P_{i,j} |i - j|,$$

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1 + (i - j)^2},$$

$$\text{Energy} = \sqrt{\sum_{i,j=0}^{N-1} P_{i,j}^2},$$

$$\text{Correlation} = \sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i - \mu_x)(j - \mu_y)}{\sqrt{(\sigma_x^2)(\sigma_y^2)}} \right],$$

where $P_{i,j}$ is the i th and j th elements in a normalized GLCM, and N is the number of gray levels.

$$\mu_x = \sum_{i=0}^{N-1} iP_x(i), \quad \mu_y = \sum_{j=0}^{N-1} jP_y(j), \quad \mu = \frac{\mu_x + \mu_y}{2},$$

$$\sigma_x^2 = \sum_{i=0}^{N-1} (i - \mu_x)^2 P_x(i), \quad \sigma_y^2 = \sum_{j=0}^{N-1} (j - \mu_y)^2 P_y(j).$$

GLCM feature sets consisting of values calculated at angles 0° , 45° , 90° , and 135° between neighboring pixels and rotation invariant GLCM feature sets consisting of the averages of all angles were constructed at the distances between pixels of 1 and 7.

2.3.2 Local binary patterns: LBP is a descriptor representing an image by encoding the relative brightness differences between a pixel and the surrounding pixels, that is, a 3×3 pixel region, in binary (Figure 3b). An image is represented as a 256-bins histogram

Figure 2: Pipeline for wood knot classification based on manually extracted features.

Figure 3: Diagrams of feature extraction algorithms used in automated classification models for wood knots.

of the index values calculated by the LBP operator in all the regions of the image. LBP descriptors from images with various radii (r) of the LBP operator and rotation-invariant uniform (RIU)-LBP descriptors were generated. The RIU-LBP is a combination of rotation-invariant LBP, which is a method of considering patterns that become identical by rotation as one, and a uniform LBP, a method that considers patterns that change from 0 to 1 or 1 to 0 three or more times in an encoded value. The RIU-LBP has only a 10-bins histogram.

2.3.3 Histogram of oriented gradients: HOG divides an image into cells of a certain size, and creates a histogram of the direction of edge pixels with a gradient magnitude over a specific value for each cell (Figure 3c). The HOG descriptor for image representation was generated by connecting the histograms of all cells. In other words, HOG is a histogram template for the direction of the edges in the image. Because HOG extracts contour information of objects, it is suitable for classifying objects with unique contours without complex patterns. To

generate HOG descriptors with the same dimensions from all images, the images were resized to 50×50 or 100×100 pixels, and classification models were built for each size. In addition, the classification performance of various cell sizes (4×4 to 16×16 pixels at intervals of 2 pixels) was investigated to determine the optimal cell size.

2.3.4 Scale-invariant feature transform: Unlike other feature extraction algorithms used in this study, SIFT, a local feature extraction technique, is an algorithm that finds blob and corner-based keypoints (also called interest points) in an image and describes the gradient orientations of their neighboring pixels as a 128-dimensional vector (Figure 3d). SIFT has proven its superiority, not only in general image classification, but also in wood identification (He et al. 2015; Hwang et al. 2020; Martins et al. 2013). The number of layers in each octave of 3, the contrast and edge thresholds of 0.04 and 10, respectively, and the sigma value of Gaussian applied to the image at the octave number zero of 1.6 were applied as parameters of the SIFT algorithm for feature extraction.

2.3.5 Dense scale-invariant feature transform: DSIFT, a dense version of SIFT, omits the keypoint detection process of SIFT. This computes the descriptors for densely sampled keypoints with the same size, orientation, and regular intervals (Figure 3e). DSIFT descriptors were generated from wood knot images by varying the step size (2–14 pixels at intervals of 2 pixels), which is the distance between the DSIFT keypoints.

2.4 Data learning

2.4.1 Classification models: An artificial neural network (ANN) was employed to build a classification model by learning the features extracted from the training set, and models based on k -nearest neighbor (k -NN) and support vector machine (SVM) were also established for performance comparison. All the models were validated by stratified three-fold cross-validation on the training.

For the ANN, a multi-layer feed-forward network with back-propagation was used for classification. The rectified linear unit (ReLU) was used as an activation function, and the cross-entropy was used as a loss function. The model optimized the loss function using stochastic gradient descent-based solvers SGD and Adam. Various ANN architectures having one and two hidden layers with different nodes were tested to find the optimal network configuration. The initial learning rates were tested as 0.0001, 0.001, 0.01, and 0.1, respectively, and the maximum number of iterations was set to 1000. The ANN architectures, solvers, and learning rates were optimized using grid searches. The classification model was determined based on the minimum loss achieved by the solver and learning rates.

k -NN, a minimum distance-based algorithm, classifies images using the distance between the query data and training data without learning the training data. The number of nearest neighbors (k) was set to odd numbers in the range of 1–15, and the optimal k was determined using a grid search. SVM is an algorithm that finds a decision boundary with the maximum margin between classes, and performs linear classification. A radial basis function (RBF) kernel is used to determine the hyperplane by projecting data onto a high-dimensional

feature space (Vert et al. 2004). The cost (C) and gamma parameters of the RBF kernel SVM were optimized using grid searches with a logarithmic grid from 10^0 to 10^5 for C and from 10^{-1} to 10^{-6} for gamma. C and gamma are parameters that control the cost of misclassification of the training data and a Gaussian kernel for nonlinear classification, respectively.

2.4.2 Evaluation metric: Because the wood knot dataset has imbalanced classes (Table 1), the F1 score was used to evaluate the classification performance of the models. In the classification of imbalanced datasets, accuracy as a performance metric often provides biased results due to oversampled classes. Therefore, the F1 score, the harmonic mean of precision and recall, is more appropriate than accuracy, and is given by:

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}},$$

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}},$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

3 Results

3.1 Image representation

The image representations of the feature extraction algorithms tested are presented in Figure 4. The decayed knot image in Figure 4a is represented differently by the feature extraction algorithms. GLCM computed the texture properties at four angles and represented the image with the simplest descriptor (Figure 4f). The image converted by the basic LBP operator showed tiny textures like a topographic

Figure 4: Image representations of the feature extraction algorithms for the decayed knot image (a). (b) LBP and (c) HOG representations, and (d) SIFT and (e) DSIFT keypoints computed from the input image. Descriptors of the (f) GLCM, (g) LBP, (h) HOG, (i) SIFT, and (j) DSIFT algorithms representing the input image.

map (Figure 4b), and its descriptor contains various pattern information of the image (Figure 4g).

The HOG feature represents the image as the size and orientation of local edges (Figure 4c), and the histogram of each cell is connected to generate a descriptor (Figure 4h). SIFT detected blob and corner-based keypoints in the image and generated descriptors for regions around the keypoints (Figure 4d and i), whereas DSIFT was densely distributed across the image (Figure 4e and j). As shown in Figure 4, the feature extraction algorithms used in this study represent the images in their own ways.

3.2 Classification performance

3.2.1 GLCM

Figure 5a and b show the F1 scores of the models trained with the GLCM features. The classification performance of models trained with the GLCM feature data, which are

linearly connected with the texture properties calculated at the four angles, was higher than that of the models trained with the rotation-invariant GLCM feature data, which is the average of the four angles. In both cases, the ANN models achieved higher performance than the others, and the F1 score of 0.793, the best performance in this study, was produced from the ANN model trained with GLCM features that combines the four angles at an inter-pixel distance of 5.

The average values of the five GLCM texture properties, dissimilarity, contrast, homogeneity, energy, and correlation, are presented in Figure 6. In the decayed and encased knots with distinct boundaries, contrast and dissimilarity values, which are texture properties related to contrast between neighboring pixels, were higher than others, whereas homogeneity values, the inverse of contrast, were relatively low. The spike knots with large changes in appearance had low energy values related to regularity between pixels, and the encased knots had low correlation values indicating a linear relationship between pixels. The GLCM texture histogram of correctly classified images

Figure 5: F1 scores of wood knot classification models using texture features computed by GLCMs. Classification performances of models that each learned the GLCM features of the four angles (a), the average of the four angles (b), the basic LBP (c), and rotational-invariant uniform LBP descriptors (d).

Figure 6: Average values of the five GLCM texture properties, dissimilarity, contrast, homogeneity, energy, and correlation, for training (a), correctly classified (b), and misclassified images (c).

(Figure 6a) resembled those of the training images (Figure 6b), whereas those of misclassified images had remarkably different patterns (Figure 6c).

The relatively low F1 scores produced from the rotation-invariant GLCM could be the result of a lack of discriminative power owing to its descriptor dimension (5-dimensional vector) lower than the descriptor combining the four angles (20-dimensional vector). If other texture properties are added, it is expected to improve the performance of the rotation-invariant feature; however, the increase in computational cost must be considered in this case.

3.2.2 LBP

The basic LBP operator producing 256-dimensional vectors from the image achieved higher F1 scores (Figure 5c) than the RIU-LBP operator producing 10-dimensional vectors (Figure 5d). In both LBP operators, the performance of SVM and ANN models was higher than that of k -NN, and the ANN model trained with the basic LBP descriptor with a radius of 7 achieved the highest performance with an F1

score of 0.744. However, the performance of models trained with LBP descriptors was lower than that trained with the GLCM feature set. In machine learning-based automated wood identification studies, it has been reported that the discriminative power of LBP is generally higher than that of GLCM (Cavalin et al. 2013; Hu et al. 2015; Martins et al. 2013; Prasetyo et al. 2010). The relatively low classification performance of LBP was attributed to the simple patterns of wood knot images. LBP extracts patterns such as flats, edges, and corners from objects in an image for image representation (Ojala et al. 2002), whereas GLCM simply represents an image as one value for each texture property.

3.2.3 HOG

Classification models trained with HOG descriptors computed from 100×100 -pixel images produced higher F1 scores (Figure 7b) than those trained with 50×50 -pixel images (Figure 7a). In the classification of 50×50 -pixel images, SVM with the pixels per cell of 8 produced the highest F1 score of 0.742, whereas in the classification of

Figure 7: F1 scores of classification models trained with the HOG descriptors extracted from wood knot images of 50×50 (a) and 100×100 pixels (b), the SIFT (dashed lines in c), and DSIFT descriptors (solid lines in c).

100×100 -pixel images, ANN with the pixels per cell of 14 gave the highest F1 score of 0.776 among models tested using the HOG descriptors. The HOG descriptor containing information about the edge of the object performed lower than the GLCM feature set, but higher than the LBP descriptor in the wood knot classification. In contrast to other texture features, the HOG algorithm exhibited a small difference in classification performance among the classifiers tested.

3.2.4 SIFT and DSIFT

The performances of the SIFT and DSIFT algorithms for wood knot classification are presented in Figure 7c. The ANN showed the highest performance in both algorithms, with F1 scores of 0.711 and 0.718 for SIFT and DSIFT, respectively. The average number of SIFT keypoints extracted from wood knot images was 103.4, far less than 936.5 for DSIFT. For the SIFT algorithm, fewer than five keypoints were detected from some small images, which would result in insufficient image representation. In addition, the simple morphological structure of the wood

knots may have an effect on the detection of a small number of SIFT keypoints. In contrast, DSIFT could have achieved slightly higher performance than SIFT because it describes all local regions in the image at regular intervals. When the number of SIFT-based features is insufficient, as in wood knot images, the use of DSIFT is worth considering.

3.3 Misclassified knots

Figure 8 shows the confusion matrices of the models that achieved the highest classification performance for each feature extraction algorithm. GLCM and HOG were more effective in classifying sound knots (Figure 8a and c), while SIFT and DSIFT were better in classifying encased knots (Figure 8d and e). The model trained with GLCM features had a particularly low classification performance for spike knots (Figure 8a). The misclassified spike knot images showed patterns in which dissimilarity and correlation values, which are texture properties representing the contrast and linear relationship of neighboring pixels

		Predicted class			
		Decayed	Encased	Sound	Spike
Actual class	Decayed	0.800	0.060	0.140	0.000
	Encased	0.218	0.636	0.109	0.036
	Sound	0.018	0.018	0.955	0.009
	Spike	0.111	0.000	0.556	0.333

		Predicted class			
		Decayed	Encased	Sound	Spike
Actual class	Decayed	0.640	0.040	0.280	0.040
	Encased	0.145	0.709	0.127	0.018
	Sound	0.080	0.054	0.839	0.027
	Spike	0.000	0.167	0.278	0.556

		Predicted class			
		Decayed	Encased	Sound	Spike
Actual class	Decayed	0.560	0.200	0.180	0.060
	Encased	0.259	0.667	0.074	0.000
	Sound	0.027	0.036	0.929	0.009
	Spike	0.056	0.000	0.111	0.833

		Predicted class			
		Decayed	Encased	Sound	Spike
Actual class	Decayed	0.400	0.120	0.440	0.040
	Encased	0.091	0.782	0.127	0.000
	Sound	0.036	0.063	0.875	0.027
	Spike	0.056	0.000	0.389	0.556

		Predicted class			
		Decayed	Encased	Sound	Spike
Actual class	Decayed	0.420	0.160	0.380	0.040
	Encased	0.091	0.782	0.127	0.000
	Sound	0.027	0.071	0.875	0.027
	Spike	0.111	0.000	0.222	0.667

Figure 8: Confusion matrices for models achieved the best classification performance for each feature extraction algorithm. (a) GLCM, (b) LBP, (c) HOG, (d) SIFT, and (e) DSIFT.

(Hall-Beyer 2017), respectively, differed from the training data (Figure 6). This result suggests that the spatial relationship of neighboring pixels is limited in representing spike knots with distinct appearances in elliptical shapes.

Decayed and spike knots showed a relatively low recall for all algorithms. Decayed knots were more likely to be misclassified as encased or sound knots, whereas spike knots were more likely to be misclassified as sound knots. The poor recall of decayed and spike knots might be because they shared many common features with other knot types and/or some of them had ambiguous knots that were difficult to classify into a particular class. The lack of training data for the spike knot class would also have affected the poor classification. In addition, the fact that there were many small knots, namely small images, in the decayed knot class may have affected the poor classification. Spike knots, which were generally poorly classified in texture features, achieved a relatively high recall in local features. This appears to be due to the morphological feature of the elliptical shape that differentiates the spike knot from others with a circular shape, and the relatively low recall in SIFT; in contrast, this appears to be due to the small number of SIFT keypoints detected from the spike knots.

4 Discussion

The performance of the texture descriptors, including GLCM and LBP, was superior to that of the local feature descriptors, SIFT and DSIFT (Table 2). However, in comparative studies of texture and local features for wood identification using computer vision and machine learning, the performance of local features such as SIFT and speeded up robust feature (SURF) was higher than those of texture features such as GLCM and LBP (Hu et al. 2015; Martins et al. 2013). The conflicting results produced from knot classification and wood identification were attributed to the differences in image scales and structural complexities of the objects. In computer vision-based wood identification, texture features are mainly applied to macroscale images such as macroscopic images, stereograms, and computed tomography images, whereas local features are applied to microscale images such as micrographs. As shape-related information of wood cells can be obtained on a microscale, local features for extracting morphological features could be more appropriate. On the other hand, on a macroscale such as wood knot images, it is presumed that the texture features were appropriate because the unique textures and patterns revealed by each type were distinctive characteristics. In addition, the

Table 2: Performance comparison of strategies tested for wood knot classification.

Feature		Feature parameters	Dsc_dim	CLS	CLS_params	F1 score
Texture	GLCM	<i>prop</i> : dis, con, hom, ene, cor θ : 0°, 45°, 90°, 135° (Rotation-invariant)	20	ANN	<i>lr</i> = 0.01 solver = Adam	0.793
			5	ANN	<i>lr</i> = 0.1 solver = Adam	0.719
	LBP	(Basic method) $r = 7, p = 8$	256	SVM	cost = 10 gamma = 0.01	0.744
		(RIU method) $r = 6, p = 8$	10	ANN	<i>lr</i> = 0.001 solver = Adam	0.613
Local feature	HOG	(50 × 50 pixels) <i>ppc</i> 8, <i>cpb</i> 2 × 2	800	SVM	cost = 10 gamma = 0.01	0.742
		(100 × 100 pixels) <i>ppc</i> 14, <i>cpb</i> 2 × 2	1152	ANN	<i>lr</i> = 0.001 solver = Adam	0.776
	SIFT	nOctaveLayers 3, CT 0.04, ET 10, sigma 1.6	128	ANN	<i>lr</i> = 0.0001 solver = Adam	0.711
	DSIFT	Step size 8	128	ANN	<i>lr</i> = 0.0001 solver = Adam	0.718

Dsc_dim, descriptor dimension; CLS, classifier; CLS_params, classifier parameters; *prop*, texture property; dis, dissimilarity; con, contrast; hom, homogeneity; ene, energy; cor, correlation; θ , angles; r , radius; p , number of points; RIU, rotation-invariant uniform; *ppc*, pixels per cell; *cpb*, cells per block; CT, contrast threshold; ET, edge threshold; *lr*, learning rate.

structural simplicity of wood knot images might be an important factor in the relative superiority of the texture features.

Notably, GLCM, which has the lowest dimensional feature among the classification strategies tested, achieved the best performance with an F1 score of 0.793 (Table 2). Because GLCM is an algorithm that computes textures from an image based on the intensity of adjacent pixels with their spatial relationship, it seems to give higher performance than LBP, a descriptor that only represents an image.

Comparing HOG and SIFT, which are methods of using gradient orientation information locally, because HOG is a type of template matching, it is difficult to classify when the object is rotated or changed in shape. By contrast, SIFT, which detects rotation-invariant features in a Gaussian pyramid, is robust to rotation and changes in the shape and size of objects (Lowe 2004). These algorithmic properties suggest that HOG is suitable for the classification of object with small shape changes, simpler patterns, and unique contours, such as wood knot images, whereas SIFT is suitable for objects with complex features and patterns such as cross-sectional micrographs.

Images containing small knots, mostly decayed and encased knots, were difficult to recognize because their textural or morphological characteristics were not sufficiently revealed. The average size of decayed and encased knot images was only 71% and 36% of that of sound knot images, respectively, and the size of misclassified images among them was 13% and 14% smaller than that of

correctly classified images, respectively. As shown in Figure 9, as the pixel resolution of the image dataset decreased, the F1 score of the model continued to decrease. Therefore, to improve classification performance, it is necessary not only to increase the image dataset, but also to improve the image quality. To further improve the classification performance, multiple feature sets in which different types of features are combined should be considered. The multiple feature set combined with GLCM

Figure 9: Effect of image pixel resolution on F1 score of models trained with GLCM features. The values in parentheses are the reduced ratio for the original size.

and LBP gave higher classification performance than the single feature set (Mahram et al. 2012), and the feature set combined with SURF and LBP improved the detection performance of wood knots (Hittawe et al. 2015).

Among the classification models tested, the ANN provided the best performance. The performance of ANN is sensitive to its network configuration (Nasir and Cool 2020). The F1 scores produced from ANN architectures tested are presented in Figure 10. ANN models trained with GLCM (Figure 10a) and DSIFT (Figure 10e) descriptors were the most sensitive to network configuration, with differences between the lowest and highest F1 scores of 9.2% and 9.6%, respectively, depending on the architecture, whereas the other models showed differences of less than 5%. Because of the different descriptor dimensions for each feature, the number of nodes in the input and hidden layers of ANNs was configured differently, but the best F1 score for each feature was produced from the architecture with two hidden layers in all models (Figure 10). In comparative studies of classifiers for wood image recognition, it has been reported that ANN models generally achieve higher recognition performance than

other classifiers, including SVM and k -NN models (Hu et al. 2015; Prasetyo et al. 2010; Tou et al. 2008; Yadav et al. 2014). These results suggest that the features extracted from wood knot images have a non-linear relationship.

5 Limitations and future studies

This study dealt with the classification of wood knot defects using manually extracted features. The performance of models trained with the imbalanced dataset could not avoid the bias towards the dominant class. Moreover, because the dataset consists only of wood knot images, various defects affecting wood quality such as checks, splits, and barks were not covered. These limitations should be overcome through sampling techniques as well as the increase in data. The species specificity of wood defects is also an issue worth investigating.

The relatively low classification performance of models trained with texture and local features gave validity to the use of CNNs for wood defect classification. For the industrial application of automated models for lumber

Figure 10: F1 scores obtained from various network architectures of ANN models achieved the best classification performance for each feature extraction algorithm. (a) GLCM, (b) LBP, (c) HOG, (d) SIFT, and (e) DSIFT.

surface quality evaluation, deep learning-based object detection models that encompass detection and classification would be more appropriate. This is because conventional image processing methods for object detection are highly dependent on image quality, binarization thresholds, and manual editing of the operator (Mallik et al. 2011; von Arx et al. 2016).

Future research is focused on the development of deep learning models for lumber quality evaluation. The key to automated quality evaluation is accurate defect detection, classification, and segmentation. To implement these functions, research is ongoing to develop a deep learning-based instance segmentation model and an automated image acquisition module.

6 Conclusions

For the automated classification of wood knots, classification models trained with descriptors of texture and local features extracted from macroscopic images of wood knots were constructed. Classification models trained with texture features achieved a higher performance than those trained with local features. Among the feature extraction algorithms investigated, GLCM gave the highest F1 score, which proves that wood knot classification is appropriate for texture classification. The discriminant power of texture features was higher than that of local features, however, encased and spike knots were classified better by local feature-based models. Therefore, it is considered that convolutional neural network models covering a wide range of features from local features to the overall shape of an object are more suitable for wood knot classification. Expanded studies on automated detection and segmentation of wood defects using deep learning-based models to achieve better performance and functionality are in progress and will be reported elsewhere.

Acknowledgments: We would like to thank Editage (www.editage.co.kr) for English language editing.

Author contributions: SH performed the experiments and analyses, and was a major contributor to the writing of the manuscript. TL, HK, HC, and JC built the image dataset and contributed to data analysis. HY designed and supervised the study. All authors have read and approved the final manuscript.

Research funding: This study was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education, Korea (NRF-2018R1D1A1B07050743).

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

References

- Affonso, C., Rossi, A.L.D., Vieira, F.H.A., and de Leon Ferreira, A.C.P. (2017). Deep learning for biological image classification. *Expert Syst. Appl.* 85: 114–122.
- Alapuranen, P. and Westman, T. (1992). Automatic visual inspection of wood surfaces. *Proceedings of the 11th IAPR international conference on pattern recognition*, August 30–September 3, 1992. IEEE, New York, pp. 371–374.
- Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-up robust features (SURF). *Comput. Vis. Image Understand.* 110: 346–359.
- Cavalin, P.R., Kapp, M.N., Martins, J., and Oliveira, L.E. (2013). A multiple feature vector framework for forest species recognition. *Proceedings of the 28th annual ACM symposium on applied computing*, March 18–22, 2013. Assoc. Comput. Mach., New York, pp. 16–20.
- Chen, H., Hu, Q., Zhai, B., Chen, H., and Liu, K. (2020). A robust weakly supervised learning of deep Conv-Nets for surface defect inspection. *Neural Comput. Appl.* 32: 11229–11244.
- Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C. (2004). Visual categorization with bags of keypoints. In: *Workshop on statistical learning in computer vision, ECCV 2004*, May 11–14, 2004. Springer, Berlin.
- Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. *Proceedings of 2005 IEEE computer society conference on computer vision and pattern recognition*, June 20–25, 2005. IEEE, New York, pp. 886–893.
- Feio, A. and Machado, J.S. (2015). In-situ assessment of timber structural members: combining information from visual strength grading and NDT/SDT methods – a review. *Construct. Build. Mater.* 101: 1157–1165.
- Gu, I.Y., Andersson, H., and Vicen, R. (2008). Automatic classification of wood defects using support vector machines. In: Bolc, L., Kulikowski, J.L., and Wojciechowski, K. (Eds.), *Computer vision and graphics, ICCVG 2008, Lecture Notes in Computer Science*, Vol. 5337. Berlin: Springer.
- Hall-Beyer, M. (2017). GLCM texture: a tutorial v. 3.0 March 2017. Available at: <https://prism.ucalgary.ca/handle/1880/51900>.
- Haralick, R.M., Shanmugam, K., and Dinstein, I.H. (1973). Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.* 6: 610–621.
- He, T., Liu, Y., Xu, C., Zhou, X., Hu, Z., and Fan, J. (2019). A fully convolutional neural network for wood defect location and identification. *IEEE Access* 7: 123453–123462.
- He, T., Liu, Y., Yu, Y., Zhao, Q., and Hu, Z. (2020). Application of deep convolutional neural network on feature extraction and detection of wood defects. *Measurement* 152: 107357.
- Hittawe, M.M., Muddamsetty, S.M., Sidibé, D., and Mériaudeau, F. (2015). Multiple features extraction for timber defects detection and classification using SVM. *Proceedings of the IEEE international conference on image processing (ICIP 2015)*, September 27–October 1, 2015. IEEE, New York, pp. 427–431.
- Hu, S., Li, K., and Bao, X. (2015). Wood species recognition based on SIFT keypoint histogram. *Proceedings of the 2015 8th*

- international congress on image and signal processing (CISP), October 14–16, 2015. IEEE, New York, pp. 702–706.
- 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. *J. Wood Sci.* 66: 16.
- Jung, S.-Y., Tsai, Y.H., Chiu, W.Y., Hu, J., S., and Sun, C.T. (2018). Defect detection on randomly textured surfaces by convolutional neural networks. *Proceedings of the 2018 IEEE/ASME international conference on advanced intelligent mechatronics (AIM)*, July 9–12, 2018. IEEE, New York, pp. 1456–1461.
- Kamal, K., Qayyum, R., Mathavan, S., and Zafar, T. (2017). Wood defects classification using laws texture energy measures and supervised learning approach. *Adv. Eng. Inf.* 34: 125–135.
- Kauppinen, H. and Silvén, O. (1996). The effect of illumination variations on color-based wood defect classification. *Proceedings of international conference on pattern recognition (ICPR 1996)*, August 25–29, 1996. IEEE, New York, pp. 828–832.
- Kim, H., Kim, M., Park, Y., Yang, S.-Y., Chung, H., Kwon, O., and Yeo, H. (2019). Visual classification of wood knots using k-nearest neighbor and convolutional neural network. *J. Korean Wood Sci. Technol.* 47: 229–238.
- Kobayashi, K., Akada, M., Torigoe, T., Imazu, S., and Sugiyama, J. (2015). Automated recognition of wood used in traditional Japanese sculptures by texture analysis of their low-resolution computed tomography data. *J. Wood Sci.* 61: 630–640.
- Lampinen, J., Smolander, S., Silven, O., and Kauppinen, H. (1994). Wood defect recognition: a comparative study. In: *Workshop on machine vision in advanced production*, Available at: <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.46.8601&rep=rep1&type=pdf>>.
- Lampinen, J., Smolander, S., and Korhonen, M. (1998). Wood surface inspection system based on generic visual features. In: Soulie, F.F. and Gallinari, P. (Eds.), *Industrial applications of neural networks*. Singapore: World Scientific, pp. 35–42.
- Lens, F., Liang, C., Guo, Y., Tang, X., Jahanbanifard, M., da Silva, F.S.C., Ceccantini, G., and Verbeek, F.J. (2020). Computer-assisted timber identification based on features extracted from microscopic wood sections. *IAWA J.* 41: 660–680.
- Liu, C., Yuen, J., and Torralba, A. (2010). Sift flow: dense correspondence across scenes and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.* 33: 978–994.
- Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60: 91–110.
- Mahram, A., Shayesteh, M.G., and Jafarpour, S. (2012). Classification of wood surface defects with hybrid usage of statistical and textural features. *Proceedings of the 35th international conference on telecommunications and signal processing (TSP 2012)*, July 3–4, 2012. IEEE, New York, pp. 749–752.
- Mallik, A., Tarrío-Saavedra, J., Francisco-Fernández, M., and Naya, S. (2011). Classification of wood micrographs by image segmentation. *Chemometr. Intell. Lab. Syst.* 107: 351–362.
- Martins, J., Oliveira, L.S., Nisgoski, S., and Sabourin, R. (2013). A database for automatic classification of forest species. *Mach. Vis. Appl.* 24: 567–578.
- Nasir, V. and Cool, J. (2020). Characterization, optimization, and acoustic emission monitoring of airborne dust emission during wood sawing. *Int. J. Adv. Manuf. Technol.* 109: 2365–2375.
- Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn.* 29: 51–59.
- Ojala, T., Pietikäinen, M., and Maenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 24: 971–987.
- Prasetyo, Khalid, M., Yusof, R., and Meriaudeau, F. (2010). A comparative study of feature extraction methods for wood texture classification. *Proceedings of the sixth international conference on signal-image technology and internet-based systems (SITIS 2010)*, December 15–18, 2010. IEEE, New York, pp. 23–29.
- Qayyum, R., Kamal, K., Zafar, T., and Mathavan, S. (2016). Wood defects classification using GLCM based features and PSO trained neural network. *Proceedings of the 2016 22nd international conference on automation and computing (ICAC)*, September 7–8, 2016. IEEE, New York, pp. 273–277.
- Ren, R., Hung, T., and Tan, K.C. (2017). A generic deep-learning-based approach for automated surface inspection. *IEEE Trans. Cybern.* 48: 929–940.
- Ruz, G.A., Estevez, P.A., and Ramirez, P.A. (2009). Automated visual inspection system for wood defect classification using computational intelligence techniques. *Int. J. Syst. Sci.* 40: 163–172.
- Souza, D.V., Santos, J.X., Vieira, H.C., Naide, T.L., Nisgoski, S., and Oliveira, L.E.S. (2020). An automatic recognition system of Brazilian flora species based on textural features of macroscopic images of wood. *Wood Sci. Technol.* 54: 1065–1090.
- Tou, J.-Y., Tay, Y.H., and Lau, P.Y. (2008). One-dimensional grey-level co-occurrence matrices for texture classification. *Proceedings of the international symposium on information technology*, August 26–28, 2008. IEEE, New York, pp. 1–6.
- Urbanas, A., Raudonis, V., Maskeliūnas, R., and Damaševičius, R. (2019). Automated identification of wood veneer surface defects using faster region-based convolutional neural network with data augmentation and transfer learning. *Appl. Sci.* 9: 4898.
- Vert, J.P., Tsuda, K., and Schölkopf, B. (2004). A primer on kernel methods. In: Schölkopf, B., Vert, J.P., and Tsuda, K. (Eds.), *Kernel methods in computational biology*. Cambridge: MIT Press, pp. 35–70.
- von Arx, G., Crivellaro, A., Prendin, A.L., Čufar, K., and Carrer, M. (2016). Quantitative wood anatomy—practical guidelines. *Front. Plant Sci.* 7: 781.
- Xie, Y.H. and Wang, J.C. (2015). Study on the identification of the wood surface defects based on texture features. *Optik* 126: 2231–2235.
- Yadav, A.R., Anand, R.S., Dewal, M.L., and Gupta, S. (2014). Analysis and classification of hardwood species based on Coiflet DWT feature extraction and WEKA workbench. *Proceedings of the 2014 international conference on signal processing and integrated networks (SPIN 2014)*, February 20–21, 2014. IEEE, New York, pp. 9–13.