A Novel Deep Convolutional Neural Network Based on ResNet-18 and Transfer Learning for Detection of Wood Knot Defects

Mingyu Gao,1 Peng Song,2 Fei Wang,3,4 Junyan Liu,3,4 Andreas Mandelis,5,6 and DaWei Qi1

1. Introduction

Wood knot defect detection is an important part in the production of wood products and finally affects the quality of wood products. Rapid detection of wood knot defects on the surface of the wood can effectively improve the qualification rate of wood products. Consequently, it is important to quickly identify the wood knot defects in a short time [1–4]. Although the traditional manual recognition is widely used and accurate, it is still a subjective [5] and inefficient method to identify wood knot defects [6]. With the rapid development of digital image processing and computer vision, artificial intelligence technology can improve the recognition speed and accuracy at a certain extent [7–9]. Among them, deep learning is the most potential method in the field of artificial intelligence.

In recent years, wood knot defect recognition based on the artificial neural network and image analysis processing has been widely studied [10–15]. Because of its simple basic structure, the neural network can fit various data in theory. Because of this, large-scale neural network combination is needed. Due to the limitation of hardware, the current tools are not enough to run this complex network, resulting in its slow evolution. At present, with the development of robots and so on, the demand for computer vision technology based on CNN (convolutional neural network) is gradually increasing. Therefore, the neural network still has a great application value in the future. In the field of wood defect detection, the accurate recognition of wood needs to collect the defect image by camera or X-ray and then recognize it by image processing and artificial intelligence. In order to accurately identify the wood knot defects, image features must be...
extracted first. For example, Lin et al. in 2015 proposed a method to classify wood knot defects by combining the aspect ratio, grayscale, and variance feature extraction method of the back propagation (BP) network [10]. The accuracy of this method can reach 86.67%. In the same year, Mu et al. proposed a wood defect classification method by extracting the perimeter, area, aspect ratio, and mean grayscale value of the defect, combined with the radial basis function (RBF) neural network with the accuracy over 85% [11]. In 2019, Ji et al. proposed a wood defect classification method based on Hu moment invariant feature extraction and a combination of wavelet moment with BP network [12]. The crack identification accuracy of this method can reach 98%. However, due to the shape of flying knot scar and hole being similar, it is easy to induce a misclassification in some cases. Due to the quite unique shape of each wood knot defect, it is difficult and complex to identify the defect by extracting the image features manually [13]. Therefore, a convolutional neural network (CNN) which can automatically learn the wood knot features is needed to replace the complex manual defect feature extraction. In 2019, Liu et al. proposed a CNN based on split-shuffle-residual (SSR) for real-time classification of rubber boards [14]. Comprehensive experiments show that the algorithm is superior than other classification methods and the latest deep learning classification network at that time has an accuracy of 94.86%, but there is still room for improvement. In 2021, a new method based on transfer learning and ResNet-34 convolutional neural network for recognizing wood knot defects was presented by Gao et al. [15]. The experimental results show that the classification accuracy of this method can reach 98.69%. Although both methods are practical, with the increase of network depth, the model parameters become more complex and the amount of calculation becomes

Table 1: Number of datasets.

<table>
<thead>
<tr>
<th>Wood knot</th>
<th>Training dataset</th>
<th>Before data augmentation</th>
<th>Testing dataset</th>
<th>Original dataset</th>
<th>Training dataset</th>
<th>After data augmentation</th>
<th>Validation dataset</th>
<th>Testing dataset</th>
<th>Total dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decayed knot</td>
<td>10</td>
<td>3</td>
<td>3</td>
<td>16</td>
<td>68</td>
<td>25</td>
<td>19</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>Dry knot</td>
<td>41</td>
<td>14</td>
<td>14</td>
<td>69</td>
<td>291</td>
<td>96</td>
<td>96</td>
<td>483</td>
<td></td>
</tr>
<tr>
<td>Edge knot</td>
<td>39</td>
<td>13</td>
<td>13</td>
<td>65</td>
<td>273</td>
<td>91</td>
<td>91</td>
<td>455</td>
<td></td>
</tr>
<tr>
<td>Encased knot</td>
<td>20</td>
<td>6</td>
<td>6</td>
<td>32</td>
<td>136</td>
<td>44</td>
<td>44</td>
<td>224</td>
<td></td>
</tr>
<tr>
<td>Horn knot</td>
<td>21</td>
<td>7</td>
<td>7</td>
<td>35</td>
<td>147</td>
<td>49</td>
<td>49</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td>Leaf knot</td>
<td>27</td>
<td>10</td>
<td>10</td>
<td>47</td>
<td>198</td>
<td>65</td>
<td>66</td>
<td>329</td>
<td></td>
</tr>
<tr>
<td>Sound knot</td>
<td>110</td>
<td>37</td>
<td>37</td>
<td>184</td>
<td>772</td>
<td>266</td>
<td>250</td>
<td>1288</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>268</td>
<td>90</td>
<td>90</td>
<td>448</td>
<td>1885</td>
<td>636</td>
<td>615</td>
<td>3136</td>
<td></td>
</tr>
</tbody>
</table>
larger. To solve these problems and improve the accuracy of the model, a high accuracy wood knot defect detection method based on the convolutional neural network is required.
In this paper, a model based on the attention mechanism and deep transfer residual convolutional neural network structure named ReSENet-18 is proposed to detect wood knot defects. This paper is arranged and structured as follows. Firstly, the dataset of wood knot defects is acquired and preprocessed. Then, the proposed ReSENet-18 model is introduced. A squeeze-and-excitation-basic block (SE-Basic-Block) is added, and the fully connected layer is replaced by a global average pooling layer to adjust the network structure. At the same time, combined with the ideology of transfer learning, the ReSENet-18 network is pretrained on ImageNet. Subsequently, the network is trained and tested by using the dataset of wood knot defects. Finally, based on a benchmark dataset, the test results are compared and analyzed with other deep learning models.

2. Image Processing and Methods

2.1. Dataset. In order to realize the classification and recognition of wood knot defects, firstly, the image information of 448 wood knot defects of spruce trees with seven kinds of knot defects were collected on the website of Computer Laboratory of Department of Electrical Engineering, University of Oulu [16–18] (shown in Figure 1), and made them into a dataset that can simulate the actual use scene of ReSENet-18 model. Then, the preprocessing operations such as image scaling and adding noise were carried out to realize data augmentation. Finally, the dataset was divided into three parts: a training set, a verification set, and a testing set for training, verification, and testing.

2.2. Data Preprocessing and Augmentation. The dataset of wood knot defects with 448 images was divided into a training set, a verification set, and a testing set according to the ratio of 6:2:2, which refers to 268 training images, 90 verification images, and 90 testing images, respectively (Table 1). The powerful generalization ability of the convolutional neural network is based on a large amount of data; thus, the model will induce the overfitting problem when the amount of data is not large enough which greatly limits the generalization ability [19–22]. Data augmentation technology [23, 24] was always used to expand the dataset of wood knot
defects using color digital image processing technology to expand the data set and add it to the original image dataset; the problem of insufficient data can be easily solved.

The preprocessing of wood knot defect images was completed by simulating the change of angle, noise, and color of different tree species. In order to simulate these changes, the images of wood knots were mirrored horizontally, rotated by 180°, and mirrored vertically to simulate different angles of actual images. By the operation of increasing hue by 10, the color of different defects in actual image acquisition is simulated. At the same time, in order to simulate the noise that may appear in the process of image acquisition, appropriate Gaussian noise and salt-and-pepper noise are added to the image of wood knot defect to further enhance the dataset. In this work, the results of the data augmentation are shown in Figure 2. After data augmentation, the size of the dataset of wood knot defects is expanded from 448 images to 3136 images (7 times expansion). The number of training set, verification set, and testing set is 1885, 636, and 615, respectively, which can effectively reduce the overfitting phenomenon of the convolutional neural network during the training phase.

2.3. Improved Deep Convolutional Neural Network- (DCNN-) Based Method

of the residual building block [25, 26] is shown in Figure 3. A kind of short-cut is used to skip the convolutional layer [27]. The input vector and the vector output through the convolutional layer can be added directly [28] and then output through the rectified linear unit (ReLU) activation function. This method can powerfully alleviate the problem of a vanishing gradient or exploding gradient caused by the increase of neural network depth and can eventually improve the recognition accuracy of wood knot defects.

The output of the residual building block is written as follows:

\[ y = F(x) + x, \]

where \( F \) presents the residual function and \( x \) and \( y \) stand for the input and output, respectively.

ResNet-18 consists of 17 convolutional layers, a max-pooling layer with the filter size of \( 3 \times 3 \), and a fully connected layer. A classical ResNet-18 model involves 33.16 million parameters, in which ReLU activation function and batch normalization (BN) are applied to the back of entire convolutional layers in “basic block.” The structure of ResNet-18 is shown in Table 2 [27].

2.3.2. SE-Basic-Block Module. The SE-Basic-Block module has been used during the champion of ImageNet 2017 classification competition [29]. The structure is shown in Figure 4, which mainly includes squeeze and excitation [30]. The input image has the size of \( W \times H \times C \), where \( W \) and \( H \) represent the width and height, respectively, and \( C \) represents the number of channels. The structure of the SE module is uncomplicated and easy to implement. It can be easily embedded into the existing network framework. The SE module mainly studies the correlation between channels, which only increases a small amount of calculation but can achieve better results.
The attention mechanism of the SE module is mainly realized by multiplying the fully connected layer and input vector for feature fusion. Assume that the size of the input image is $H \times W \times C$, after passing through the global pooling layer and the fully connected layer, the input image is stretched to $1 \times 1 \times C$ and then multiplied with the original image to give weights to each channel. In the denoising task, each noise point is given weight, the low weight noise points are removed automatically, and the high weight noise points are retained. During this process, the network running efficiency can be improved, and the recognition accuracy is improved [31]. As shown in Figure 4, by processing the feature map of convolutional, a one-dimensional vector with the same number of channels is obtained as the evaluation score of each channel [32], and then, the score is used for the corresponding channel to get the result.

The SE module can be embedded into the residual basic block of ResNet-18. Figure 5 shows the combined structure of the SE module and residual basic block module.

2.3.3. Transfer Learning of DCNNs. Due to the small size of the data in this experiment and a certain depth of the proposed network (ReSENet-18), it is easy to induce the overfitting problem in the training process, which leads to a poor recognition ability [33, 34]. In this case, the transfer learning is used to pretrain the deep learning model and then retrain for the wood knot defect detection task using the dataset in this study, which can make our model converge rapidly; thus, a lot of training time can be saved. The deep learning model includes a hierarchical architecture with various layers to learn the complex features of images with wood knot defects [35, 36]. Finally, all these layers are connected to the final fully connected layer classifier to obtain the final results. In the transfer learning, ResNet, VGG and AlexNet models have been trained in ImageNet [37], so that the better classification performance of wood knot defects can be achieved with less training time.

2.3.4. Global Average Pooling. The fully connected layer is usually used as a classifier of CNN, but too many parameters of the fully connected layer will increase the calculation amount of the network and thus slow down the training speed and also easily appear the overfitting problem [38]. Global average pooling (GAP) is a global average of all pixels in the feature map of each channel and outputs the output of each feature map [39–41]. GAP directly removes the features of black box in the fully connected layer and gives each channel practical significance; then, the vectors composed of these output features will be sent to the classifier for classification directly [42]. Figure 6 shows the comparison between the fully connected layer and the global average pool layer.

2.3.5. Overall Architecture. ReSENet-18 is a deep neural network based on the residual structure and attention mechanism (Figure 7). The main features of the architecture are described below.

ReSENet-18 network based on deep learning consists of a ResNet-18, a residual basic block and a squeeze-and-excitation (SE) module. The ReSENet-18 model has 22 layers, including 8 parts: Conv1, SE-Basic-Block, Conv2_x, Conv3_x, Conv4_x, Conv5_x, a global average pooling layer, and a fully connected layer. The first part (Conv1) includes a convolutional layer, a batch normalization layer, a ReLU

---

**Table 5: The evaluation index values of network.**

<table>
<thead>
<tr>
<th>Classes</th>
<th>Model</th>
<th>$P$</th>
<th>$R$</th>
<th>$F1$</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horn knot</td>
<td>LeNet-5</td>
<td>70.89%</td>
<td>84.85%</td>
<td>77.24%</td>
<td>1.88%</td>
</tr>
<tr>
<td></td>
<td>AlexNet</td>
<td>98.39%</td>
<td>92.42%</td>
<td>95.31%</td>
<td>0.91%</td>
</tr>
<tr>
<td></td>
<td>VGGNet-16</td>
<td>98.15%</td>
<td>80.30%</td>
<td>88.33%</td>
<td>2.33%</td>
</tr>
<tr>
<td></td>
<td>GoogLeNet</td>
<td>98.25%</td>
<td>84.85%</td>
<td>91.06%</td>
<td>1.81%</td>
</tr>
<tr>
<td></td>
<td>MobileNet V2</td>
<td>90%</td>
<td>95.45%</td>
<td>92.64%</td>
<td>0.55%</td>
</tr>
<tr>
<td></td>
<td>ReSENet-18</td>
<td>97.01%</td>
<td>98.48%</td>
<td>97.74%</td>
<td>0.18%</td>
</tr>
<tr>
<td>Leaf knot</td>
<td>LeNet-5</td>
<td>85.39%</td>
<td>91.2%</td>
<td>88.20%</td>
<td>6.40%</td>
</tr>
<tr>
<td></td>
<td>AlexNet</td>
<td>96.8%</td>
<td>96.8%</td>
<td>96.8%</td>
<td>2.22%</td>
</tr>
<tr>
<td></td>
<td>VGGNet-16</td>
<td>95.33%</td>
<td>98%</td>
<td>96.65%</td>
<td>1.41%</td>
</tr>
<tr>
<td></td>
<td>GoogLeNet</td>
<td>88.93%</td>
<td>99.6%</td>
<td>93.96%</td>
<td>0.30%</td>
</tr>
<tr>
<td></td>
<td>MobileNet V2</td>
<td>96.06%</td>
<td>97.6%</td>
<td>96.82%</td>
<td>1.68%</td>
</tr>
<tr>
<td></td>
<td>ReSENet-18</td>
<td>99.20%</td>
<td>99.20%</td>
<td>99.2%</td>
<td>0.55%</td>
</tr>
</tbody>
</table>
activation function, and a max-pooling layer. The convolutional layer with the kernel size of $7 \times 7$, stride of 2, padding of 3, and the max-pooling layer with the kernel size of $3 \times 3$, stride of 2, padding of 1 were employed. Adding the max-pooling layer helps to reduce the dimensions and the parameters of model, to expand the receptive fields, and to retain important feature information. The second part (SE-Basic-Block) consists of a residual basic block and a squeeze-and-excitation (SE) module. There are two convolutional layers in the second part. The SE module was embedded into the residual basic block to form a SE-Basic-Block. The structure of the SE-Basic-Block module is shown in Figure 5. In the proposed SE-Basic-Block module, two convolutional layers with the kernel size of $3 \times 3$ and the stride of 1 were used. The first convolutional layer is followed by a BN layer and a ReLU activation function, while the second convolutional layer is only followed by a BN layer. As discussed above, the SE module mainly includes two parts. The first is squeeze, which makes the input image global average pooling; then, the feature map is compressed into a $1 \times 1 \times C$ vector. The second is excitation, which is composed of two fully connected layers and two activation functions (ReLU and Sigmoid). The input of the first fully connected layer is $1 \times 1 \times C$, and the output is $1 \times 1 \times C \times 1/r$, where $r$ is a scaling parameter which is used to reduce the number of channels so as to reduce the amount of calculation. The input of the second fully connected layer is $1 \times 1 \times C \times 1/r$, and the output is $1 \times 1 \times C$. In this paper, $r = 16$ is used. After getting the vector of $1 \times 1 \times C$, the initial feature map and the vector of $1 \times 1 \times C$ will be scaled. The size of the original feature map is $W \times H \times C$, the weight value of each channel output by the SE module is multiplied by the two-dimensional matrix of the corresponding channel of the original feature map, and the final output result is obtained. Parts three to six (Conv2_x, Conv3_x, Conv4_x, and Conv5_x) are shown in Figure 2. The seventh part (global average pool) uses AdaptiveAvgPool function, and the output size of this layer was set to $1 \times 1$. The eighth part (fully connected layer) is the classifier of ReSENet-18. Its output was set to 7, which corresponds to the types of datasets to train and classify.

ReSENet-18 takes RGB image with the random size as input, and then, the image is adjusted to $85 \times 85$ in batch. The input layer of ReSENet-18 is followed by a series of convolutional blocks and a subsampling layer. The CNN structure used in this paper is a variant of ResNet-18, and the feature extraction part of this network is similar to ResNet-18. We used 17 convolutional layers of ResNet-18 to self-study the features of input RGB images from low to high. With the deepening of convolutional layers, the resolution of feature map is reduced, and more abstract high-level
features are extracted. Then, inspired by the success of SENet, we use a SE-Basic-Block module to improve the receptive field. The SE-Basic-Block was embedded between Conv1 and Conv2_x. To improve the sensitivity of the model to channel features, the original features are recalibrated in the channel dimension, so that the model can automatically learn the importance features of different channels. Then, the fully connected layer is replaced by the global average pooling layer to reduce the training parameters, to accelerate the convergence speed of the model, and to improve the classification accuracy of the model.

Finally, we use transfer learning in ReSENet-18 to maximize the collected data and prevent overfitting problem. During the training phase, the weights of 17 convolutional layers except SE-Basic-Block are initialized by the pretraining model of ResNet-18. In this paper, all the parameters would

![Figure 12: Prediction results of the ReSENet-18 model and its component models.](image)

Table 7: Wood knots classification results of ReSENet-18 and its component networks.

<table>
<thead>
<tr>
<th>Method</th>
<th>Actual category</th>
<th>Decayed knot</th>
<th>Dry knot</th>
<th>Edge knot</th>
<th>Predict category</th>
<th>Encased knot</th>
<th>Horn knot</th>
<th>Leaf knot</th>
<th>Sound knot</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ReSENet-18 (our method)</td>
<td>Decayed knot</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Dry knot</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Edge knot</td>
<td>0</td>
<td>0</td>
<td>91</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Encased knot</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Horn knot</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leaf knot</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>65</td>
<td>1</td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sound knot</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>248</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResNet-18</td>
<td>Decayed knot</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Dry knot</td>
<td>4</td>
<td>86</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Edge knot</td>
<td>0</td>
<td>2</td>
<td>89</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Encased knot</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Horn knot</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>1</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leaf knot</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>42</td>
<td>2</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sound knot</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>241</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENet</td>
<td>Decayed knot</td>
<td>16</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Dry knot</td>
<td>0</td>
<td>95</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Edge knot</td>
<td>0</td>
<td>1</td>
<td>89</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Encased knot</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Horn knot</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Leaf knot</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>65</td>
<td>0</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sound knot</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>247</td>
<td>250</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 12: Prediction results of the ReSENet-18 model and its component models.](image)
not be frozen. After loading the pretraining weights, the current dataset of wood knot defects was used to retrain the whole model, which can not only improve the accuracy and speed of training but also improve the recognition ability of the model in the current dataset of wood knot defects. This is very important for effective and stable feature learning.

2.4. Training. The proposed ReSENet-18 was used and trained on one GPU (GTX 960M 2G). The experimental environment is presented in Table 3. The parameter configuration is shown in Table 4. The model using the Adam optimization algorithm and the cross-entropy loss function was trained for 200 epochs, whose batch size is 128 and learning rate is 1e-4.

The flow diagram of the detection process of wood knot defects is shown in Figure 8. First, the images of knot defects were collected from logs. The original datasets were classified by experienced professionals according to the types of defects. Then, the datasets were divided into a training dataset, a verification dataset, and a testing dataset. Subsequently, the proposed ReSENet-18 model was trained on the dataset of wood knot defects. Finally, the model is used to detect the defect types of each image in the testing dataset.

Figure 9 shows the process of training the model using the training and validation datasets. The best accuracy is 99.062%, the best loss is about 0.044, and the overall accuracy in the test phase is about 99.02%.

3. Experimental Results and Discussion

3.1. Comparisons of Model Performance. To evaluate the performance of the proposed model, the dataset was randomly divided and trained 10 times in our case. The classification accuracy of these 10 models is 99.02%, 98.20%, 98.20%, 98.20%, 98.20%, 98.20%, 98.36%, 97.71%, 97.87%, 99.02%, and 98.20%, respectively. The average classification accuracy of the 10 models is 98.30 ± 0.16% and the variance is 0.40%, which indicates a good stability. Taking the first model with the accuracy of 99.02% as an example, the confusion matrix is established by analyzing the predicted labels and true labels of the testing dataset, as shown in Figure 10. All the correct predictions are on the square of the diagonal.

Figure 10 shows that the recognition accuracy of the model for dry knot, edge knot, and horn knot are 100%. In the testing set, the total number of images is 611. The classification accuracies for decayed knot defect and encased knot defect are both 95%, which is due to the small number of decayed knot images and the quite shape difference of encased knot. For leaf knot defect, the classification accuracy is 98% due to the similarity of geometric features between the horn knot and the leaf knot which is easy to be mixed. The classification accuracy of sound knot is 99% which is due to the largest number of sound knot images.
To evaluate the performance of the ReSENet-18, the precision \((P)\), recall \((R)\), f1-score \((F1)\), and false alarm rate \((FAR)\) were applied for the evaluation shown as follows:

\[
P = \frac{T_{ii}}{T_{ii} + T_{ij}}, \tag{2a}
\]

\[
R = \frac{T_{ii}}{T_{ii} + T_{ji}}, \tag{2b}
\]

\[
FAR = \frac{T_{ij}}{T_{ij} + T_{jj}}, \tag{2c}
\]

\[
F1 = \frac{2 \cdot P \cdot R}{P + R}, \tag{2d}
\]

where \(T_{ii}, T_{ij}, T_{ji},\) and \(T_{jj}\) represent the confusion matrix components.

Table 5 shows the precision, recall, f1-score, and false alarm rate of ReSENet-18 for the seven types of wood knot defect and the other five models for comparison. It can be seen from Table 5 that the four indicators of ReSENet-18 are the best in the recognition of five knots (decayed knot, dry knot, edge knot, encased knot, and horn knot) compared with other five classical CNN models, i.e., LeNet-5, AlexNet, VGGNet-16, GoogLeNet, and MobileNet V2. In the recognition of leaf and sound knots, some indicators of ReSENet-18 are slightly worse than other models. For example, the precision of ResNet-18 is higher than that of LeNet-5 and MobileNet V2, but slightly worse than that of AlexNet, VGGNet-16, and GoogLeNet. Among the other three indicators \((R, F1,\) and \(FAR)\), ReSENet-18 is still the best of the six models. In the recognition of sound knot, recall and false accept rate of ReSENet-18 are slightly worse than GoogLeNet and precision and f1-score are better than GoogLeNet. Compared with LeNet-5, AlexNet, VGGNet-16, and MobileNet V2, all the
indicators of ReSENet-18 are better than them. Based on the above analysis, although there is still room for improvement in a few indicators, compared with the other five methods, it can be seen that ReSENet-18 still has a good performance in the identification of wood knot defects.

3.2. Comparison of the Accuracy of Different Models. The ReSENet-18 model was modified, and the test results are shown in Table 6. The accuracy of without ReLU, BN, or SE Basic-Block models is 84.78%, 96.07%, and 98.53%, respectively. The accuracy of adding a BN layer and a ReLU is 96.24%. The accuracy of adding a logsoftmax classifier and replacing Adam with NLLLoss, adding a ReLU and a fully connected layer, and adding a softmax classifier are 14.08%, 14.08%, and 14.40%, respectively. By comparing the performance of different structures of the ReSENet-18 model, the experimental results show that the ReSENet-18 model has the highest recognition accuracy.

3.3. Convergence and Prediction Accuracy Analysis of ReSENet-18 Network Model

3.3.1. Comparison with ReSENet-18 Model and Its Component Models. As demonstrated above, ReSENet-18 is composed of ResNet-18, SE module, and residual basic block. Therefore, to test the performance of the ReSENet-18 model, it was compared with ResNet-18 (composed of residual basic block) and SENet (composed of SE module). Figure 11 shows the loss curve and accuracy curve of ResNet-18, SENet, and ReSENet-18 during the training phase, which are trained by the wood knot defects dataset. One could learn from Figure 11 that our proposed method has the lowest loss value and highest accuracy compared with ResNet-18 and SENet.

![Figure 16: Training results of the ReSENet-18 model with and without data augmentation: (a) training loss value and (b) validation accuracy value.](image_url)

![Figure 17: Prediction results of the ReSENet-18 model with and without data augmentation.](image_url)
The classification results of wood knot defects are shown in Table 7, in which the italic entries represent the number of knot defects correctly identified by the corresponding model, and bold entries represent the total number of the wood knot defects. From Table 7 and Figure 12, ResSENet-18 has the best recognition effect among seven kinds of wood knot defects. Compared with other networks, the ResNet-18 network is relatively shallow in depth, so some degrees of underfitting phenomenon might appear, which leads to a low accuracy on the testing set. The SENet network has the largest number of layers among the three networks, but it can be seen from Table 7 and Figure 12 that the result of identification of SENet is not the best among them due to the increase of network layers and the appearance of overfitting phenomenon. The accuracy of the proposed model with lightweight of ResNet-18 in the testing dataset reaches 99.02% (Figure 12). At the same time, it can combine the features of channel into the network, which improves the feature extraction ability.

Based on the above analysis, ResSENet-18 has been proved to have the highest accuracy and fastest convergence speed in the wood knot defect dataset than other models.

### 3.3.2. Comparison with Classical CNN Model.

Figure 13 shows the training of ReSENet-18 and other five CNN models which was mentioned in Section 3.1. These networks are trained through the dataset of wood knot defects. It can be seen that the ReSENet-18 network has the highest accuracy and the fastest convergence speed than other models on the wood knot defect dataset.

Table 8 compares the number and accuracy of the six network models on the testing dataset. The results show that the LeNet-5 model has the minimal training parameters, which may lead to the underfitting of the network which leads to the lowest accuracy. The parameters of GoogLeNet and MobileNet V2 models are slightly more than LeNet-5, but they are more complex than LeNet-5. It can be seen from Table 8 that their accuracy is improved compared with that of LeNet-5. VGGNet-16 has the maximum parameters among the six models, and AlexNet follows. However, the accuracy is lower than that of ReSENet-18 even through the increase of parameters and longer training time. Compared with VGGNet-16, ReSENet-18 is a kind of lightweight network. At the same time, it can use SE-Basic-Block to weight and recalibrate features. It has stronger feature extraction ability and higher accuracy. It can be seen that, compared with other models, we have improved the performance of the ReSENet-18 model by adding appropriate parameters, while maintaining the robustness and efficiency of the model. Among the six models, the recognition accuracy of the ReSENet-18 model is the highest.

### 3.3.3. Transfer Learning.

A pretraining model of ResNet-18 which includes 1.2 million color images and 1000 categories is used in this study. The weight of the pretrained model is taken as the initial weight of the dataset of wood knot defects. Figures 14 and 15 show the influence and prediction results of transfer learning on the classification accuracy and convergence speed of the ReSENet-18 model. It can be seen that the convergence speed and accuracy of the model have been improved after using transfer learning. The experimental results show that the accuracy of the model with transfer learning is 2.29% higher than that of the model without transfer learning on the testing dataset. Therefore, better convergence can be achieved using transfer learning.
3.3.4. Data Augmentation. Figure 16 shows the convergence and recognition accuracy curves of ReSENet-18 with and without data augmentation. Under the same experimental conditions, ReSENet-18 was trained on 3136 images after data augmentation, and the final classification accuracy in testing dataset reached 99.02% while the accuracy was 80.68% before data augmentation, which is shown in Figure 17.

3.3.5. Comparison of Optimization Algorithms. The optimization algorithm has an important influence on the model performance. In this study, the Adam optimization algorithm is used and compared with SGD, AdaGrad, and AdaMax, which are shown in Figure 18. The results show that the model with Adam algorithm has fastest convergence speed. Table 9 shows the prediction results of these four optimization algorithms under the same environment. The results show that the accuracy in the testing phase is 70.21% for SGD algorithm, 91.98% for AdaGrad algorithm, 95.91% for AdaMax algorithm, and 99.02% for Adam algorithm. It can be seen that the ReSENet-18 model has the best training effect using the Adam optimization algorithm.

3.3.6. Recognition Results of Different Kinds of Wood Knot Defects. “Correct recognition” was used to mark in green, and the “Wrong recognition” was marked in grey in this study. Details of the identification, such as the name and probability of wood knot defects, are displayed next to each label. Figure 19 shows seven wood knot defects and the corresponding identification results.

It can be seen that most of the wood knot defects in the image were correctly identified. Due to the shape of some wood knot defects being similar to other defects, there is no clear feature to extract under this background to induce a few regions incorrectly identified. In addition, the shape of the defect is blurred due to the low resolution of some images, which also makes the extracted features different from those in the training set. In the most cases, our method (ReSENet-18) still has a high accuracy.

4. Conclusions

In conclusion, a novel convolutional neural network model ReSENet-18 is proposed. In the feature extraction part of the network, the SE module is embedded into the residual basic blocks to form SE-Basic-Block. The classifier of the network selects the global average pool to replace the fully connected layer after the convolutional layer at the end to speed up the convergence speed and reduce the model parameters. 2521 images of wood knot defects were used for training after 200 training epochs. Experimental results show that the accuracy of ReSENet-18 in the test phase reaches 99.02%, which is 8.19% higher than the classical ResNet-18 (90.83%). In addition, when various wood knot defects are detected by this method, a large amount of image preprocessing and manual feature extraction are not required, which greatly improves the recognition efficiency. This means that ReSENet-18 has a potential application in wood nondestructive testing and knot defect identification, and it provides a feasible solution for future wood knot defect identification.

Data Availability

The datasets, codes, and weight files used to support the findings of this study are available from the corresponding author upon request.
**Conflicts of Interest**

The authors declare that there are no conflicts of interest regarding the publication of this paper.

**Acknowledgments**

This work was supported by the Project funded by the China Postdoctoral Science Foundation under Grant 2020M670902 and Heilongjiang Postdoctoral Fund under Grant LBH-Z20156. J. Liu acknowledges the Strategic Cooperation Program of the World Top Universities funded by Harbin Institute of Technology. A. Mandelis acknowledged the Natural Sciences and Engineering Research Council of Canada for a Discovery grant.

**References**


[29] H. Ma, G. Han, L. Peng, L. Zhu, and J. Shu, "Rock thin sections identification based on improved squeeze-and-excitation


