Mask Region-Based Convolutional Neural Network (R-CNN) Based Image Segmentation of Rays in Softwoods

Hye-Ji YOO¹ · Ohkyung KWON²† · Jeong-Wook SEO³†

ABSTRACT

The current study aimed to verify the image segmentation ability of rays in tangential thin sections of conifers using artificial intelligence technology. The applied model was Mask region-based convolutional neural network (Mask R-CNN) and softwoods (viz. Picea jezoensis, Larix gmelinii, Abies nephrolepis, Abies koreana, Ginkgo biloba, Taxus cuspidata, Cryptomeria japonica, Cedrus deodara, Pinus koraiensis) were selected for the study. To take digital pictures, thin sections of thickness 10–15 μm were cut using a microtome, and then stained using a 1:1 mixture of 0.5% astra blue and 1% safranin. In the digital images, rays were selected as detection objects, and Computer Vision Annotation Tool was used to annotate the rays in the training images taken from the tangential sections of the woods. The performance of the Mask R-CNN applied to select rays was as high as 0.837 mean average precision and saving the time more than half of that required for Ground Truth. During the image analysis process, however, division of the rays into two or more rays occurred. This caused some errors in the measurement of the ray height. To improve the image processing algorithms, further work on combining the fragments of a ray into one ray segment, and increasing the precision of the boundary between rays and the neighboring tissues is required.

Keywords: instance segmentation, Mask region-based convolutional neural network (R-CNN), rays, quantitative wood anatomy

1. INTRODUCTION

The anatomical features of woods vary with tree species. Therefore, by observing the anatomical features of wood in cross, tangential, and radial sections of the wood under light microscope, different types of wood can be identified (IAWA Committee, 1989, 2004).

The anatomical features of softwoods typically used for wood identification are the presence or absence of intercellular resin canals, bordered pit’s lines or spiral thickening observed in axial tracheids, axial parenchyma cells, and the shape of cross field fits, rays, and aspirated pits (Choi et al., 2022; Eom and Park, 2018; Kwon et al., 2020; Lee and Bae, 2021; Lee et al., 2021a, 2021b; Nam and Kim, 2021; Park et al., 1987). Moreover, the size of wood cells or tissues can also be used to identify wood (Lee et al., 2009; Seo and Eom, 2017; Seo et al., 2014), but such data are insufficient to draw

---

Date Received October 16, 2022, Date Revised November 10, 2022, Date Accepted November 18, 2022

¹ Department of Forest Products, Chungbuk National University, Cheongju 28644, Korea
² National Instrumentation Center for Environmental Management, Seoul National University, Seoul 08826, Korea
³ Department of Wood and Paper Science, Chungbuk National University, Cheongju 28644, Korea

† Corresponding author: Ohkyung KWON (e-mail: zoom@smu.ac.kr, https://orcid.org/0000-0002-6307-0060)
‡ Corresponding author: Jeong-Wook SEO (e-mail: jwseo@chungbuk.ac.kr, https://orcid.org/0000-0002-4395-0570)

© Copyright 2022 The Korean Society of Wood Science & Technology. This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
any inference.

To apply the size of wood tissues to species identification, considerable measurement data are required, which requires a huge amount of time. Also, long working hours cause both physical and mental fatigue, which in turn, may reduce measurement accuracy. To address such issues, automatic wood identification technique based on image processing is gaining popularity in recent years (Kwon et al., 2017; Lopes et al., 2021; Ravindran et al., 2021).

Quantitative wood anatomy involves quantitative analysis of anatomical features of wood and evaluation of its geographical origin and growing condition based on the anatomical features (Gebregeorgis et al., 2021; von Arx et al., 2016). To note, the anatomical structure of wood varies from tree species to species, while external environmental factors influence the size of wood cells (da Silva et al., 2021; Hwang et al., 2020; Kim et al., 2018; Seo and Eom, 2017; Seo et al., 2014). Therefore, based on this, the quantitative data can be used to identify a wood and its geographical origin. The quantitative data can also be used to control the distribution of wood of endangered species in the wood market (de Palacios et al., 2020; Lee et al., 2020; Parades-Villanueva et al., 2018; Savero et al., 2020; Yu et al., 2017).

The size of rays in a tangential section of wood is large enough to be identified under a loupe and is different for different tree species (Alves and Angyalossy-Alfonso, 2000; Burgert and Eckstein, 2001). This signifies that the axial size of the rays in the tangential section can be used as a criterion to identify wood. Furthermore, not too many studies have been done on application of ray size in wood identification.

Deep learning is a part of machine learning that enables automatic measurement, a vast amount of data acquisition and analysis through learning. Therefore, deep learning has the potential to automatically identify wood species efficiently (de Geus et al., 2021; Fabijańska et al., 2021; Fathurahman et al., 2021; Wu et al., 2021).

To apply deep learning in wood identification using size of rays in tangential section, quantitative data on ray size are needed.

The present study was conducted to confirm the detectability of rays in the tangential section of conifers using mask region-based convolutional neural network (Mask R-CNN), which is an instance segmentation model among image segmentation models, and the accuracy of the instance segmentation. This work is expected to contribute to automatic wood identification means and bring improvement in the strategy of using a large amount of quantitative wood anatomical data for wood identification.

2. MATERIALS and METHODS

2.1. Wood samples and their digital image preparation

2.1.1. Wood samples

Softwoods such as P. jezoensis, L. gmelinii, A. nephrolepis, A. koreana, G. biloba, T. cuspidata, C. japonica, C. deodara, and P. koraiensis were selected as experimental samples (Table 1).

Cubes of length 1 cm were cut out of the experimental

<table>
<thead>
<tr>
<th>Family</th>
<th>Scientific name</th>
<th>Common name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ginkgoaceae</td>
<td>Ginkgo biloba</td>
<td>Maidenhair tree</td>
</tr>
<tr>
<td>Pinaceae</td>
<td>Abies koreana</td>
<td>Korean fir</td>
</tr>
<tr>
<td>Pinaceae</td>
<td>Abies nephrolepis</td>
<td>Khingan fir</td>
</tr>
<tr>
<td>Pinaceae</td>
<td>Cedrus deodara</td>
<td>Himalayan cedar</td>
</tr>
<tr>
<td>Pinaceae</td>
<td>Larix gmelinii</td>
<td>Dahurian larch</td>
</tr>
<tr>
<td>Pinaceae</td>
<td>Picea jezoensis</td>
<td>Dark-bark spruce</td>
</tr>
<tr>
<td>Pinaceae</td>
<td>Pinus koraiensis</td>
<td>Korean pine</td>
</tr>
<tr>
<td>Taxaceae</td>
<td>Taxus cuspidata</td>
<td>Rigid-branch yew</td>
</tr>
<tr>
<td>Taxodiaceae</td>
<td>Cryptomeria japonica</td>
<td>Japanese cedar</td>
</tr>
</tbody>
</table>
woods to make cross thin sections for examination under optical microscope. The cubes were softened in a mixture of glycerin and distilled water (1:4) adjusted at 60°C for 1 to 3 days. The softened cubes were cut into sections of thickness 10 to 15 μm using a small sliding microtome (GSL1, WSL), and then stained with a mixture (1:1) of astra blue (0.5%) and safranin (1%) to contrast colorations between cellulose (blue) and lignified cell walls (red). The stained thin sections were dehydrated by sequentially immersing in 30%, 50%, 70%, and 100% ethanol to minimize bubble generation while mounting with a cover glass. Euparal was used for mounting.

2.1.2. Digital images

Digital images of cross, tangential, and radial sections were obtained using a slide scanner (Axio Scan.Z1, Zeiss, Oberkochen, Germany). Magnification of the objective lens was 20× and numerical aperture (NA) was 0.8. The images cropped to 2,048 × 2,048 pixels were used for training the model and the size per pixel was 0.220 μm × 0.220 μm.

2.2. Annotation

The training course for instance segmentation requires an annotation marking the boundary of the target object area. Only tangential section images were taken, and Computer Vision Annotation Tool (CVAT) was used to annotate the rays in the tangential section. The total number of tangential images taken was 524, out of which 400 images were used for training and 124 for verification. The total number of annotations was 633, out of which 427 were used for training and 206 for verification. The annotated images were evenly selected from all tangential images to avoid overestimation of any particular image.

2.3. Test image set

To quantitatively test the performance of the trained model, test images were prepared using the images that had not been used for training purpose. The boundary of the rays in the test images were marked to produce the ground truth (GT) data. The GT data were used to verify the accuracy of the annotation result by CVAT. The accuracy was evaluated as the mean average precision (mAP), which is a quantitative indicator used to select a training model for measuring the ray height.

2.4. Image segmentation and training

In this study, Mask R-CNN model was used for image segmentation. Mask R-CNN is an extended version of Fast R-CNN, one of the existing object detection models, and is used to divide the boundary of an object in the detected area (He et al., 2017). Object detection is a technology to detect the location information of an object in an image and classification of objects in a detected area. On the other hand, instance segmentation is a technique to distinguish individual objects from overlapping of the same type of objects.

The Mask R-CNN model used for training is based on Python, Tensorflow, and Keras (GitHub, 2022). For transfer learning, a weight file trained using MS COCO (Microsoft Common Objects in Context) dataset was used.

The training of the model in the present study were conducted using Tensorflow 1.14.0, Keras 2.1.6, and Python 3.7.5 for segmentation of rays in the tangential section. Computers with Intel Xeon (2.2 GHz, 10 Core/20 Thread, 13.75 MB Cache), 192GB RAM, and Nvidia RTX 2080Ti were used for training.

3. RESULTS and DISCUSSION

3.1. Training of Mask region-based convolutional neural network (R-CNN)

The hyper-parameters and basic values of Mask R-CNN used in this study are as follows.
BACKBONE = resnet50
DETECTION_MIN_CONFIDENCE = 0.7
IMAGE_MIN_DIM = 512
IMAGE_MAX_DIM = 512
LOSS_WEIGHTS: rpn_class_loss = 1.0
LOSS_WEIGHTS: rpn_bbox_loss = 1.0
LOSS_WEIGHTS: mrcnn_class_loss = 1.0
LOSS_WEIGHTS: mrcnn_bbox_loss = 1.0
LOSS_WEIGHTS: mrcnn_mask_loss = 1.0

Among the above factors, the model was trained by varying the values of IMAGE_MIN_DIM and IMAGE_MAX_DIM, which determine the size of the input image, and LOSS_WEIGHTS: mrcnn_mask_loss, which greatly affects the alignment accuracy of the generated mask. The learning rate, momentum, and weight decay were 0.001, 0.9, and 0.0001, respectively. The batch size set in this study was 4, and learning was conducted for 10 to 20 epochs. The input image sizes were 512 × 512 and 1,024 × 1,024.

As a result of training under the above conditions, the highest mAP was resnet50 as a backbone, and 0.837 was found when the input image size was 1,024 × 1,024 (Table 2). The correlation analysis with GT also showed significant results, and it was confirmed that the time taken to derive the results was half that of GT. So, it was verified that the applied model could be used to automatically measure the height of rays.

3.2. Instance segmentation

When the height of the rays was high, there were rays divided into two or more rays and recognized as separate rays (Left in Fig. 1). These were called as fragmented rays in this study. To minimize the decrease in the average height of the rays due to fragmentation, image processing was done to combine the fragmented rays into one. Through image processing, a large number of fragmented rays was corrected to form combined rays (Right in Fig. 1). However, not all fragmented rays could be corrected which needs to be improved in the future.

3.3. Comparison of ray heights by tree species

For G. biloba, T. cuspidata, and C. japonica, there was a difference of more than 100 between the numbers of rays of GT and instance segmentation (Table 3). The average height of the rays in these species did not exceed 100 μm (Fig. 2 the uppermost row). Except for P. koraiensis, the difference in the number of rays from GT and instance segmentation was relatively low. The reason for observing more differences in tree species having rays lower than 100 μm might be poor perfor-

<table>
<thead>
<tr>
<th>Image size</th>
<th>Backbone</th>
<th>Mask loss</th>
<th>Epoch</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>512 × 512</td>
<td>resnet50</td>
<td>1.0</td>
<td>8</td>
<td>0.616</td>
</tr>
<tr>
<td>512 × 512</td>
<td>resnet50</td>
<td>1.0</td>
<td>12</td>
<td>0.673</td>
</tr>
<tr>
<td>1,024 × 1,024</td>
<td>resnet50</td>
<td>1.0</td>
<td>8</td>
<td>0.837</td>
</tr>
<tr>
<td>1,024 × 1,024</td>
<td>resnet50</td>
<td>1.0</td>
<td>12</td>
<td>0.609</td>
</tr>
<tr>
<td>1,024 × 1,024</td>
<td>resnet50</td>
<td>2.0</td>
<td>7</td>
<td>0.777</td>
</tr>
<tr>
<td>1,024 × 1,024</td>
<td>resnet50</td>
<td>2.0</td>
<td>12</td>
<td>0.836</td>
</tr>
</tbody>
</table>

Table 2. Hyper-parameters and mean average precision (mAP) used to train the Mask R-CNN model

**Fig. 1.** Examples of the fragmented rays (left) and corrections of the same rays (right).

**Table 3.** Comparisons of the number of rays from the ground truth and instance segmentation, and the mean heights from the ground truth

<table>
<thead>
<tr>
<th>Wood species</th>
<th>Ground truth</th>
<th>Instance segmentation</th>
<th>Height of rays (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Abies koreana</em></td>
<td>382</td>
<td>344</td>
<td>155.5</td>
</tr>
<tr>
<td><em>Abies nephrolepis</em></td>
<td>320</td>
<td>295</td>
<td>138.9</td>
</tr>
<tr>
<td><em>Cedrus deodara</em></td>
<td>256</td>
<td>243</td>
<td>207.4</td>
</tr>
<tr>
<td><em>Cryptomeria japonica</em></td>
<td>500</td>
<td>400</td>
<td>77.3</td>
</tr>
<tr>
<td><em>Ginkgo biloba</em></td>
<td>447</td>
<td>246</td>
<td>71.4</td>
</tr>
<tr>
<td><em>Larix gmelinii</em></td>
<td>289</td>
<td>293</td>
<td>158.0</td>
</tr>
<tr>
<td><em>Picea jezoensis</em></td>
<td>359</td>
<td>396</td>
<td>134.2</td>
</tr>
<tr>
<td><em>Pinus koraiensis</em></td>
<td>197</td>
<td>85</td>
<td>146.7</td>
</tr>
<tr>
<td><em>Taxus cuspidate</em></td>
<td>501</td>
<td>384</td>
<td>84.1</td>
</tr>
</tbody>
</table>

**4. CONCLUSIONS**

The present study confirmed that Mask R-CNN, trained to detect rays in the tangential section, can be used to measure rays’ height. The highest mAP of the

...
The height of rays measured through Mask R-CNN showed a very high correlation with the results from GT, and the time to obtain the result using the model took less than half of the time required for the GT.

The frequency distribution of the ray height was verified to be unique according to the wood species. Based on these results, it was concluded that statistical values such as average values, skewness, or kurtosis of the height of rays from different wood species can be used as a basis for identification of the wood species. To apply these values, however, more improved object segmentation model and a large number of data on the height of rays are needed.

The major problems in the measurement of the height of rays using Mask R-CNN were the undetected rays and fragmentation of the rays. These two issues caused reduction in the mean ray height in each tree species. The fragmentation mostly occurred in tree species with ray height greater or equal to 100 μm. For more accuracy and reliable results, further research is necessary on improving the detection performance of the trained model and reducing the fragmentation of rays.

**CONFLICT of INTEREST**

No potential conflict of interest relevant to this article was reported.
ACKNOWLEDGMENT

Yoo, H.Y. and Seo, J.W. were supported by the project of ‘Collecting Wood Samples and Establishing the Image Database of their Surfaces and Kwon, O. was supported by the project of ‘Establishment of Standard Database for Wood Identification and Development of Automatic Identification of Wood Species’, which were funded by Korea Forest Service.

REFERENCES


IAWA Committee. 1989. IAWA list of microscopic features for hardwood identification. IAWA Bulletin

