# A Study on tree species discrimination using machine learning in forestry

Tsuya, Kohei ; Fujii, Nobutada ; Kokuryo, Daisuke ; Kaihara, Toshiya ; Sunami, Yasuhiro ; Izuno, Reini ; Mano, Masahito

(Citation)
Procedia CIRP,99:703-706

(Issue Date)
2021-05-03

(Resource Type)
journal article

(Version)
Version of Record

Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 99 (2021) 703–706

www.elsevier.com/locate/procedia

14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME '20

# A Study on tree species discrimination using machine learning in forestry

Kohei Tsuya[a, *], Nobutada Fujii[a], Daisuke Kokuryo[a], Toshiya Kaihara[a]
Yasuhiro Sunami[b], Reini Izuno[b], Masahito Mano[c]

[a]*Kobe University,1-1,Rokkodai,Nada,Kobe,Hyogo,Japan*
[b]*TACHIBANA CONSULTANT,10-3,Kouno,Tachibana,Anan,Tokushima,Japan*
[c] *FIELDCOM.INC,365-2,Hirota hirota,Minami awaji,Hyogo,Japan*

* Corresponding author. Tell.:+81-78-803-6053; *E-mail address:* tsuya@kaede.cs.kobe-u.ac.jp

**Abstract**

As the present state of forestry in Japan, the period of full-scale utilization of planted forests is approaching; consolidation of forest practices aimed at the stable and efficient supply of domestic timber is being promoted. Visualization of forest boundaries is important, however it takes a lot of cost and time because forest surveys are conducted by human's exploration. This study proposes an efficient tree species identification method using deep learning method from aerial images of forests using drones for the purpose of visualizing forest boundaries. The effectiveness of the proposed method is confirmed by experiments using actual images taken by drones.



*Keywords:* Forestry, Machine learning, Convolutional neural network, Drone, Tree species classfication

## 1. Introduction

As the present state of forestry in Japan, the period of full-scale utilization of planted forests is approaching [1]. However, there is a problem of low motivation of forestry workers and inefficiency of the forestry industry. These problems have led to a decline in the forestry population, forestry output, and forest area, and an increase in the amount of abandoned land [2]. Also, the forest size is very small, and 60% of forest owners are said to have no intention of felling. For this reason, forestry managers (such as material producers) have the intention of expanding the scale of their business, but it is difficult to secure a business site [3]. Therefore, consolidation of forest practices is being promoted for the purpose of securing a business site and stable and efficient supply of domestic timber .

Consolidation of forest practices means that the forestry company puts together the forest land of the adjacent owner to carry out the business. Therefore, it is important to understand the forest information that is used by forestry companies to judge the status of forests and the value of consolidation of forest practices. Forest information refers to information such as forest ownership, usage, tree species, tree height, etc. Therefore, forest surveys are being conducted by human's exploration, however it is not progressing to grasp forest information because of the cost and time [4].

Therefore, this study focuses on tree species in forest information, and uses convolutional neural networks (CNN) from aerial images of forests using drones to perform efficient tree species discrimination for the purpose of visualizing forest boundaries. There have already been previous studies on image discrimination using drones and CNN [5]. In the previous study, the training images and the testing images are cropped from a single image taken by drone. However, in practice, it is necessary to evaluate the testing images, which is different from the shooting conditions of training images. Shooting conditions means the location, time and date, weather conditions etc. at the timing of shooting. Because the color of the trees varies depending on the season, sunlight, and other factors due to differences in shooting conditions, general-purpose learning is


10.1016/j.procir.2021.03.094

required. Therefore, this study aims to propose a method that is robust to different shooting conditions by using CNN.

In this paper, as the first step, training and testing images are different images taken on the same condition, and a computer experiment is conducted to verify the effectiveness of the proposed method.

## 2. Proposed method

The overview of the proposed method is shown in Fig. 1.

The proposed method outputs the tree species of the tree image ; the tree images are input to the CNN after image processing and image discrimination is performed.

The proposed method is divided into five steps.

- **Step1** : Image cropping
- **Step2** : Image labeling
- **Step3** : Data Augmentation
- **Step4** : Image size changing
- **Step5** : Image discrimination

The proposed method is also divided into two phases. The first phase is the image processing of Step 1 to 4 and second phase is image discrimination of Step5.

### *2.1. Flow of image processing*

The flow of the image processing is shown below.

#### *2.1.1. Image cropping (Step1)*

One aerial image taken by drone is converted into an ortho image, and as shown in Fig. 2, the tree image is cropped after mesh division into 5 m squares.

#### *2.1.2. Image labeling (Step2)*

As shown in Fig. 3, the cropped tree image is labeled with three types of Hinoki (Japanese cypress), Sugi (Japanese cedar), and other trees. Also, it is divided into training images and testing image at a number of 4: 1 to the label.

#### *2.1.3. Data Augmentation (Step3)*

Data augmentation (DA) is performed to expand the existing images and increase the number of images. In general, many training images are needed to enhance accuracy of image discrimination using machine learning such as CNN [6]. In the proposed method, flip horizontal, image rotation, cutout, scale change, contrast change, and brightness change are performed for a total of six types of expansion, and the training images are inflated to 64 times the number of existing images.

#### *2.1.4. Image size changing (Step4)*

Change the input image size to 75 pixels x 75 pixels for the purpose of reducing the calculation time for image discrimination.

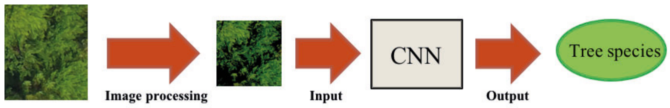


Fig. 1. Overview of the proposed method

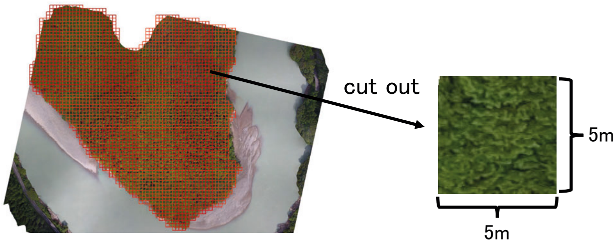


Fig. 2. Image cropping

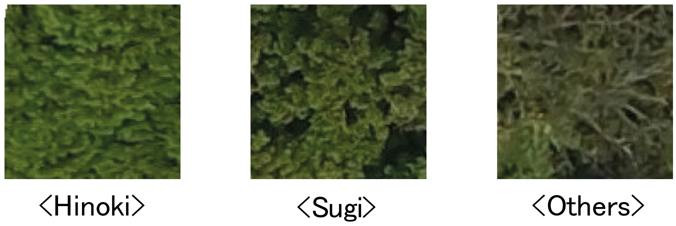


Fig. 3. Image labeling

### *2.2. Image discrimination (Step5)*

The CNN model used in the proposed method is shown in Fig. 4.

As shown in Fig. 4, there are two convolutional layers, two pooling layers and three fully connected layers.

#### *2.2.1. Convolutional layer*

The convolutional layer has the role of extracting the features of the image by performing the convolution operation. The tensor calculated in the convolutional layer is called a feature map. If the input tensor size is $H \times W \times C$, convolutional filter size is $L \times L \times K$, feature map is calculated such as Equation (1).

$$y_{i,j,k} = \sum_{c+1}^{C} \left( \sum_{p+1}^{L} \sum_{q+1}^{L} x_{i+p,j+q,c} \, h_{p,q,c,k} \right) + b_k \qquad (1)$$

The meaning of the letters in Equation (1) are shown below.

- $y_{i,j,k}$ : the value $i$ -th row and $j$ -th column of the $k$ -th feature map
- $x_{i,j,c}$ : the value in the $i$ -th row and $j$ -th column of the $c$ -th channel of the input tensor
- $h_{p,q,c,k}$ : the value in the $i$ -th row and $j$ -th column of the $c$ -th channel of the $k$ -th convolutional filter
- $b_k$ : bias for $k$ -th feature map

The example of the convolutional layer is shown Fig. 5.

In this study, input first tensor size is $75 \times 75 \times 3$, first convolutional filter size is $5 \times 5 \times 16$, and the second convolutional filter size is $5 \times 5 \times 32$.

In addition, the value of the feature map is input to the ReLU function expressed in the following equation (2) and output to the next layer.

$$f(x) = \max(0, x) \tag{2}$$

*2.2.2. Pooling layer*

The pooling layer has the role of making the features extracted by the convolutional layer more emphasize.

In this study, the maximum pooling is used such as Fig. 6 and the maximum pooling filter size is $2 \times 2$.

*2.2.3. Fully connect layer*

The fully connected layer has a role in discrimination which label image the input image is based on the features obtained in the convolutional layer and the pooling layer.

The example of fully connect layers is shown Fig. 7. The circles in the figure represent neurons, which contain image feature. In addition, each neuron is bonded to a neuron in the next layer, and the bond strength is represented by weight.

*2.2.4. Learning algorithm*

In this study, the stochastic gradient descent (SGD) is used as a learning algorithm and update weights to minimize the objective function. The objective function is a loss function represented in Equation (3).

$$E_i = -\boldsymbol{l}_i \log(\boldsymbol{y}_i) \tag{3}$$

The meaning of the letters in Equation (3) are shown below.
q

- $i$ : $i$-th image
- $E_i$ : loss function of the $i$-th image
- $\boldsymbol{l}_i$ : correct label of the $i$-th image
- $\boldsymbol{y}_i$ : output value of the $i$-th image

SGD is a method for randomly selecting one training image from the image data set and iterating the operation of updating each weight using the gradient descent method represented in Equation (4).

$$\boldsymbol{w}_{t+1} \leftarrow \boldsymbol{w}_t - \epsilon \frac{\partial E_i}{\partial \boldsymbol{w}_t} \tag{4}$$

The meaning of the letters in Equation (4) are shown below.

- $i$ : $i$-th image
- $t$ : $t$-th update
- $\boldsymbol{w}_t$ : weight of $t$-th update
- $E_i$ : loss function of $i$-th image
- $\in$ : learning rate

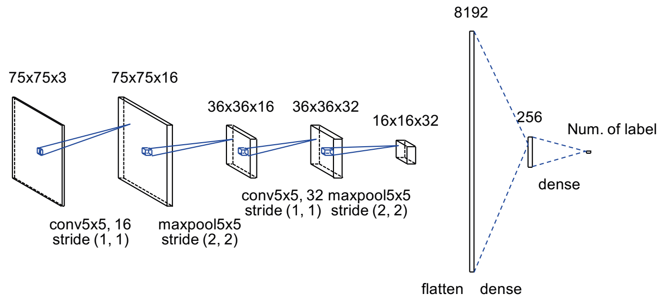


Fig. 4. Model of CNN

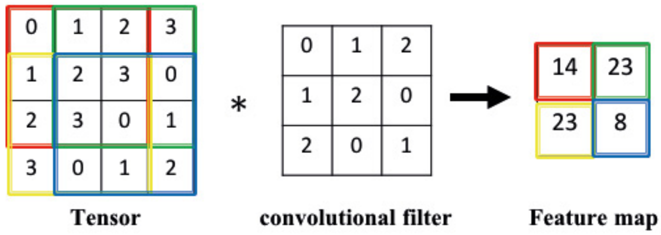


Fig. 5. Example of the convolutional layer

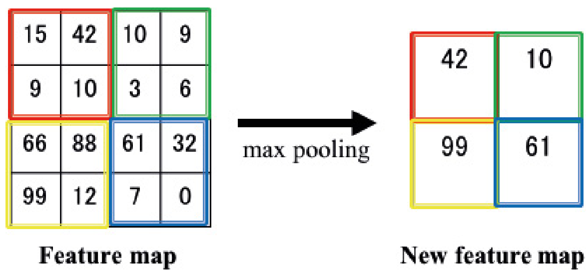


Fig. 6. Example of pooling layer

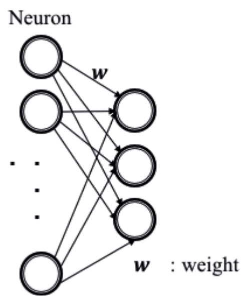


Fig. 7. Example of fully connected layer

## 3. Computational experiment

*3.1. Experimental overview*

Using the proposed method, tree species discrimination is performed for a total of three types of tree images, which are Hinoki, Sugi, and other trees cropped from aerial images of forests using drones.

### 3.2. Experimental conditions

- The learning rates : 0.005
- Batch size : 128
- The number of epochs : 20
- The number of trials : 50
- The number of labels : 3
- The number of training images (without DA) : 1921
- The number of training images (with DA) : 122944
- The number of testing images : 478

Table 1 shows the number of training images and the number of testing images with and without DA for each tree species used in the experiment.

Table 1. The number of images in the dataset used in the experiment

| Tree species | Hinoki | Sugi | Other trees |
|---|---|---|---|
| Num. of training images (No using DA) | 995 | 447 | 479 |
| Num. of training images (Using DA) | 63680 | 28608 | 30656 |
| Num. of testing images | 248 | 111 | 119 |

### 3.3. Experimental result

First, the result for the discrimination rate is described. Table 2 summarizes the mean and standard deviation of the discrimination rates for the 50 trials, and shows the results for 20th epoch.

Table 2. Discrimination result(%)

| Tree species | Hinoki | Sugi | Other trees | Total |
|---|---|---|---|---|
| Ave. | 94.7 | 65.9 | 96.1 | 88.4 |
| S.D. | 1.35 | 4.61 | 1.75 | 1.15 |

From Table 2, the discrimination rate of Hinoki was 94.7%, the discrimination rate of Sugi was 65.9%, the discrimination rate of other trees was 96.1%, and the total discrimination rate of those three tree species is 88.4%. A high discrimination rate was obtained with Hinoki discrimination rate of 94.7% and other trees discrimination rate of 96.1%, but the discrimination rate of Sugi was 65.9%, which was lower than that of Hinoki and other trees. The reason is considered to be because 995 Hinoki teaching images and 447 Sugi teaching images were used, so there were fewer Sugi teaching images compared to Hinoki, and the learning was biased towards Hinoki. However, the number of images of other trees is smaller than that of Hinoki, and the discrimination rate is high at 96.1%. The reason is considered that the discrimination rate of other trees were higher than that of Sugi because it was easier to grasp the features and progressed in learning well.

Next, the number of predictions for each label on a given trial is shown in Table 3. In Table 3, The column shows correct labels of testing images and the row shows the predicted labels of testing images.

Table 3. The number of predictions for each label on a given trial

| | Hinoki | Sugi | Other trees | Discrimination rate(%) |
|---|---|---|---|---|
| Hinoki | 239 | 5 | 4 | 96.3 |
| Sugi | 36 | 70 | 5 | 63.1 |
| Other trees | 8 | 2 | 109 | 91.6 |
| Total | | | | 87.4 |

From Table 3, the number of images of Sugi identified as Hinoki is 36, and the number of images of other trees was 8. Also, the number of images of Sugi identified as other trees is 4, and the number of images of other trees identified as Sugi is 2. Many of the images of Sugi and other trees are identified as Hinoki. This result also shows that learning is biased towards Hinoki, which have many training images.

## 4. Conclusion

In this paper proposed a tree species discrimination method using CNN from aerial images of forests using drones. It is considered that the discrimination of tree species will lead to visualization of forest boundaries. In the computational experiments, it was confirmed that the discrimination rate of Hinoki and other trees was high and the discrimination rate of Sugi was low because of the bias learning towards Hinoki.

In the future, it is necessary to devise image processing and CNN models in order to reduce the bias in the discrimination rate from the tree species. Discrimination method with robustness over the shooting image conditions should be also developed.

## References

[1] I Otsuka, Y Hori, S Yamada, S Iwagana, T Amano, T Komaki, H Mochida. Diverse aspects of the forestry management reorganization — Suggestion for forestry business —. In: Bulletin of FFPRI, 2018; Vol.17, No.3.
[2] H Harayama, A Uemura. The challenge of low-cost forestry for reforestation [Translated from Japanese.] . In: J-STAGE, forestry science, 2017; Vol. 80, pp. 2-5.
[3] Forest Agency. Current status and issues of forest and forestry policy -New forest management system- [Translated from Japanese.]. 2019.
[4] Y Tachiki, T Owari. Applicability of DGPS and LPS to Forest Survey.In : Hokkaido University Collection of Scholarly and Academic Papers, Research Bulletin of the Hokkaido University Forests, 2000, Vol. 57, pp. 95-104.
[5] M Onishi, T Ise. Automatic classification of trees using a UAV onboard camera and deep learning. In arXiv:1804.10390, 2018.
[6] J.Wang and L.Perez. effectiveness of data aug- mentation in image classification using deep learning . arXiv:1712.04621, 2017.