Forest Classification by Tree Species with UAV data and U-Net

Ei Matsui (1), Masahiko Nagai (1), Yumiko Nagai (1), Tsuyoshi Eguchi (1), Shoji Okamoto (1), Yoshiki Mizukami (1)

[1] Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi, Japan
Email: g504wd@yamaguchi-u.ac.jp

KEY WORDS: Forestry, UAV Data, Convolution Neural Network, Deep Learning, Photogrammetry

ABSTRACT:
Classification of tree species is required for estimating forest volume and provides information for forestry workers. In this study, we assess the potential of a U-net, a recent deep learning algorithm. High resolution images (3 cm) from the UAV are used to segment for 2 types of tree conifers species.
As the results, the spatial mapping of tree species can be helped to get borderline of tree species, and deep learning algorithm can support applications for tree species mapping.

1. Introduction

In Japan, the declining timber prices and decreasing the number of forestry workers become problems, thus the unmaintained forests have been increasing. The forest is an important role for preventing environmental problems such as preventing global warming and landslide disasters, recharging water sources and conversing biodiversity (Learning Museum of the Forest and Foresty).

At the stage of forest maintenance planning, the Forest GIS is one of the useful tools for managing forest information, such as forest types, species, and coverages. However, in many cases, since there is no notification from forest owners, forest information is not updated, so the data become far from the real forest condition. In the Forest GIS, each polygon has trees volume estimated by tree age and locational condition. However, trees volume is conventionally estimated by field survey in the case of timber trading, thus the Forest GIS information is not used for foresters in Japan. In addition, the survey for estimating the trees volume must use lots of man powers and be costly. Forest survey conducted by airborne Light detection and ranging (LiDAR) has been receiving attention in recent year because it can measure tree height and tree crown shape (Wannasiri, 2013). However, survey by LiDAR is also costly. In the real forest survey, it is difficult and vast expensive to measure all the tree forms. So, Unmanned Aerial Vehicle (UAV) which can measure tree information such as tree height and crown size can be significant method.

Recently, the development of image recognition technology using deep learning has been developing remarkably. The image recognition technology using Convolution Neural Network has 3 kinds of technical characteristics as followings. First, image classification can be done by adding label to the raw data as supervised data and automatically learn the whole feature by training. Second, object detection can be done by acquiring the position coordinate of the object using recognizing the object in the image (Fujiyoshi, 2017). Third is segmentation which trains the object area of the image each pixel and predicts label per pixel and makes the contour of the object. U-net of segmentation algorism achieves a per pixel classification, predicting the object area of each pixel to belong to single class or Multi classes (LISA lab, 2018) (EyadHaj, 2015), (Falk, 2019), (Wagner, 2019). These techniques can be applied to the data set which normally acquired from remote sensing, satellite, aircraft, UAV. In Japan, deep learning research for target to forest has been started. Freg (2015) resulted with greater precision of the grasping the land coverage changes in urban area by analyzing UAV data using Random Forest and Texture analysis, and verified the effectiveness of the high-resolution UAV data analysis.

Morioka (2019) and Mizoguchi (2016) resulted the tree classification from UAV images and point cloud using deep learning. The tree species classification of these researches has been done manually with Empirical knowledge which is particularly needed to classify mixed forests of tree species.

For tree species classification, individual tree volume can be calculated by using the tree volume equation which is developed by the ground sampling data (Kim's K, M, 2016).

Depending on the tree species, the coefficient of tree volume equation is different, thus the information regarding to the tree species is one of the essential information for tree form estimation. Thus, we consider that U-net is effective for differentiate the tree species precisely. So, in this study we build U-net model, verify the precision of the recognition of conifer tree species and conclude its validity.

2. Study area and methods

2.1 Study area UAV images

The study area is located in Yamaguchi prefecture, Japan. This area is managed by Yamaguchi Prefectural Forestry Guidance Institute. There are some species of trees, e.g. hardwood, bamboo, cedar and cypress in this area. We select 1.6 ha of the forest for the study, because there are various tree species mixed including conifers in the area as shown in Fig.1. The background image of Fig.1 is ortho image made from RGB images processed by SfM (Structure from Motion).

Figure 1: Forest area in study, yellow polygon (background: ortho image)

2.2 UAV images

We use UAV Phantom 4 Pro flight, which summary is described in Table 1.

Forest area has 100 to 120 meters altitude differences from takeoff point to northeast direction. The UAV photography is decided to perform photogrammetry using SfM processing. SfM is a photogrammetric range imaging technique for estimating position of attitude of camera and three-dimensional structures of the object from several images taken by plural photographing points (Jonathan, 2015).

In order to keep higher ground resolution and safe flight standards, the ground altitude for UAV flight plan is set up as 70 meters. While moving the UAV, images are taken by every 2 seconds and both overlap and side wrap are set up as 80%. The photographed angle is directly below the UAV. The photographed images are taken as RGB images with 5,472 x 3,648 pixels.

Table 1: UAV flight summary

UAV	Phantom 4 Pro
Camera model	FC6310
Focus	9 mm
Altitude from ground	70 m
Range horizon	121.1 m
Range vertical	90.8 m
Ground resolution	3 cm
Sampling date	11/Dec/2018
Ortho image resolution	3 cm
Width / Height	5,472 pixel / 3,648 pixel

2.3 Forest mask data

We make Forest mask data as training data for U-net Model in this study.

We use the RGB images as the background and create conifer polygons by visually with Labelbox tool (Labelbox, Inc). The created polygons of tree species are converted to mask image of conifer. Fig.2 shows the polygon from image using Labelbox, which conifer mask include the cedar and cypress class.

Species of conifer are distinguished from crown shape and tree height of them, because conifer crown is sharper and taller than hardwood. There are two species as conifer in this study area. One is cedar whose each leave is looked growing in quantity, which seems to be like one branch. The another in cypress that has flat leaves and it spreads more than cedar from the air (FFPRI, 2019). We distinguish between these species using UAV images based on centimeter ground resolution.

Figure 2: Polygon of tree specie

2.4 Data preprocessing

The training data are created by 12 data sets from UAV images and forest mask data as true data. Images are inputted as original resolution in order to improve accuracy of discrimination. Each UAV image is divided into three sizes, which are 256, 512, and 1,024 pixels in order to increase training data for U-net (Fig.3). And also the divided image is rotated 90, 180 and 270 degrees, as different images. The numbers of each data are shown in Table 2.

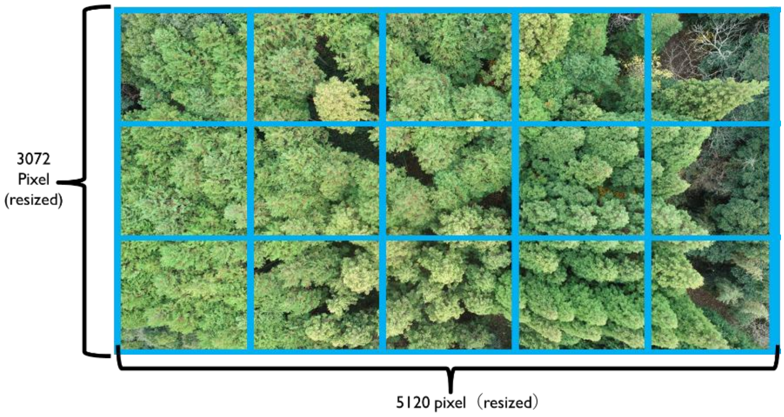


Figure 3: Divided UAV image (example of 1,024 pixels size)

Table 2: The number of each divided images size

Size	Vertical	Horizontal	counts
256	12	20	12,096
512	6	10	2,880
1,024	3	5	720

3. U-Net Model

3.1 Architecture

In this study, we apply single class-segmentation using U-net for the tree species. We use the sigmoid function, which is commonly used as activation function in neural networks for outputting the forest species mask images. U-net model referred by Zhixuhao (2019) can minimize the number of convolution layers and reduce the risk from overfitting. Fig.4 shows the U-net model architecture.

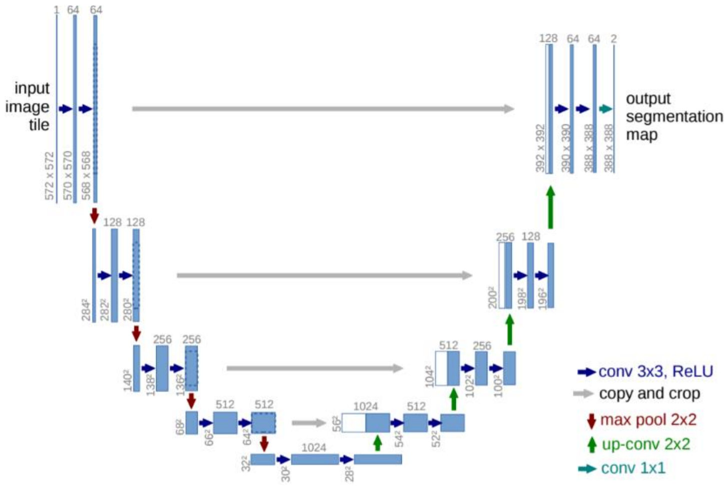


Figure 4: U-net model architecture (Ronneberger, 2015)

The data for training and validation is “forest mask” and “RGB image” as input data, which are divided by 3 types of image size, 256, 512 and 1,024 pixels. 60% of data are applied as training dataset and remains, which are used as validation dataset. The computing resource uses Google Colaboratory. We use python language as processing code. Deep learning flamework is Keras and Tensorflow, which are used for backend. Graphics Processing Unit (GPU) is NVIDIA Tesla K80, and GPU memory is 12GB GPU upper limit. So, we set training max size images per twelve batch.

3.2 Segmentation accuracy assessment

U-net performance is calculated by Intersection over Union (IoU), not by accuracy, for reducing the impact of prediction errors (Preferred Networks & Kikagaku, 2018). The IoU is the object class (eqn 1), which is the number of pixels labeled as object in both the prediction and the reference, divided by the number of pixels labeled as object in the prediction and in the reference. The loss is computed as the percentage of errors classified by pixels.

$$\mathbf{IoU} = \frac{\boldsymbol{Y_{true}} \cap \boldsymbol{Y_{pred}}}{\boldsymbol{Y_{true}} \cup \boldsymbol{Y_{pred}}} \qquad \boldsymbol{eqn}\ \mathbf{1}$$

3.3 Prediction

For prediction, UAV image is divided by 3 sizes, each images position is to create an overlap between the 50% patches. Because of the original size (5,472 × 3,648 pixels), UAV image is combined from divided predicted images, predicted image keeps patch size borderlines. Since U-net has a problem that the prediction accuracy of the patch size boundary is low, the prediction images are overlapped (Ronneberger, 2015). Fig.5 shows the original image and conifer mask image by combining predictions with overlapping patch sizes.

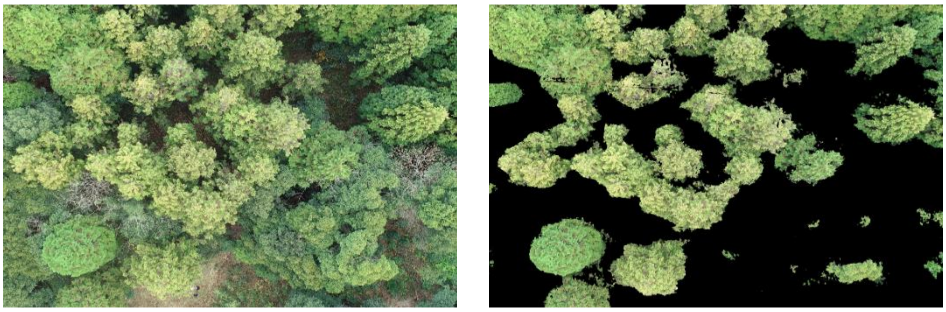


Figure 5: Left: original image, Right: combined predicted image

4. Results

4.1 Model convergence details

In the case of model loss and IoU of convergence, time for convergence is less than 6 hours for making the model.

The best models are obtained after 108 epochs with 12 images set per batch. For the conifer of segmentation, the IoU is 0.96 when image size is 512 × 512 pixels, as shown in Table 3.

Table 3: U-net model convergence result

Size	Epoch	Train	Validation	IoU	Loss
256	113	8,064	4,032	0.72	0.23
512	108	2,016	864	0.96	0.05
1,024	121	576	144	0.96	0.02

4.2 Considerations about predicted images

The values regarding to the both IoU of 512 and 1,024 pixel size are 0.96, and the loss of 512 pixel size resulted as higher than that of 1,024 size. For 512 size, some errors of labeling have been identified. For the hardwood, the main errors appeared when the crown had a similar spectral response and contained less shade due to a highly closed canopy. In Fig.6, the hardwoods has similar to the conifer structure, which are visible in the right bottom of the left original image, but the model results have error point showing by polygon in the right.

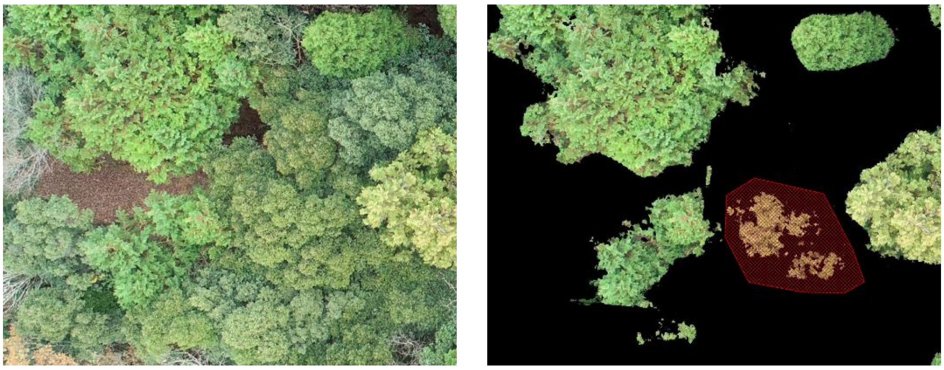


Figure 6: Left: original image,   Right: corresponded prediction image

An example of conifer trees manual segmentation and U-net segmentation is shown in Fig.7. The crowns of conifer trees are sharper than hardwood. The border of the conifer is not sharp in the image and there is a variation between the manual and automatic segmentation. Even though this segmentation can be considered good by visual interpretation, we consider that low accuracy can result from inaccurate manual delineation of forest mask.

Figure 7: Left: visually manual delineation in red polygon, Right: prediction image

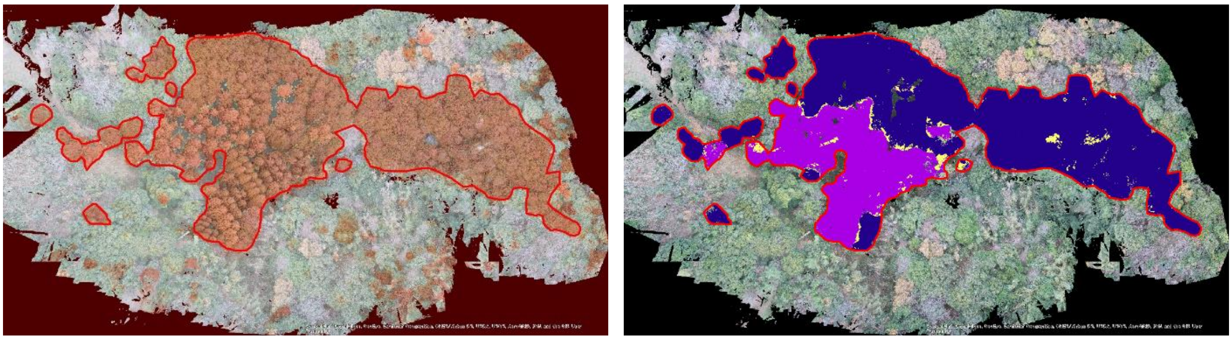
5. Discussion

5.1 Mapping of tree species

While logging, foresters consider about forest volume and price fluctuations due to demand for each species in the timber market. Therefore, the occupation ratio for each tree species is researched, the tree species in the logging area are mapped to estimate forest volume of each species.

In order to mapping the conifers species, UAV images are converted to ortho image by SfM processing. There are 177 UAV images taken from the forest area and predicted them to output conifer masks. RGB channels of UAV images are added to the conifer mask. Thus, we make 4 channel images. 4 channel images are processed using SfM in Pix4D software. SfM processing setting of Pix4D has RGB channels, which keeps the weight in each pixel to calc texture. And it keeps the pixel values in the multi channels Images. SfM processing exports the point cloud and ortho image. Fig.8 shows the conifer, cedar, and cypress channels in the ortho image.

In this study area, conifer area is 0.69 ha confirmed by visually distinguished in the 1.6 ha forest area. In Fig.8(A) shows the visually distinguished area with red line polygons, and the area labeled as conifer by U-net with 0.67 ha. We can predict 97% of the entire conifers in the study area. Fig.8(B) shows that conifer area labeled by U-net having 0.20 ha of cedar and 0.49 ha of cypress. The unified area labeled with cedar and cypress is 0.02ha.



(A) Red fill is conifer area using U-net **(B) Conifer species map using U-net**
conifer polygon by visually cypress cedar cedar and cypress
Figure 8: Conifer map (background: ortho image)

This study area is mixed forest that have some tree species of hardwood, bamboo and conifer. To train other species, we must prepare training dataset of them. In this study we focus on only conifer for U-net training, and prediction error has hardwood. If we prepare hardwood training dataset, there is a possibility that U-net can train the feature of conifer and hardwood. In other UAV images, the object is expected to change depending on the weather conditions, height of the sun, seasonality, flight altitude, and ground resolution due to different flight conditions. As the future study, we consider to examine the verification of IoU with these conditions. This verifying will be useful for making flight guidelines for tree species mapping.

5.2 Contributing of estimated forest volume

In this study, U-net makes it easier to classify conifer of forest tree species, which has been obtained by visual discrimination. This created visual discrimination method that requires empirical knowledge of tree species making boundaries to mixed forests.

For foresters, forest volume actual estimation based on tree height and diameter obtained from laser data or field sampling area results is applied.

Low cost volume estimation method based on data obtained from UAV-SfM for single tree species study (Takeuchi, 2019) is studied and the future study can be expected to be applied to mixed forests. Also, UAV has been widely used, and forestry operators start verifying pre-cutting surveys, bridge design for carrying out timber, process management, logging certification, (Forestry Agency, 2018). Spreading UAV utilization leads the development of low manpower and efficient methods as mentioned this study, which will spread and extend from various viewpoints.

In addition, forest managers in the region are recommend to update the forest GIS database by receiving the annual forest renewal status from the logging site with UAV. Collecting freshness of data can contribute to a sustainable forest management plan.

6. Conclusions

In conclusions, extraction of conifer in the forest is successfully done by U-net of deep learning technique with UAV imagery. Also, we can classify the area of cedar and cypress which is different volumes.

For this application, ground image resolution of UAV imagery is 3cm. The accuracy of extraction of conifer is approximately 96% with training data samples from original twelve images. Usually, it is very difficult to identify cedar and cypress by manual interpretation which require experimental knowledge. Therefore, this result is very useful for forest monitoring.

For future study, we are planning to apply this result to other forest area. We would like evaluate robustness of our algorithm. Also, we would to apply other species with the same U-net and image size.

7. Acknowledgments

The authors are grateful to Yamaguchi Prefectural Forestry Guidance Institute to support research field in this study area. We also thank to Aim Co., Ltd. for sharing their dataset with us.

8. Conflict of Interest

The authors declare no conflict of interest.

References

Reference from Journals:

Said, E, Fahmy, G, Nassar D, Ammar, H., 2015. Dental X-ray Image Segmentation. ISBI.

Falk, T, Mai, D, Bensch, R, Çiçek, Ö, Abdulkadir, A, Marrakchi, Y, Böhm, A, Deubner, J Jäckel, Z, Seiwald, K, Dovzhenko, A, Tietz, O, Bosco, D C, Walsh, S, Saltukoglu, D, Tay, L T , Prinz, M, Palme, K, Simons, M, Diester, I, Brox T & Ronneberger,O., 2019. Deep learning for cell counting, detection, and morphometry. Nature Methods.

Feng, Q, Liu J, Gong, J., 2015. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. Remote sensing ISSN 2072-4292.

Kim's, K M., 2016. Estimate forest stand volume with high accuracy using low-cost digital aerial photographs. FFPRI.

Fujiyoshi, H, Yamashita, T., 2017. Image Recognition by Deep Learning. ISSN.

Jonathan, P D, Olano, M, Ellis, C Eile., 2015. Optimal Altitude, Overlap, and Weather Conditions for Computer Vision UAV Estimates of Forest Structure.

Mizoguchi, T, Ishii, A, Nakamura, H, Inoue, T, Takamatsu, H., 2016. Individual tree species classification on laser scan of forest by deep learning. The Japan Society for Precision Engineering.

Morioka, Y, Yone, Y., 2019. Classification of UAV images using machine learning. The Japanese Forest Society.

Ronneberger, O, Fischer P, Brox T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Germany : University of Freiburg.

Takeuchi, F, Kondou, S, Takiya M, Ebina M, Zen K, Fuzisawa R, Ishihama N, Ohno, Y., 2019. Estimation of single tree volume of todo fir plantation using UAV-SfM and Deep Learning. The Japanese Forest Society.

Wagner H, F, Sanchez, A, Tarabalka, Y, Lotte G, R, Ferreira P, M, Aidar P, M, Gloor, E, Phillips & Aragao E O C, L., 2019. Using the U-net convolutional network to map forest types and disturbance in the Atlantic rainforest with very high resolution images. Remote Sensing in Ecology and Conservation.

Wannasiri, W, Nagai, M, Honda, K, Santitamnont, P, Miphokasap, P., 2013. Extraction of Mangrove Biophysical Parameters Using Airborne LiDAR. ISSN.

References from Other Literature:
Forestry and Forest Products Research Institute, 2019. I want to know the current state of forestry.

Forestry Agency, 2018. Social implementation of smart forestry.

Yamanashi Forest Environment Department, 2019. 6th Yamanashi Prefectural Forest Research Institute Basic Research Plan.

Reference from Websites:
Abadi, 2015. TensorFlow. Retrieved from https://www.tensorflow.org.

Labelbox, Inc. Labelbox. Retrieved from https://app.labelbox.com.

Learning Museum of the Forest and Forestry. Current state of forestry in Japan. Retrieved from https://www.shinrin-ringyou.com/ringyou/

LISA lab, 2018. U-Net. Retrieved from http://deeplearning.net/tutorial/unet.html.

Preferred Networks & Kikagaku, 2018 medical-ai-course-materials. Retrieved from https://japan-medical-ai.github.io/medical-ai-course-materials/notebooks/Image_Segmentation.html

Zhixuhao, 2019. unet for image segmentation. Retrieved from https://github.com/zhixuhao/unet