A Study on Automatic Tree Detection and Carbon Dioxide Absorption Estimation Using Drone Images and Deep Learning Method
（ドローン画像とディープラーニングを用いた樹木検出および CO2 の吸収量の推定に関する研究）

VON Chhaysoneath

概要：本研究では、スマートフォレストを目指すため、東京都市大学横浜キャンパスの保全林を対象に、低層リモートセンシング技術と深層学習アルゴリズム、U−Net セマンティックセグメンテーションを利用して、シラカシとバンブーの 2 樹種を自動検出し、CO2 吸収量の推定を試みた。本研究では、1178 枚の空中画像を収集し利用した。まず、U−Net でのトレーニングと検証を行うために、画像を 512x512 ピクセルのサイズに分割し、トレーニングおよび検証を行った。自動抽出の精度が 89 パーセントに達した。次に、多視点ステレオ写真測量と既存のデータを使用して、選択された領域で検出された樹木による二酸化炭素の吸収量を推定した。その結果、7 本樫の木で年間約 13.6 キロの CO2 を吸収できることがわかった。

Summary: This study aims to introduce a smart forest monitoring technique by utilizing low-rise remote sensing technology and deep learning algorithms, U-Net semantic segmentation, to detect two tree species Chinese Green Oak (シラカシ) and Bamboo (竹) at Tokyo City University Yokohama Campus. DJI Mavic Mini2 is used to collect 1178 aerial imageries of size 4000 x 2250 pixels. 31 images are split into sizes of 512 x 512 pixels for training and verification in U-Net. The method exhibits a relatively remarkable result of 89% accuracy. Next, the study estimates the fixed amount of carbon dioxide absorption of detected trees in the selected area using multi-viewpoint stereo photogrammetry and existing data. As a result, seven samples matured oak trees could absorb approximately 13.6 kg per year. To conclude, this research has proven valuable and convenient for the innovation of smart forest monitoring systems despite finite time and small resources, the result of the study is noteworthy, practical, and achievable.

キーワード: 樹木検出、ドローン画像、ディープラーニング、画像セグメンテーション、U−Net

Keywords: Tree detection, Drone Image, Deep Learning, Image Segmentation, U-Net

1. Introduction

In a simple term, "Environment" refers to the surroundings, including the entire physical and biotic conditions affecting the reaction in the organism. The leading culprits causing constant changes in the global environmental problems are human civilization, especially globalization. Elements that are considered to provoke environmental issues are pollution, ozone depletion, acid rain, deforestation, biodiversity loss, and climate change.

It is undeniable that climate change harms every country on the continent. According to United Nations, 2019 was recorded as the second warmest year and the highest concentration of carbon dioxide levels and other greenhouse gases in the atmosphere. In Sustainable Development Goals (SDGs) 13, all countries have the responsibility to take action in combating climate change and its impact.

1.1 Research Background

Up to now, various deep neural network architectures have been utilized to deal with image segmentation problems (Zaitoun & Aqel, 2015). Recent development in Artificial Intelligence has guided the utilization of core algorithms so-called Deep Learning (DL) which have been applied in diverse area to attain remarkable result in various studies.

1.2 Research purpose

In this research, there are two primary purposes to be achieved. First and foremost is utilizing low-rise remote sensing technology, drone, deep learning algorithms, and the U-Net semantic segmentation model to detect target tree species. Secondly, we estimate a fixed amount of carbon dioxide that a tree can absorb from a sample target tree species at Tokyo City University, Yokohama Campus to automatically detect two tree species, Quercus Myrsinifolia known as Chinese Green Oak (シラカシ), and Bambusoideae known as Bamboo (竹) .

1.3 Research Flow

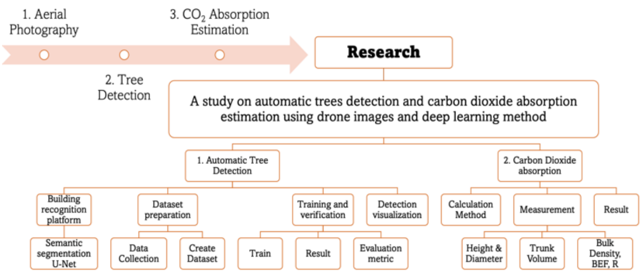

Figure 1. Research flow

In order to carry out this research, there are three essential steps. The first one is taking aerial photography, after that tree detection is performed by using deep learning and finally the co2 absorption estimation is performed. Figure 6 shows the whole procedure for this research with detailed procedure.

2. Methodologies

2.1 Automatic Tree Detection: U-Net

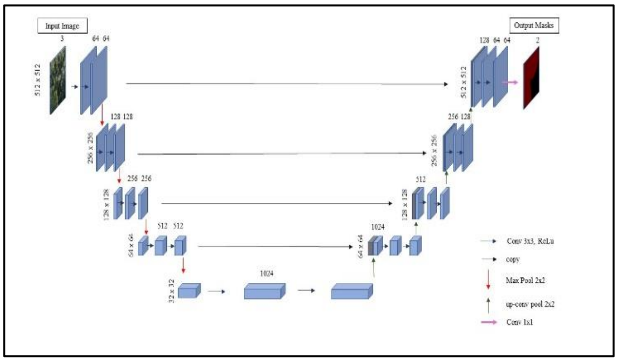

Figure 2. U-Net Architecture Used in the study

In this study, U-Net, a convolutional network for multi-class image segmentation was used. U-Net model architecture encompasses of encoding path and a decoding path. The encoding path extracts hierarchical features/patterns given labeled input data, while the decoding path learns spatial information required to reconstruct the original input (Ronneberger, Fischer, & Brox, 2015). The network performs per-pixel classification and anticipates the probability of each pixel that belongs to a specific class. This study adopted the U-Net architecture accordingly from (Ronneberger, Fischer, & Brox, 2015) by employing the same filter and three-band RGB images as the input.

2.2 Carbon Dioxide Absorption Estimation

There are many possible ways to calculate the amount of carbon dioxide trees can absorb depending on the data type one has. However, the estimation method used in the study is followed by the **Forest Research and Management Organization of Japan** to calculate the amount of carbon dioxide fixed by trees.

$$C = V \times D \times BEF \times (1+R) \times CF \times 44/12$$

C = Carbon fixation amount (kg)
V = trunk volume (m^3)
D = Bulk density (kg/m^3)
BEF = Magnification factor
R = Underground/aboveground ratio
CF = Carbon content

2.2.1 Measurement

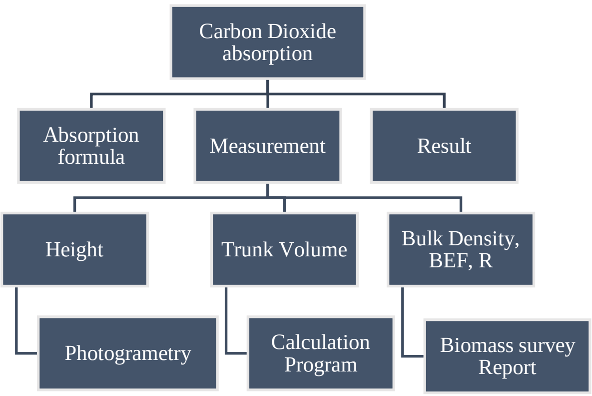

Table 1. Carbon dioxide absorption flow

Firstly, all the 1178 images are being processed by photogrammetry in Agisoft's Metashape photoscan to generate following results: tie points, dense cloud, mesh (3D model), Digital Surface Model (DSM), Digital Elevation Model (DEM), and orthomosaic. Therefore, the tree height of 10 sample oak trees have been calculated. With the height and diameter at breast height, we calculate the volume of the tree by using the "Stem Volume Calculation Program" of Forest Research and Management Organization of Japan (2).

Meanwhile, the rest of the calculation parameter such as biomass expansion factor (BEF) [above-ground biomass/below-ground biomass], root-to-shoot ratios (R) and Wood density (D) were extracted from the "Classification in Forest Status Survey" (3) with parameter as following D=0.469, BEF=1.37, R=0.26, CF=0.48. Ultimately, the carbon dioxide that trees can absorb has been estimated individually and in sum.

2.3 Data Collection

Figure 3. DJI Mavic Mini 2 used in the study

This study uses DJI Mavic Mini 2 for data collection (Fig.2) The DJI Mini 2 is very light and weighs less than 199 g. Its small figure and weight make it very convenient to use.

To be able to compute in U-Net specially to process in photogrammetry, a large amount of data is needed to learn, test, and verify. Image used in this study is drone imageries acquired by DJI Mavic Mini 2 flying 30 meters above ground at protection forest of Tokyo City University (TCU) Yokohama Campus. The FC7303 camera with a 4-millimeters focal length built-in points straight down to collect data at this altitude. In total, 1178 drone imageries were collected (*fig.4*).

Figure 4. Sample Image Collection

2.4 Dataset Preparation for Tree Detection

2.4.1 Image Selection

Among the 1178 images that have been collected, only 31 images were used to perform automatic tree detection in U-Net. The 31 images are high-quality images and contain enough sample target tree species.

2.4.2 Image Split

31 images are selected from total collected imageries of size 4420 by 2250. To run in the deep learning, the images are split into smaller pixels generate a total of 992 datasets of size 512 by 512 pixels (*fig.5*).

Figure 5.Sample data split

2.4.3. Data annotation: Labelme

In this study, the annotation tool, LabelMe, is used to draw polygons and assign classes manually (*fig.6*).

Figure 6. Sample image with annotation

2.4.4 Data Augmentation

Data augmentation is a method to artificially expand the amount of data by producing new data from existing data. This study employed data augmentation to substantially increase the training sample size to alleviate the limitation and overfitting problems such as rotating, cropping, and filtering as shown in *fig.7*.

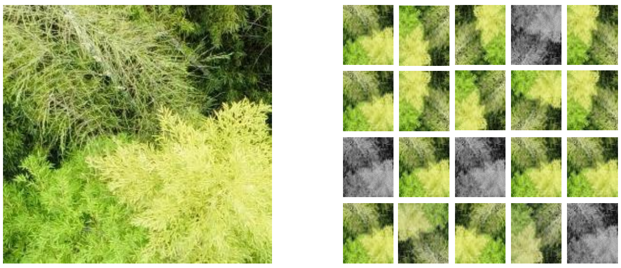

Figure 7. Data Augmentation

3. Data Experiment and Result

3.1 Automatic Tree Detection

3.1.1 Training

The architecture used a sigmoid activation function to ascertain that output pixel values range between 0 and 1. This training used an input size of 512 x 512 pixels for the segmentation of 520 training samples. 90% of images were used for training and 10% were used for validation. During the training, the Adam optimization was used with parameters as follows: learning rate 1e-4 (0.0001). The training was trained for 200 epochs where each epoch comprises 128 batches size.

Figure 8 demonstrates the characteristic of the loss in the training process. It shows that the loss dropped dramatically at the initial stage and gradually stabilized until 50 epochs. The fluctuation similar to a bump occurs due to the frozen training batch. It is noticed that the lowest loss was measured after around 90 epochs.

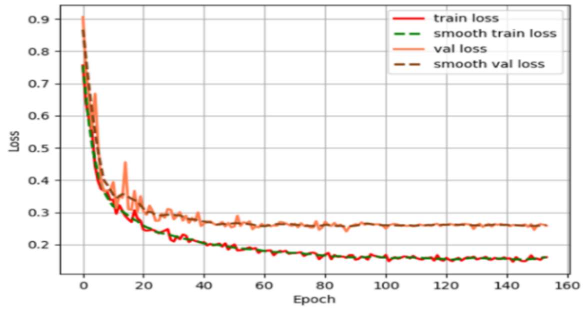

Figure 8. Loss trend chart

3.2.2 Validation

To ensure the preciseness of the model, various alternatives for estimating the accuracy of semantic segmentation are used. In this study, to quantify the result, mean pixel accuracy (mPA), mean intersection over union (MIoU), the weight Precision, and mean Recall (mRecall) were used. As a result, the model performed well and was able to detect the target tree species precisely with a decent accuracy rate as shown in table2.

Quantitative evaluation of model performance					
U-Net	mPA	mIou	mRecall	mPrecision	Overall Accuray
Quercus myrsinifolia (シラカシ)	96%	84%	96%	87%	
Bambusoideae (bamboo)	86%	76%	86%	87%	
Background	85%	78%	85%	90%	
Overall	89%	79%	89%	88%	89%

Table 2. Quantitative evaluation of model performance

3.2.3 Detection

We carried out the detection result to validate the accuracy by performing the visualization prediction on images containing the target species. The image on the left side refers to the sample training before the detection prediction whereas the image on the right side refers to the sample training data after the prediction.

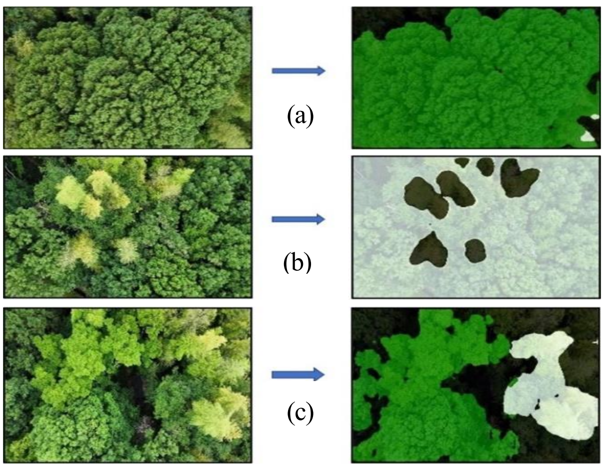

Figure 9. Tree species detection of (a) Green Oak, (b) Bamboo, (c) Both species

3.2 Carbon Dioxide Absorption Estimation

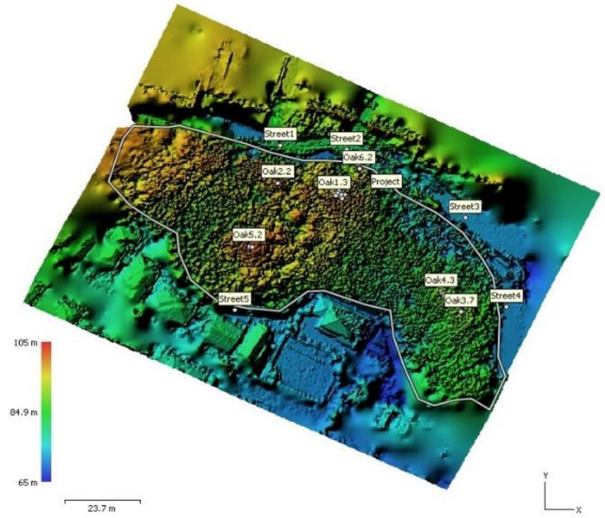

Figure 10. Tree sample selection on DSM model

Within the selected area, 5613 square meters, DSM is generated to estimate tree height by subtracting the altitude level at the trees' end with the ground's altitude level. Since the selected study area is not flat, the ground's level is marked in accordance with the tree (point) as shown in *fig.10* and the result is generated in table 3.

No	Trees	DEM Altitude	Tree Height (m)	DBH (cm)	Volume (m^3)	D	BEF	1+R	CF	CO_2(kg)
1	Oak1	97	15	43	0.87	0.469	1.37	1.26	0.48	1.240
2	Oak2	99	16	45	0.87	0.469	1.37	1.26	0.48	1.240
3	Oak3	90	18	51	1.47	0.469	1.37	1.26	0.48	2.095
4	Oak4	90	18	51	1.47	0.469	1.37	1.26	0.48	2.095
5	Oak5	101	26	73	4.30	0.469	1.37	1.26	0.48	6.127
6	Oak6	94	12	34	0.44	0.469	1.37	1.26	0.48	0.627
7	Oak7	94	21	59	2.30	0.469	1.37	1.26	0.48	3.277
8	Oak8	93	20	57	2.13	0.469	1.37	1.26	0.48	3.035
9	Oak9	99	16	45	1.00	0.469	1.37	1.26	0.48	1.425
10	Oak10	98	13	37	0.50	0.469	1.37	1.26	0.48	0.712
Total										21.873

Table 3. Carbon Dioxide Absorption Estimation on Sample Trees

4. Discussion

Throughout this research analysis, U-Net delivered a decent and acceptable result with an accuracy of 89%. The result of this study can be proven that U-Net architecture is an efficient and lightweight model for tree detection. This study uses only two species, but the model can be used on multiple classes, making it an efficient and resourceful technique for smart forest monitoring. However, although the model executes favorable accuracy results of 89%, it is speculated that the performance is limited and cannot classified from non-training data source which is quite challenging experience. The metadata needs to be from the surrounding location or apply the same data collection parameter. Otherwise, the detection rate is low or unrecognizable. The remaining 11% errors can be enhanced by getting more training data with precise annotation.

In the second objective, the height obtained by DSM is relatively accurate. To quantify tree height results, another method is used which is by laser find ranger. The difference was 1 meter for a tree with 12m in height. No module or calculation formula estimate precise carbon dioxide absorption due it its complication of data collection. In this study, the tree's circumference is measured by the tape meter of one tree to estimate the rest of the tree according to height. Circumference cannot be measured by photogrammetry is due to the crowdedness of the forest crown covering up what is below the tree. This can be improved by choosing a sparser tree location or selecting the conifers instead of broad leaf trees. In addition to the above method, more ground truth data is needed to quantify the result of absorption estimation.

5. Conclusion

This study's first objective is to use a small number of aerial photographs to run in deep learning: U-Net to automatically detect two tree species: Oak and Bamboo, as sample tree types. Secondly, to unmanned aerial vehicles and the stereo photogrammetry method to generate a Digital Surface Model (DSM) to calculate the height of the selected sample trees. To conclude, this research is carried out to examine the remote sensing advancement by integrating aerial photographs and deep learning, and the contribution to achieving carbon neutrality by 2050. With finite time and small resources, the result of the study is noteworthy, practical, and achievable. However, this research also has its limitations as described, which can be improved by collecting more ground truth data. Overall, this research has proven valuable and convenient for the innovation of smart forest monitoring systems.

References

(1) Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer Assited-Intervention* (pp. pp. 1-8). Springer, Cham. Retrieved from https://arxiv.org/abs/1505.04597

(2) Kazuo Hosoda, Yasushi Mitsuda, Toshiro Iehara "Differences between the current standing tree trunk volume table and the calculated values by the volume formula and how to correct them" Journal of Forest Planning Society Vol. 44, No. 2: 23-39, published in December 2010

(3) Greenhouse Gas Inventory Office of Japan and Ministry of the Environment, Japan (eds.), National Greenhouse Gas Inventory Report of JAPAN 2022, Center for Global Environmental Research, Earth System Division, National Institute for Environmental Studies, Japan, 2022.