Evaluation of Automatic Detection of Uncultivated Land

Using Machine Learning for Japanese Municipality

Jun S. [1*]

[1]First Author's Affiliation: Associate Professor, Faculty of Information & Communications,

Bunkyo University, Japan

*sakuraij@bunkyo.ac.jp

Abstract *In Japan, the increase in abandoned farmland is becoming increasingly serious due to the aging of farmers and the lack of successors, and countermeasures to this problem have become an important policy issue. For this reason, each of the local governments conducts an annual survey of the use of agricultural land, covering all agricultural land in Japan. In this survey, farmland is judged in stages according to the degree of disrepair, and in particular, it is important to determine whether farmland is in disrepair or not. Currently, workers actually visit all farmland to conduct field surveys, but this is considered a problem because of the large human cost involved. Therefore, this study proposes an automatic method for determining cultivated and uncultivated land by machine learning using aerial photographs. The procedure is as follows: First, a large amount of polygon data extracting the extent of farmland from aerial photographs is generated using GIS. Next, for each farmland, it is classified into cultivated and uncultivated land based on the results of a utilization survey. Then, by using a convolutional neural network, a learning model is constructed to classify cultivated land and uncultivated land. Furthermore, in order to examine the feasibility of using this method and the issues involved, the accuracy of the method is verified for agricultural lands in a specific municipality in Japan. As a result, a certain degree of accuracy was confirmed, indicating the possibility of using machine learning. On the other hand, regarding farmland that could not be detected correctly, the necessity of using photogrammetry technology to estimate the height of trees and shrubs was suggested.*

Keywords: uncultivated land, image classification, machine learning, GIS

Introduction

In recent years, the increase in abandoned cultivated land has become a serious issue due to the aging of agricultural workers and the shortage of successors, making the development of countermeasures a critical policy issue. Against this backdrop, the "Act for Promotion of Farmland Bank Activities [1]" was enacted in 2014, leading to the establishment of Farmland Bank in all prefectures. These Farmland Banks are conducting a business of leasing farmland from owners that cannot be cultivated due to reasons such as aging, and then subleasing it to motivated farmers. To promote this system, it is essential to accurately grasp the current status of farmland. Therefore, Agricultural Committees established in each municipality conduct an annual Farmland Utilization Survey. Since the fiscal year 2021, this survey has been integrated with the dilapidated farmland survey conducted by municipal departments, resulting in the definition of new

classification methods based on the degree of farmland dilapidation. Specifically, all farmland nationwide is classified into "Cultivated Land," "Category 1 Idle Farmland," which refers to land unlikely to be cultivated in the future, "Category 2 Idle Farmland," which is significantly underutilized compared to surrounding farmland, and "Farmland Difficult to Restore," which is heavily degraded and challenging to restore to farmland. Furthermore, "Category 1 Idle Farmland" is further divided into "Category 1 Idle Farmland (Green)" and "Category 1 Idle Farmland (Yellow)" based on criteria such as the proliferation of shrubs exceeding human height, resulting in a total of five categories. This classification process is carried out through on-site visual inspections by multiple agricultural committee members but is seen as problematic for requiring substantial human resources and for potential errors and omissions in identifying the locations of the land. Therefore, the establishment of a framework that utilizes information technology to support the Farmland Utilization Survey is highly desired.

In terms of existing related research, since the Farmland Utilization Survey is unique to Japan, various examples exist within the country. For example, Fukumoto et al. (2016) [2] devised a method for inspecting dilapidated farmland using Google Street View on-site images; however, it has been pointed out that because this method requires visual confirmation, it is not an automated survey method, and that the limited availability of images from the latest fiscal year has been identified as a challenge. Additionally, there have been studies that attempted to extract potentially dilapidated rice paddies using satellite data (Fukumoto et al., 2014) [3] and to identify potentially dilapidated farmland using aerial laser data (Tsukahara et al., 2020) [4]. These studies aimed to detect farmland with a potential for dilapidation through image processing. However, satellite and aerial laser data are extremely expensive and have limited measurement periods, making it virtually impossible for local governments to obtain data that corresponds to their survey periods. Furthermore, as a model project in Kanagawa Prefecture (Simmon Co., Ltd., 2020) [5], a survey of dilapidated farmland using drone aerial imagery has been proposed. However, this was merely an attempt to determine whether dilapidated farmland could be visually identified from aerial images, and it has not yet led to the establishment of an automated method. To address these challenges, the author has devised a prototype method for automatically determining cultivation status using aerial photographs [6]. However, due to the use of low-resolution free aerial photographs, the determination accuracy was only about 80%, resulting in the challenge of some farmland not being correctly determined.

Therefore, this study proposes and verifies the accuracy of an automatic method for determining cultivated land and uncultivated land using machine learning with high-resolution aerial photographs, and examines the potential applications and challenges for supporting the Farmland Utilization Survey.

Research Overview

This study proposes a method for automatically determining cultivated land and uncultivated land using machine learning with high-resolution aerial photographs. As illustrated in Figure 1, the processing flow of this study is divided into three main parts: the data generation part, the learning part, and the estimation part. The details of these processes are described below.
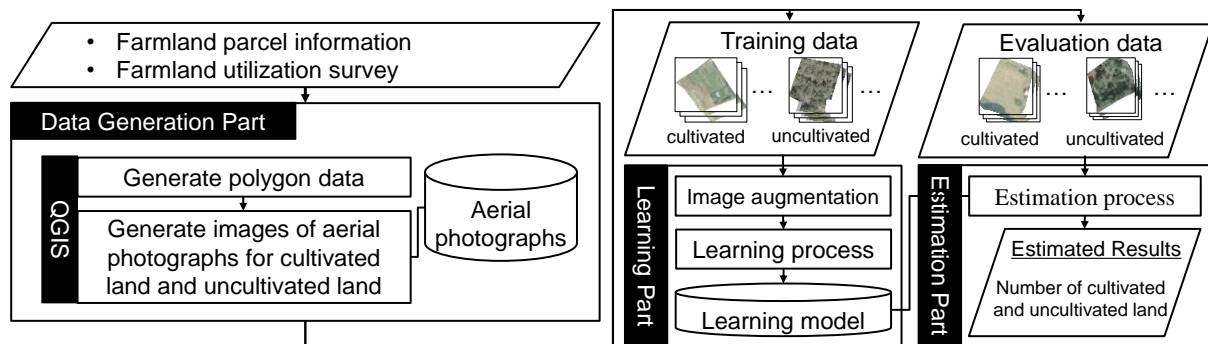

Figure 1: System Flow

a.   Operational Environment and Data Used

As the operational environment for this study, QGIS 3.16, a GIS software, was used in the data generation part to generate and process data by leveraging various GIS functionalities. For the machine learning part and estimation part, Google Colaboratory Pro, which allows for writing and executing Python code within a browser, was utilized, with machine learning libraries TensorFlow and Keras employed for processing. The data used in this study is shown in Figure 2.
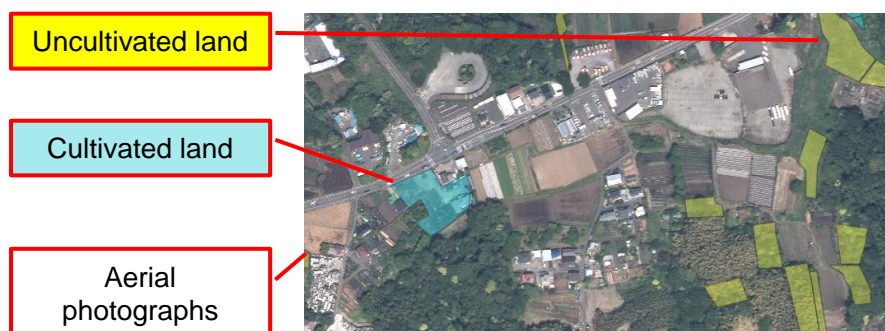

Figure 2: Data Used

First, the correct data used to determine whether the target farmland is categorized as cultivated land or uncultivated land was based on the results of the Farmland Utilization Survey conducted by the Agricultural Committee in 2021. In this survey, all farmland, except for cultivated land, that was judged as dilapidated was defined as uncultivated land Next, as data for delineating farmland parcels, farmland parcel information referred to as "Fude Polygon [7]" from the Ministry of Agriculture, Forestry and Fisheries was utilized. Furthermore, aerial photographs taken in 2021 with a resolution of 25 cm, known as "GEOSPACE Aerial Photo 2500", were purchased and employed as the aerial photograph data.

b.   Data Generation Part

As shown in Figure 3, the data generation part uses QGIS to generate polygon data for cultivated land and uncultivated land.
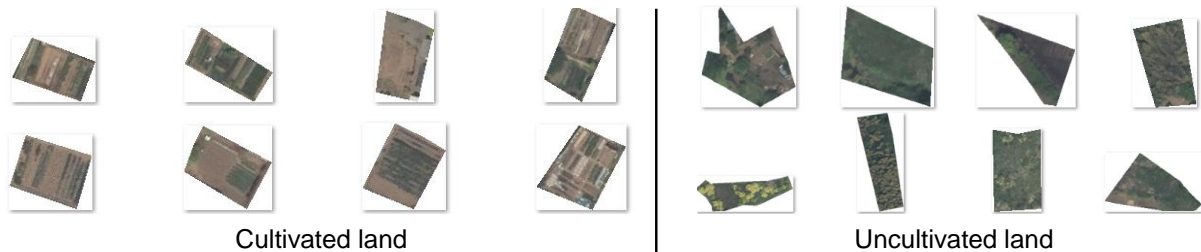

Figure 3: Generated data

Specifically, the layer splitting function based on attributes in QGIS is used to generate polygon data for each parcel classified as cultivated land or uncultivated land. Then, the clipping function with mask layers in QGIS is used to generate images of aerial photographs by clipping within the boundaries of each polygon for both cultivated land and uncultivated land.

c.   Learning Part and Estimation Part

Based on the images of cultivated land and uncultivated land generated in the data generation part, the data were classified into labeled training data for learning and evaluation data for estimation. In the learning part, image augmentation was first applied, expanding the original images by tilting them in 5-degree increments from -20 degrees to 20 degrees. Next, through the learning process, using convolutional neural networks (hereinafter referred to as CNN), a method of deep learning, the labeled training data were trained. Given the small size of the original input images, the input image size was set to 128×128 pixels with three channels, and the model structure consisted of four convolutional layers and two fully connected layers. In the estimation part, the learning

model generated in the learning part was used to estimate the results of the evaluation data, determining whether the farmland is cultivated land or uncultivated land.

Evaluation Experiment

In this experiment, the proposed method was used to evaluate the estimation accuracy of determining cultivated land and uncultivated land using two types of aerial photographs, thereby verifying the feasibility of automatic determination.

As the evaluation procedure, first, 270 polygons of cultivated land and 270 polygons of uncultivated land were extracted from fields in Chigasaki City, Kanagawa Prefecture. Then, 54 polygons of cultivated land and 54 polygons of uncultivated land were used as evaluation data, while the remaining 216 polygons of each were used as labeled training data, and the learning and estimation processes were carried out. The evaluation data underwent 5-fold cross-validation, resulting in the estimation of 270 images each for cultivated land and uncultivated land. The results were evaluated using precision, recall, and F-measure. The evaluation results are shown in Table 1.

Table 1: Evaluation Results

Category	Total	Estimated number	Correct number	Precision	Recall	F-measure
Cultivated land	270	268	236	0.88	0.87	0.88
Uncultivated land	270	272	238	0.88	0.88	0.88
All	540	540	474	0.88	0.88	0.88

From these results, it was found that the F-measure was 0.88 for both cultivated land and uncultivated land, indicating that the method correctly determined about 90% of the images as either cultivated land or uncultivated land. While a certain level of accuracy was ensured, approximately 10% of the determinations were incorrect. To investigate the causes, selected images from each result are presented in Table 2.

Table 2: Examples of evaluation results

Survey Results	Judgment Results		Picture
Cultivated	Cultivated	Correct	Cultivated
Cultivated	Uncultivated	Incorrect	Trees
Uncultivated	Uncultivated	Correct	Forest or Concrete
Uncultivated	Cultivated	Incorrect	Land

This table presents selected images from the results of the Farmland Utilization Survey and the determination results from the proposed method. First, focusing on the images where both the survey results and the determination results correctly identified cultivated land, the fields were visibly tilled, with clear boundaries between rows. In contrast, when the survey results indicated cultivated land but the determination results mistakenly identified uncultivated land, many images showed parts of the fields covered with trees. Next, focusing on the images where both the survey results and the determination results correctly identified uncultivated land, it was observed that almost the entire area was covered with trees or had been paved with concrete, indicating that the land was no longer

farmland. On the other hand, when the survey results indicated uncultivated land but the classification results mistakenly identified cultivated land, it was found that although the land was not tilled as farmland, it was covered with soil without any vegetation.

Based on the above findings, it was concluded that while there are instances where the presence of trees or bare soil leads to incorrect determination, it was found that the method generally succeeds in correctly determining clear cases of cultivated land and uncultivated land.

Conclusion

This study proposed a method for determining between cultivated land and uncultivated land using high-resolution aerial photographs and verified the determination accuracy of this method. The following insights were obtained from the results:

The use of high-resolution aerial photographs resulted in a high F-measure of 0.88 for both cultivated land and uncultivated land, indicating a high level of determination accuracy.

Cases were observed where the method failed to accurately determine cultivated land with trees present in parts of the field and uncultivated land consisting solely of soil.

From these findings, it can be concluded that the feasibility of utilizing machine learning for the automatic determination between cultivated land and uncultivated land has been demonstrated to a certain extent. However, for farmland that could not be accurately determined, efforts should be directed toward exploring ways to improve accuracy by employing drones to capture higher-resolution photographs and utilizing photogrammetry techniques to estimate the height of shrubs and other vegetation.

Acknowledgements

References

[1] Ministry of Agriculture, Forestry and Fisheries of Japan. Act on Promotion of Interim Agricultural Land Management Projects. Retrieved September 10, 2024, from https://www.maff.go.jp/j/keiei/koukai/kikou/kikou_law.html

[2] M. Fukumoto. and S. Shindo., (2016). Method to Check Mistakes and Omissions of Survey Data for Locations of Abandoned Farmlands Utilizing Google Earth. *Technical report of the National Institute for Rural Engineering*, No. 218, pp. 19-28

[3] M. Fukumoto. and H.Yoshisako., (2014). Screening Method of Paddy Fields Targeted for Field Surveys during Investigations on the Status of Generation and Elimination of Devastated Farmlands Using Multitemporal Satellite Data and Paddy Parcel Boundary Data. *Transactions of the Japanese Society of Irrigation, Drainage and Rural Engineering*, Vol. 82，No. 5，pp. 339-346

[4] M. Tsukahara. and B. Ota.. Using aerial laser data to identify degraded farmland. Retrieved September 10, 2024, from https://www.ajiko.co.jp/dl_data/pdf_ff2020/p64-65.pdf

[5] Simmon Co., Ltd.. Survey of devastated farmland using drones. Retrieved September 10, 2024, from https://www.pref.kanagawa.jp/documents/67972/02_nouti.pdf

[6] J. Sakurai., (2023). Investigation of Support for Survey of Farmland Use by Machine Learning Using Aerial Photograph. *Bunkyo University Information and Communication Studies*, Vol. 67，pp. 15-24

[7] Ministry of Agriculture, Forestry and Fisheries of Japan. farmland parcel information. Retrieved September 10, 2024, from https://www.maff.go.jp/j/tokei/porigon/index.html