# MAPPING AQUACULTURE RAFTS FROM A GEOEYE-1 IMAGE USING U-NET

Tatsuyuki Sagawa[1] and Hiroki Murata[2]
[1]Tottori University of Environmental Studies, 1-1-1 Wakabadai-kita, Tottori, Tottori 689-1111, Japan,
Email: sagawa-t@kankyo-u.ac.jp
[2] Graduate School of Agriculture Science, Tohoku University, 468-1 Aramaki Aza Aoba, Aoba-ku, Sendai 980-0845, Japan,
Email: murata.ykhm@gmail.com

**KEY WORDS:** aquaculture raft, GeoEye-1, U-Net

**ABSTRACT:** Monitoring aquaculture facilities is important for managing marine resources and protecting the environment. For example, large amounts of excrement from aquaculture rafts have negative effects on the environment; an appropriate number of rafts is thus required. There are many aquaculture rafts for oysters and scallops in Yamada Bay on the Sanriku Coast of Japan. In this study, we aimed to map these aquaculture rafts using a high-resolution GeoEye-1 satellite image and deep-learning method. U-Net is a fully convolutional network (FCN) often used for segmentation and object-detection tasks. Raft areas were extracted from the GeoEye-1 image using a U-Net model. Polygon data for the rafts were created manually based on visual inspection and used as truth data. These raft polygon data were used to train and validate the U-Net model. Analysis of the GeoEye-1 image using U-Net also provided raft polygons. Through comparison of these raft polygons, precision and recall values, and the F1 score, were calculated as accuracy indices. The proposed method mapped rafts in Yamada Bay with high accuracy (maximum F1 score = 0.990) and is expected to promote efficient monitoring of aquaculture facilities.

## 1. INTRODUCTION

Many aquaculture facilities, such as rafts and fish tanks, are installed along the Japanese coast. In 2019, the production value of Japanese aquaculture was 501 billion yen, accounting for 34% of the value of all fishery products (Fishery Agency, 2021).

Marine aquaculture affects the surrounding environment, and monitoring of aquaculture facilities is important for stable management of marine resources and environmental protection. Water temperature and dissolved oxygen affect the growth of aquaculture targets, and red tides and waves can cause considerable damage to aquaculture facilities. In turn, aquaculture has a negative impact on the environment because organic matter, such as the excrement of target fish and shellfish that accumulates on the seabed, consumes oxygen and generates toxic sulfides. In Japan, deterioration of the fishing-ground environment is prevented by controlling the amount of aquaculture, depending on environmental carrying capacity, and disease control measures. However, marine aquaculture facilities are not always managed properly. At some installations the number of facilities greatly exceeds the limit stipulated by the fisheries cooperative (Mutsuda et al., 2011), while other facilities are installed outside of demarcated fishing areas.

Satellite remote sensing is an efficient method for monitoring aquaculture facilities and the marine environment. Murata et al. (2021) mapped aquaculture rafts and seagrass beds using high-spatial-resolution satellite images to assess the impact of rafts on seagrass beds. However, the distribution and number of rafts were determined by visual inspection, which takes considerable time and effort because there are almost 100 rafts in each satellite image.

Various satellite image analysis methods have been used for extracting aquaculture facilities, such as rafts, including object-based (Wang et al., 2018) and deep-learning (Shi et al., 2018) methods. Wang et al. (2018) extracted aquaculture rafts from Gaofen-2 multiband images (spatial resolution = 3.2 m) and had an F1 score of 0.84. Shi et al. (2018) proposed a raft-extraction method based on a homogeneous convolutional neural network (HCN); an accuracy of 0.723 was achieved using the Gaofen-1 wide-field-of-view camera (spatial resolution = 16 m), as measured using Intersection over Union. Both methods were highly accurate, but direct comparison of their results is difficult due to the use of different accuracy indices. In these studies, the size of the target rafts and details of undetected rafts and misdetected objects were not clearly described. The accuracy of these methods may depend on the type of raft and bay. Further studies using various rafts and satellite images are required to confirm the accuracy and robustness of the methods.

In this study, we aimed to accurately map aquaculture rafts in Yamada Bay, on the Sanriku Coast of Japan, using a satellite image. A high-spatial-resolution GeoEye-1 satellite image was obtained and the U-Net deep learning method was applied for image analysis. The accuracy of raft mapping was evaluated by comparing the results with raft

information obtained via visual interpretation of the satellite image.

## 2. STUDY AREA AND DATA

The Sanriku Coast of Japan is a thriving area for marine aquaculture. There are many aquaculture rafts for oysters and scallops in Yamada Bay. In this study, we aimed to map these rafts using a satellite image. A typical aquaculture raft for oysters or scallops is 4 m wide and 12 m long. In addition to these rafts, there are also 50- to 100-m-long longline aquaculture facilities with ropes and buoys in Yamada Bay (Figure 1).

Figure 1. Aquaculture rafts (front) and longline aquaculture facilities (back) in Yamada Bay (photo was taken on January 18, 2019).

The GeoEye-1 satellite was launched on September 6, 2008 and has a high-resolution optical sensor. The spatial resolutions of the multispectral and panchromatic bands are 2.0 and 0.5 m, respectively. An image covering Yamada Bay obtained by GeoEye-1 on May 20, 2018 was used to map the rafts (Figure 2).

Visual inspection of the satellite image revealed 823 rafts in Yamada Bay. The raft polygon data were considered as truth data, and were used for training the deep-learning model and evaluating the detection results.

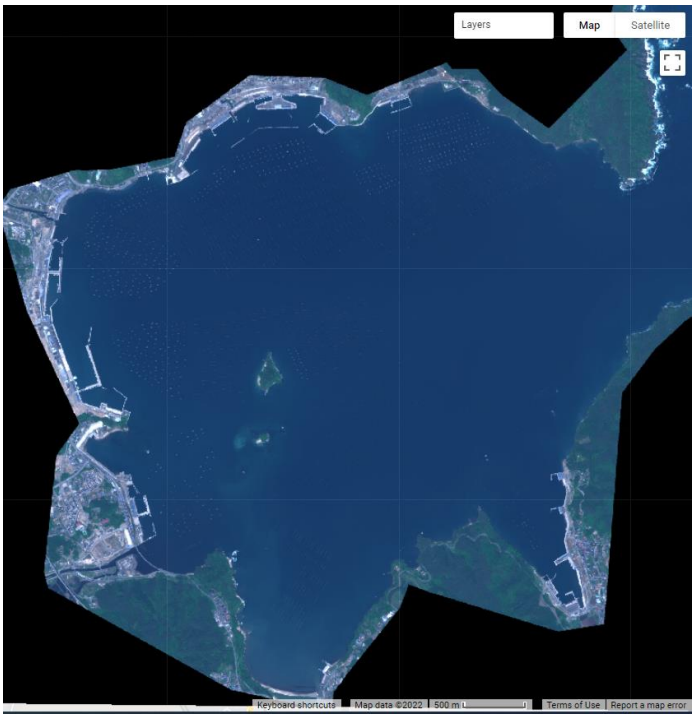



Figure 2. GeoEye-1 image of Yamada Bay on May 20, 2018.

## 3. METHOD

### 3.1 U-Net model for raft detection

Deep learning methods have progressed rapidly and are being applied in various fields. In remote sensing, a large amount of satellite derived data are widely available, and deep learning methods are also suitable for analyzing these complicated big data. U-Net is a fully convolutional network (FCN) often used for segmentation or object detection. U-Net has been applied to analyze satellite data for land cover classification (Gian et al., 2020) and building detection (Alsabhan et al., 2022).

In this study, U-Net was applied to extract raft areas from a satellite image. Figure 3 shows the U-Net framework developed for raft extraction. A 64 × 64 pixel satellite image was used as input. There were five channels corresponding to the panchromatic and multispectral bands. The resolution of the multispectral images was adjusted to that of the panchromatic image (0.5 m) by resampling using the nearest neighbor method. A 64 × 64 pixel area corresponds to 32 × 32 m on the ground, and can include an aquaculture raft and the surrounding environment (Figure 4). The output is also a 64 × 64 pixel image; the probability of an aquaculture raft being present is expressed as a value between 0 and 1.

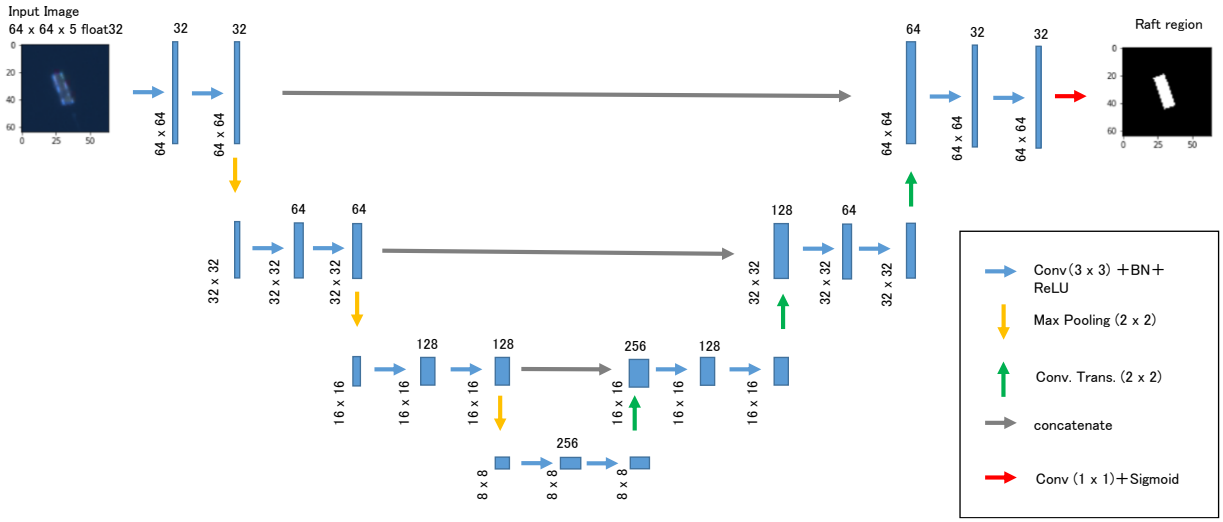



Figure 3 U-Net framework developed for raft extraction.

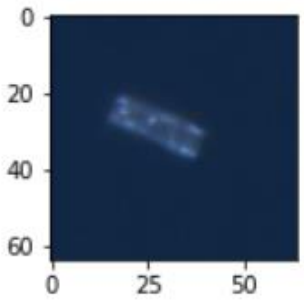


Figure 4 A 64 pixels x 64 pixel image including an aquaculture raft.

### 3.2 Analysis environment

Google Earth Engine (GEE) and Google Colaboratory (GC) were used for satellite image analysis and deep learning. GEE is a suitable environment for comparing satellite images and raft polygon data, and was used to create a dataset for deep learning, as well as to show the results and evaluate the accuracy of raft extraction. TensorFlow and Keras are open-source libraries for machine learning and artificial neural networks that can be used with GC. A U-Net model was developed using GC and a dataset created with GEE. Satellite image analysis was conducted using the U-Net model and GC, while post-processing (such as accuracy assessment) was conducted using GEE.

### 3.3 Analysis Flow

Figure 5 shows the analysis flow. Using GEE, multispectral band satellite images were resampled by the nearest neighbor method to 0.5 m resolution, to match the resolution of the panchromatic band and create a five-band image with 0.5 m resolution. A labeled raft image (0.5 m resolution) was also created through visual inspection of the raft polygons; values of 1 and 0 denote raft areas and non-raft areas, respectively.

First, 1,000 sampling points were generated randomly within each raft polygon. An additional 1,000 sampling points were generated randomly in areas outside of the raft polygons, including land, breakwaters, ships, and other aquaculture facilities.

For each sampling point in the five-band satellite image, a 64 × 64 pixel image was extracted. The image dataset was randomly separated into training and test datasets, which were exported to Google Drive for use with GC. A satellite image of the analysis area was also exported to Google Drive, too.

Using GC, the U-Net model was developed and trained with the training and test datasets. The satellite image was analyzed and the raft probability image was exported.

The pixels in the raft probability image have values of 0–1. A raft area was extracted from the raft probability image based on a threshold. A majority filter measuring 3 × 3 pixels was applied to the extracted raft image to reduce noise. Raft polygons were created from the raft image; those less than half the normal raft size (48 m2) were removed.

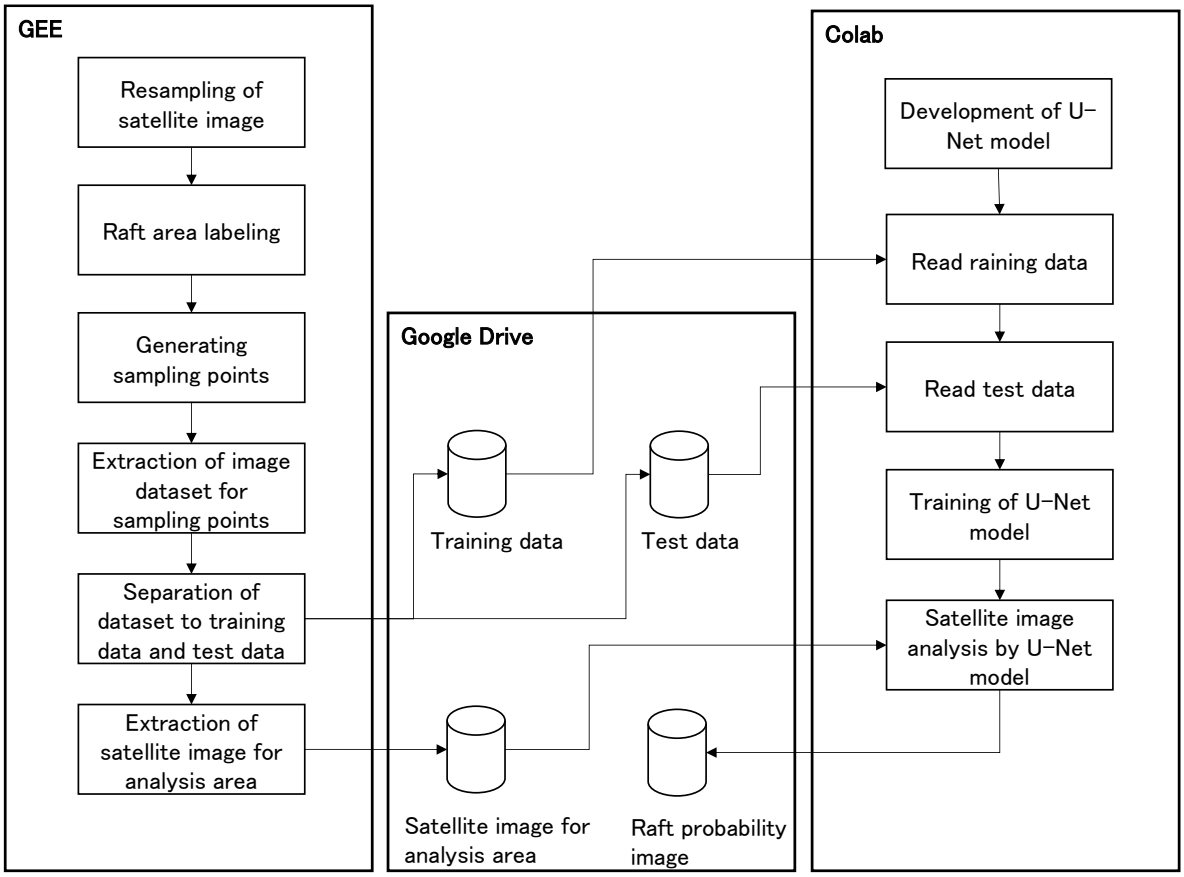



Figure 5 Analysis Flow.

### 3.4 Accuracy assessment

The U-Net-derived raft polygons were compared with those used as truth data described above, and the true-positive (TP), false-positive (FP), and false-negative (FN) polygons were counted. As accuracy indices, precision and recall values, and the F1 score, were calculated using the following equations.

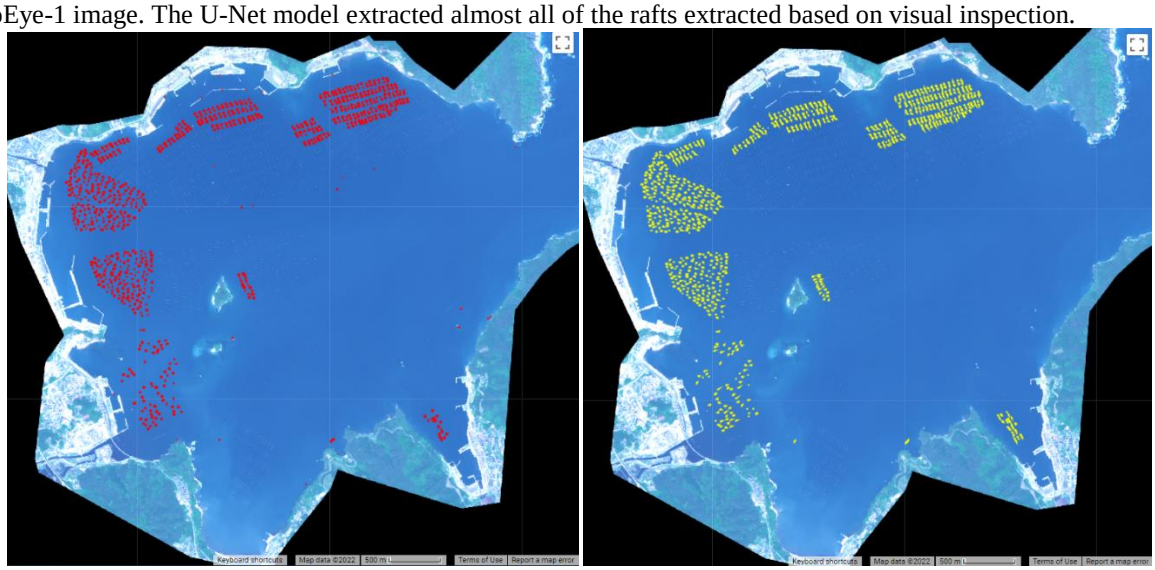
$$\text{Precision} = TP/(TP+FP) \quad (1)$$

$$\text{Recall} = TP/(TP+FN) \quad (2)$$

$$F1 = (2 \times \text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \quad (3)$$

## 4. RESULTS AND DISCUSSION

Figure 6 shows the distribution of raft polygons derived from the U-Net analysis, and through visual inspection of the GeoEye-1 image. The U-Net model extracted almost all of the rafts extracted based on visual inspection.



(a) U-Net model analysis (b) Visual inspection

Figure 6 Rafts polygons identified based on U-Net model analysis and visual inspection of the GeoEye-1 image.

As explained in the Method section, the raft area was extracted from the raft probability image based on a threshold. As the accuracy of the extracted raft area depends on this threshold, several were examined and that providing the maximum F1 score was identified. First, the threshold was changed from 0.1 to 0.9 in 0.1 increments; accuracy indices were calculated for each threshold (Figure 7). Then, near the maximum F1 score, the threshold was adjusted again (in 0.01 increments; Figure 8). The precision value increased from 0.844 to 0.999, while the recall value decreased from 0.989 to 0.971, as the threshold changed from 0.1 to 0.95. The maximum F1 score of 0.990, obtained when the threshold was 0.92, is higher than the score of 0.84 reported in Wang et al. (2018). Table 1 summarizes the numbers of TP, FP and FN when the F1 score was maximum.

Figure 9 shows GeoEye-1 multiband sensor images and the raft detection results for FP. A breakwater edge and small islands can be seen; these are visually similar to rafts, such that misdetection was difficult to avoid. However, the breakwater is obviously longer than a raft and is distinguishable when viewed over a large area.
Figure 10 shows the GeoEye-1 multi band sensor images, raft detection results for FN, and raft polygons based on visual inspection. The outlines of undetected rafts in the image are blurred; these rafts may sink deeper than the detected ones. A lower threshold improved the detection rate slightly, but nine rafts were still not detectable when the threshold was 0.1. Another detection approach may be required for these rafts because they differ visually from the detected ones.

Although there were several misdetected objects and undetected rafts, our results were highly accurate compared with previous studies. A high-spatial-resolution satellite image and image analysis based on deep learning with U-Net enhanced the performance. There are many deep-learning methods besides U-Net and the HCN proposed by Shi et al. (2018). However, CNN-based networks are considered efficient for raft extraction from satellite images. The input data for deep learning are also important. Shi et al. (2018) used a Gaofen-1 satellite image, but converted 16-bit four-band data into 8-bit red, green, blue (RGB) three-band data for use as input data for deep learning. By contrast, we merged four multispectral bands and one panchromatic band from GeoEye-1; the dynamic range was the same as the original 11-bit depth of the input data. Although the input data contained more information due to this process, this necessitates more time for learning and analysis. Thus, either the amount of data or processing time must be prioritized depending on the purpose and analysis environment.

Data extraction from aquaculture facilities using deep learning methods exhibits accuracy close to that achieved

through traditional visual inspection, while the ability to process many images is superior to that of humans. Also, more high-spatial-resolution satellite images are also becoming available. Consequently, automatic monitoring of aquaculture facilities using satellite images may be possible in the near future. Satellite-derived information is expected to enhance the management of aquaculture facilities. For example, the number and position of rafts can be checked quickly and regularly. Automatic analysis of satellite images also facilitates studies of the relationship between aquaculture facilities and the surrounding environment, such as that of Murata et al. (2021).

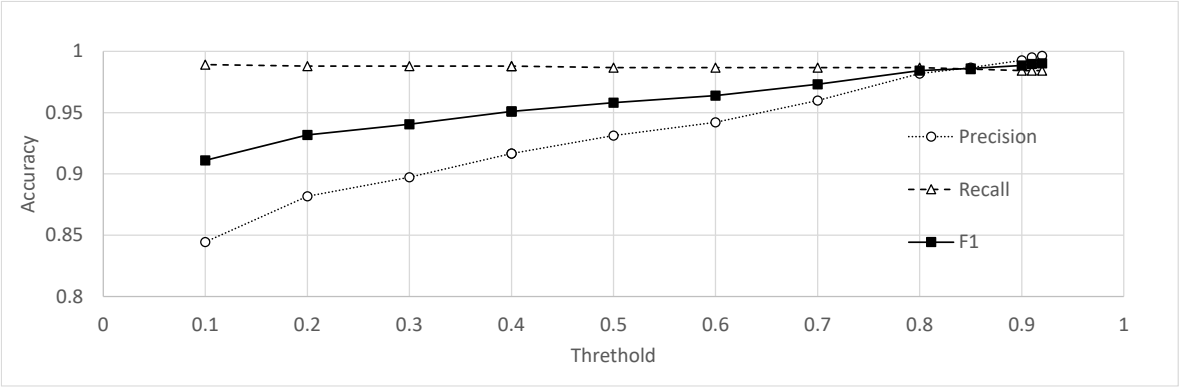



Figure 7 Accuracy indices and thresholds for raft extraction (ranging from 0.1 to 0.95).

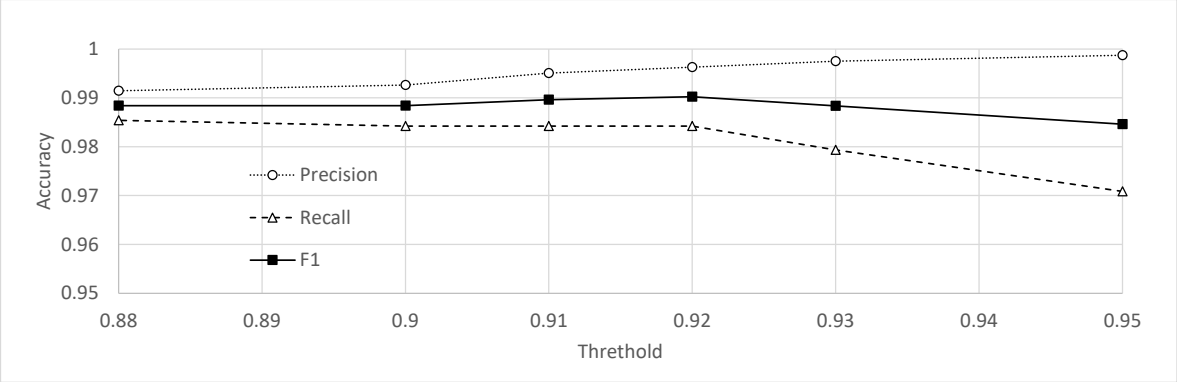



Figure 8 Accuracy indices and thresholds for raft extraction (ranging from 0.88 to 0.95).

Table 1 Accuracy indices when the F1 score is maximal.

| F1 score (Maximum) | Precision | Recall | Number of TP | Number of FP | Number of FN |
|---|---|---|---|---|---|
| 0.990 | 0.996 | 0.984 | 810 | 3 | 13 |


Figure 9 Multiband sensor images and raft detection results for false-positives.

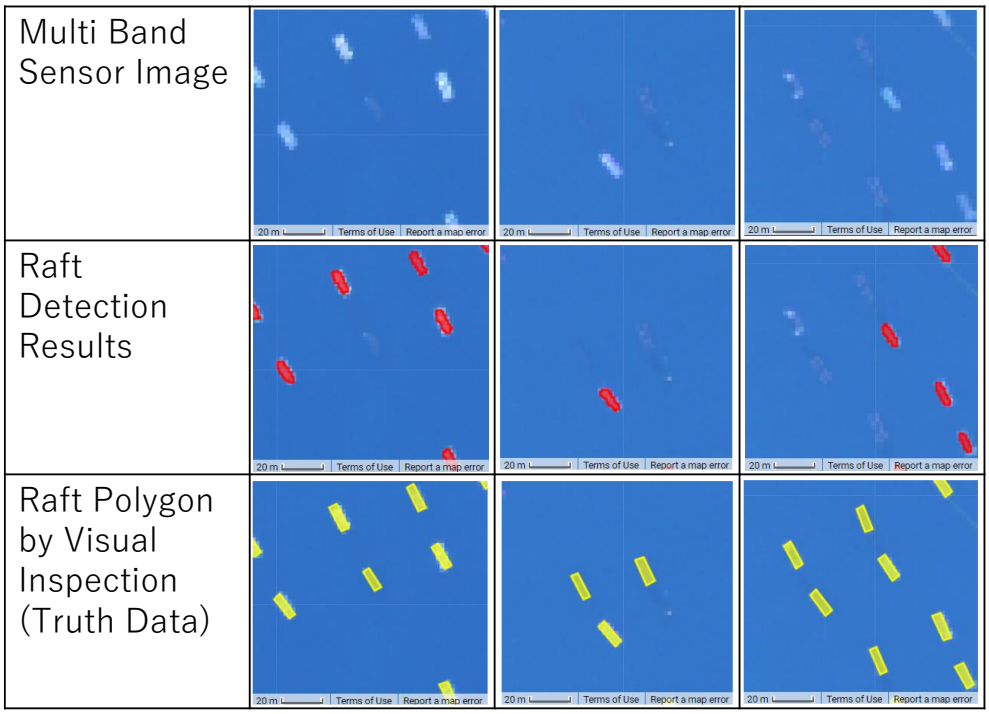


Figure 10 Multiband sensor images and raft detection results for false-negatives.

## 5. CONCLUSION

In this study, aquaculture rafts were extracted from a GeoEye-1 image using U-Net, and a maximum F1 score of 0.990 was obtained. This high accuracy value demonstrates the utility of the U-Net model. Deep-learning-based image analysis methods can map various marine aquaculture facilities with accuracy equal to or better than that based on visual inspection. Also, high-spatial-resolution satellite images are becoming increasingly available. Consequently, automatic monitoring of aquaculture facilities using satellite images may be possible in the near future. Satellite-derived information should facilitate the management of aquaculture facilities and scientific research on their relationship with the surrounding environment.

## 6. REFERENCES

Fisheries Agency, 2021. White paper on fisheries in FY2020, from https://www.maff.go.jp/e/data/publish/attach/pdf/index-199.pdf

Mutsuda, H.; Murakami, K.; Doi, Y.; Yamamoto, T.; Kawaguchi, O., 2011. Seasonal change of oyster raft placement and seawater exchange in Etajima Bay. Journal of Japan Society of Civil Engineers, Ser. B3 (Ocean Eggineering), 67 (2), pp. I_364-I_369.

Murata, H.; Hara, M.; Yonezawa, C.; Komatsu, T., 2021, Monitoring oyster culture rafts and seagrass meadows in Nagatsura-Ura Lagoon, Sanriku Coast, Japan before and after the 2011 Tsunami by remote sensing: Their recoveries implying the sustainable development of coastal waters. PeerJ, 9, pp. e10727.

Wang, Z.; Yang, X.; Liu, Y.; Lu, C., 2018, Extraction of coastal raft cultivation area with heterogeneous water background by thresholding object-based visually salient NDVI from high spatial resolution imagery. Remote Sensing Letters, 9, pp. 839-846.

Shi, T.; Xu, Q.; Zou, Z.; Shi, Z., 2018, Automatic raft labeling for remote sensing images via dual-scale homogeneous convolutional neural network. Remote Sensing, 10, pp. 1130.

Giang, T.L.; Dang, K.B.; Toan Le, Q.; Nguyen, V.G.; Tong, S.S.; Pham, V.-M., 2020, U-Net convolutional networks for mining land cover classification based on high-resolution UAV imagery. IEEE Access, 8, pp. 186257-186273.

Alsabhan, W.; Alotaiby, T.; Dudin, B., 2022, Detecting buildings and nonbuildings from satellite images using U-Net. Computational Intelligence and Neuroscience 2022, pp. 1-13.