Species level mapping of a multispecific seagrass bed using UAV and deep learning technique

Satoru Tahara [1] , **Kenji Sudo** [2] , **Takehisa Yamakita** [3] , **Masahiro Nakaoka** [Corresp. 2]

[1] Graduate School of Environmental Science, Hokkaido University, Sapporo, Hokkaido, Japan
[2] Akkeshi Marine Station, Field Science Center for Northern Biosphere, Hokkaido University, Akkeshi, Hokkaido, Japan
[3] R&D Center for Submarine Resource, Japan Agency for Marine Earth Science and Technology, Yokosuka, Kanagawa, Japan

Corresponding Author: Masahiro Nakaoka
Email address: nakaoka@fsc.hokudai.ac.jp

Seagrass beds are essential habitats in coastal ecosystems, providing valuable ecosystem services, but are threatened by various climate change and human activities. Seagrass monitorings by remote sensing have been conducted over past decades using satellite and aerial images, which have too low resolution to analyze changes in the composition of different seagrass species in multispecific beds. Recently, UAVs have allowed us to obtain much higher resolution images, which is promising in observing fine-scale changes in seagrass species composition. Furthermore, image processing techniques based on deep learning can be applied to discrimination of seagrass species that were difficult based only on color variation. In this study, we conducted mapping of a multispecific seagrass bed in Saroma-ko Lagoon, Hokkaido, Japan, and compared the accuracy of the three discrimination methods of seagrass bed areas and species composition, i.e., pixel-based classification, object-based classification, and the application of deep neural network. We set five taxonomic classes, two seagrass species (*Zostera marina* and *Z. japonica*), brown and green macroalgae, and no vegetation for creating a benthic cover map. High-resolution images by UAV photography enabled us to produce a map at fine scales (<1 cm resolution). The application of a deep neural network successfully classified the two seagrass species. The accuracy of seagrass bed classification was the highest (82%) when the deep neural network was applied. Our results highlighted that a combination of UAV mapping and deep learning could help monitor the spatial extent of seagrass beds and classify their species composition at very fine scales.

Species level mapping of a multispecific seagrass bed using UAV and deep learning technique

Satoru Tahara[1], Kenji Sudo[2], Takehisa Yamakita[3], Masahiro Nakaoka[2]

[1]Graduate School of Environmental Science, Hokkaido University, Sapporo, Hokkaido, Japan
[2]Akkeshi Marine Station, Field Science Center for Northern Biosphere, Hokkaido University, Akkeshi, Hokkaido, Japan
[3]R&D Center for Submarine Resource, Japan Agency for Marine Earth Science and Technology, Yokosuka, Kanagawa, Japan

Corresponding Author:
Masahiro Nakaoka[2]
Akkeshi Marine Station, Field Science Center for Northern Biosphere, Hokkaido University, Akkeshi, Hokkaido, Japan
Email address: nakaoka@fsc.hokudai.ac.jp

Abstract

Seagrass beds are essential habitats in coastal ecosystems, providing valuable ecosystem services, but are threatened by various climate change and human activities. Seagrass monitorings by remote sensing have been conducted over past decades using satellite and aerial images, which have too low resolution to analyze changes in the composition of different seagrass species in multispecific beds. Recently, UAVs have allowed us to obtain much higher resolution images, which is promising in observing fine-scale changes in seagrass species composition. Furthermore, image processing techniques based on deep learning can be applied to discrimination of seagrass species that were difficult based only on color variation. In this study, we conducted mapping of a multispecific seagrass bed in Saroma-ko Lagoon, Hokkaido, Japan, and compared the accuracy of the three discrimination methods of seagrass bed areas and species composition, i.e., pixel-based classification, object-based classification, and the application of deep neural network. We set five taxonomic classes, two seagrass species (*Zostera marina* and *Z. japonica*), brown and green macroalgae, and no vegetation for creating a benthic cover map. High-resolution images by UAV photography enabled us to produce a map at fine scales (<1 cm resolution). The application of a deep neural network successfully classified the two seagrass species. The accuracy of seagrass bed classification was the highest (82%) when the deep neural network was applied. Our results highlighted that a combination of UAV mapping and deep learning could help monitor the spatial extent of seagrass beds and classify their species composition at very fine scales.

Introduction

Seagrasses are angiosperms that inhabit relatively shallow environments along tropical and subarctic coasts, and about 60 species are known worldwide (Short et al., 2007). Seagrasses usually form ~~seagrass~~ beds composed of single or multiple species. While seagrass beds play an essential role in providing valuable ecosystem services, they ~~have been reported to be~~ declining in many parts of the world due to natural and human-induced disturbances (Short & Wyllie-Echeverria, 1996; Waycott et al., 2009, Sudo et al., 2021). Since seagrass distribution and abundance show significant spatiotemporal variability (Tomasko et al., 2005), long-term monitoring of spatial information at each location is essential for ~~deep understanding and~~ appropriate management.

Monitoring of seagrass beds has been conducted using ground-based field surveys (Short et al., 2006), optical remote sensing with aircraft (Kendrick et al., 2000; Sherwood et al., 2017), satellites (Xu et al., 2021; Zoffoli et al., 2021), and acoustic remote sensing (Gumusay et al., 2019). Field surveys can provide detailed information on seagrass cover, species composition, and biomass. However, they are time-consuming and labor-intensive, and the survey area is limited. In contrast, remote sensing methods can obtain large/wide areal distribution information with less effort than field surveys. In addition, it is possible to analyze long-term temporal changes by using aerial photographs (Yamakita, Watanabe & Nakaoka, 2011). While many results have also been reported using satellite data for long-term monitoring (Lyons, Phinn & Roelfsema, 2012; Calleja et al., 2017; Zoffoli et al., 2020; Xu et al., 2021), several limitations have been pointed out for traditional optical remote sensing. The biggest problem is the resolution. The most commonly used satellite data, the Landsat series, provides data over a wide area at a low cost but has a spatial resolution of 30 m which is too low compared to detailed fine-scale information obtained by in-situ field surveys. Phinn et al. (2008) has reported that higher spatial and spectral resolutions are needed for more accurate detailed mapping. Studies using commercial high-resolution satellite images such as WorldView2 and RapidEye have reported high mapping accuracy (Coffer et al., 2020). However, these commercial satellite images are too expensive for long-term, broad-scale monitoring.

In recent years, UAVs (Unmanned Aerial Vehicles, or drones) have been increasingly used in field research due to some advantages compared with conventional remote sensing (Nowak et al., 2019). High spatial resolution data are available by low altitude UAV flights. Frequent flight is possible because the no-cloud sky is unnecessary like satellite, and operation cost is low. It is also possible to adjust the survey time and day, which is impossible with satellites in a fixed orbit. In seagrass research, UAVs have been used for detailed bed mapping (Duffy et al., 2018; Nahirnick et al., 2019; Hobley et al., 2021). Nonetheless, most of these studies mapped seagrass beds consisting of only a single species or conducted mapping without species discrimination.

Seagrasses have different morphologies and life histories depending on the species (Duarte, 1991), and when they live nearby, mapping them by species is necessary to obtain more accurate information such as estimating biomass (Knudby & Nordlund, 2011). ~~It is also known that~~ different species provide different ecosystem services (Mtwana Nordlund et al., 2016) and respond differently to changes in the environment (Roca et al., 2016). Thus, developments of detailed methods that can discriminate different seagrass species are promising for more effective monitoring of seagrass beds~~. It is~~ also ~~helpful~~ for monitoring and managing invasive species (Kumar et al., 2019).

Few studies performed species discrimination of seagrasses with UAV images. Román et al. (2021) showed that seagrass bed mapping, including seagrass discrimination, can be performed with high accuracy using a UAV-mounted ten band multispectral camera and automatic classification based on machine learning algorithms. Chayhard et al. (2018) showed that visual interpretation could be applied to classify seagrass species with different morphology, such as long leaves type (*Enhalus acoroides*) and short leaves type (*Halodule pinifolia* and *H. uninervis*), even using the RGB images taken by UAVs. The camera installed in the consumer-grade UAV is an RGB sensor, and the use of a multispectral camera is costly. Therefore, there is a need for developing methods for seagrass species discrimination using image data with limited spectral resolution but high spatial resolution.

In general, spatial distribution mapping of seagrass beds by optical remote sensing is carried out using classification algorithms (Diesing et al., 2016). Classification algorithms classify the image into several classes such as seagrass, bare sand, and macroalgae by computer. Classification algorithms can be divided into supervised classification and unsupervised classification depending on whether training data are used or not. In supervised classification, which uses ~~ground truth~~ data obtained from field surveys as training data, there are two types of classification: (1) pixel-based classification which classifies each pixel, and (2) object-based classification which classifies each object by grouping similarly colored neighboring pixels. It has been reported that object-based classification provides higher accuracy for high spatial resolution images than pixel-based classification (Gao & Mas, 2008). These classification methods have been used to analyze optical remote sensing data based only on limited image information such as the color, object shape, and size. On the other hand, in ultra-high-resolution UAV images, more features are available, such as the pattern, texture, and location of the objects in the image. A deep neural network (DNN) can automatically extract these various features using a convolutional neural network (CNN), the basic network used for DNN image processing (Traore et al., 2018). The image-to-image translation is one of the applications of DNN. This model is trained with supervised data for transforming the input image into a corresponding output image using the extracted features (Isola et al., 2017). It can be used for semantic segmentation of input images and has also been applied to seagrass bed mapping by remote sensing (Yamakita et al., 2019).

This study aimed to use UAV images and image analysis techniques to create a detailed multispecific seagrass map. The study site was set in a seagrass bed of Saroma-ko Lagoon in northeastern Japan where several seagrass and seaweed species are mixed. We got RGB images by consumer-grade UAV and created a benthic map including the following plant taxa; (1) eelgrass *Zostera marina*, (2) dwarf eelgrass *Z. japonica*, (3) green algae (*Chaetomorpha crassa*, *Cladophora* sp.), and (4) a brown algae (*Cystoseira hakodatensis*). The accuracies of mapping were compared among three methods, (1) conventional pixel-based supervised classification, (2) object-based supervised classification, and (3) image-to-image translation based on DNN method.

Materials & Methods

In this study, we first undertook UAV photography and transect surveys in the field to create reference data, then conducted image analysis in the laboratory. The overall workflow is shown in Fig. 1.

Fieldwork

Fieldwork was carried out on July 9, 2019 at Saroma-ko Lagoon in eastern Hokkaido, Japan (Fig. 2). Saroma-ko Lagoon is a brackish lagoon of about 152 km^2 and is connected to the Sea of Okhotsk by two channels, one about 300 m in width and another 50 m. The maximum depth of the lagoon is 19.6 m. Three species of seagrasses (*Zostera japonica*, *Z. marina*, and *Z. caespitosa*) occur along the intertidal and shallow subtidal zones of the lagoon (Biodiversity Center of Japan, 2008). ~~The present~~ study was conducted in a seagrass bed at the eastern coast of the lagoon (Fig. 2).

The transect survey and UAV photography were conducted during a low tide. In the transect survey, a transect line was set perpendicular to the shoreline from the shallowest end in the east to the deepest part of the bed in the west until no seagrass appeared (about 600 m offshore). A total of 86 quadrats of 0.25 m^2 were placed haphazardly along the transect to cover all present seagrasses and macroalgae along the transect, and species and cover were recorded. Surveys were conducted by wading, snorkeling, and SCUBA diving.

UAV photography was conducted from shore using a quadcopter Mavic2 pro (DJI Co. Ltd). The flight area was set at 580 m offshore and 90 m wide, including a measuring tape used for the transect. We took the images with the RGB sensor camera equipped with the Mavic2 pro at a nadir angle. The flight was automated using DroneDeploy (DroneDeploy Co. Ltd.). DroneDeploy enables automatic flight and photography by specifying the flight area, altitude, and overlap rate (front and side) between images. To ensure sufficient spatial resolution for seagrass species identification and to enable orthorectification, we used the setting for DroneDeploy as follows: altitude 30 m, front overlap 80 % and side overlap 70 %. The camera settings were set manually before the shooting and were not changed (aperture: f/2.8, shutter speed: 1/400 s, and ISO: 200).

Image pre-processing

The captured UAV images were orthorectified using the SfM-MVS processing software Metashape ver. 1.7.1 (Agisoft Co. Ltd.). Through SfM-MVS processing, we can produce an orthoimage from overlapped images(Verhoeven et al., 2013). Then, the images were cropped for subsequent analyses.The orthoimage was first converted to a benthic cover map by visual interpretation. As a result of the transect survey, three species of seagrass (*Z. marina*, *Z. japonica*, *Z. caespitosa*), green algae (*Chaetomorpha crassa*, *Cladophora* sp.), a brown alga (*Cystoseira hakodatensis*), and red algae (Ceramiaceae gen spp.) were observed. Three seagrass species were continuously mixed and the dominant species changed with water depth; *Z. japonica* (intertidal), *Z. marina* (shallower subtidal), and *Z. caespitosa* (deeper subtidal). *Zostera caespitosa* was difficult to distinguish from *Z. marina* without observing the belowground part, so the area offshore of 300 m from the shoreline where *Z. caespitosa* occurred was cropped and excluded from subsequent analysis of orthoimage. This cropping resulted in a total area of 7,884 m^2, 291 m along the depth axis and 27 m horizontally to the depth axis. As for macroalgae, red algae were found only in a limited area and were not distinguishable from other vegetation by the naked eye, so they were excluded from the classification. Green algae were combined into one class because it was difficult to distinguish the two species.

These resulted in five taxonomic classes in this study (*Z. marina* (ZM), *Z. japonica* (ZJ), green algae (GA), brown algae (BA), and no vegetation (NV)). The interpreter who conducted a field survey could distinguish these five classes on orthoimage and hand-traced the boundaries of each class on an image editing software, Paint. NET ver.4.2.16 (dotPDN LLC.). This study used the maps created by visual interpretation as ground-truth images for training and accuracy verification data. To examine the credibility of the visual interpretation, we compared the ground-truth images with the data obtained from the transect survey. For the comparison, the location of each quadrat was first identified on the orthoimage based on the measurement tape used for the transect installation, and the dominant vegetation classes (ZM, ZJ, GA, BA) were examined. Next, the area corresponding to the quadrat area was cropped from the ground-truth image. The dominant taxonomic classes were examined in the same way and compared with the results of the transect survey. In all cases, however, if the coverage of the dominant class was less than 10%, the no vegetation class (ND) was considered the dominant class.

Mapping method comparison

Mapping by visual interpretation is highly accurate but requires extensive labor. This study compared three mapping methods (pixel-, object-based classification and image-to-image translation based on DNN) to find a more efficient and reproducible method. All ~~methods~~ are supervised methods, which means that by training the computer using ~~some of the data as~~ training data, mapping can be done automatically for the rest of the data. In this study, we trained each method using the ground-truth image by visual interpretation. About half of the orthoimage (54%) was used as a training area and the rest (46%) as a validation area from which accuracy assessment was conducted for each method.

a). Conventional mapping (Pixel-based and object-based classification)

Pixel-based and object-based classification ~~is a~~ standard mapping method for remote sensing images (Dat Pham et al., 2019). It is a supervised classification in which data in some areas are used as training data to classify data in other areas. In this study, the training data for empirical mapping and classification were created on ArcGIS pro ver. 2.8.1 (Esri Co. Ltd.). Pixel-based classification classifies each pixel, while object-based classification classifies each object. An object is a collection of similarly colored neighboring pixels created by the segmentation of the input image. For segmentation, three parameters were adjusted until the object became an appropriate size (Spectral detail: 20, Spatial detail: 5, Minimum segment size: 500).

In this study, the algorithm used for classification was ~~the~~ support vector machine (SVM), which was used in seagrass mapping and reported to be sufficiently accurate (Pottier et al., 2021). SVM is not sensitive to training data size and does not assume the probability distribution of the data (Mountrakis, Im & Ogole, 2011). The training data were polygons created from a ground-truth image by uniformly selecting a representative area of each specific class. The area (number) of training data for each class (ZM, ZJ, GA, BA, and NV) was 140 m^2 (9), 23.7 m^2 (11), 8.01 m^2 (7), and 1.24 m^2 (9), respectively.

b). Image translation based on deep learning (pix2pix)

Pix2pix is an image-to-image translation model based on conditional generative adversarial networks (cGANs) (Isola et al., 2017). cGANs are the application of CNN and have two networks: generator and discriminator. The generator transforms the input image, and the discriminator classifies translated image as fake or real by comparing it with the ground-truth image. The generator and discriminator compete with each other, and the generator comes to transform the image into a more realistic one. This model can also be used for remote sensing mapping by translating images to classified images and showed higher accuracy than other deep learning models (Isola et al., 2017). Pix2pix has been applied to various examples, including seagrass mapping for black-and-white aerial photography (Yamakita et al., 2019).

The translation process in pix2pix requires the size of the input image to be 256 x 256 pixels. Therefore, the training and validation data were sliced to an appropriate size beforehand. After slicing the orthoimages, number of training and validation data were 980 and 840. In general, DNNs are trained more robustly with increasing training data. Therefore, we added flipped copies of the training data to increase the data for training ~~in this study~~. We added horizontal, vertical, and simultaneous horizontal and vertical flipped copies of the training data. After all, the number of training data was 3920.

GANs-based networks often suffer from a problem called mode collapse (Goodfellow, 2016). This occurs when the training data contain a lot of similar ground-truth images. In such cases, the translated image by the network would also result in similar images. In the study area, the percentage of the ZM area is high, and ~~a lot~~ of ground-truth data of the training data are dominated by ZM only, which can cause ~~the~~ mode collapse. We divided the training data ZM into three subclasses to solve this problem. We reduced colors in the orthoimage of the ZM area to three by posterization and assigned a subclass to each of them. This prevented homogenization of the ground truth image (Fig. 3).

Accuracy assessment

Accuracy assessment was performed by comparing the mapping results of each method in the validation area with the ground truth data. Five thousand random points were extracted in the validation area, and a confusion matrix was created for each resulting map. The confusion matrix was used to calculate the overall accuracy (OA) and Kappa coefficient (K) for all classes and the user accuracy (UA) and producer accuracy (PA) for individual classes. OA represents the ratio of the pixel classified correctly. K is a statistic value that expresses the degree of agreement between data, taking into account coincidence (Cohen, 1960). $K = 0$ means that the degree of agreement is equal to that obtained by chance, and positive values indicate a degree of agreement greater than chance, with the maximum value of 1. In general, the relationship between K and strength of agreement is < 0.00: poor, 0.00-0.20: slight, 0.21-0.40: fair, 0.41-0.60: moderate, and 0.61-0.80: substantial (Landis & Koch, 1977). UA is the ratio of each class assigned by the correctly classified mapping, and PA is the ratio of each class assigned by ground truth that is correctly classified.

Results

Image pre-processing

The flight time of the UAV photography was 22 minutes, and 406 out of 534 taken images were used to orthorectification. The remaining 128 images were taken in deep water where the seagrass was submerged entirely, and they could not be used for the synthesis because there were few matching points.

The spatial resolution of the created orthoimage was 8.13 mm/pix (Fig. 4 A). After cropping orthoimage for analysis, the number of quadrats for the transect survey included in image was 42. Among them, those dominated by *Z. marina*, *Z. japonica*, green algae, brown algae, red alga and no vegetation were 18, 11, 4, 1, 3, and 5, respectively. In the ground-truth image, a part of *Z. japonica* (4/11) and green algae (1/4) were misclassified as adjacent *Z. marina* or no vegetation. Similarly, the red algae that did not appear in the ground-truth images were classified as *Z. marina* or no vegetation (Table 1). Overall accuracy and Kappa of visual interpretation data (Fig. 4 B) validated by the field data were 0.786 and 0.687, respectively.

Mapping and accuracy assessment

The results of mapping generally agreed among the three different methods although some misidentifications were observed (Fig. 5). For example, small gaps (no vegetation) in deeper parts were not identified by the pixel-based and object-based methods. In contrast, GA in the shallower parts were overestimated by the pixel-based method.

Accuracy assessment showed that the values of OA and K were highest for DNN, followed by the object-based, and the pixel-based methods (Table 2). K value for DNN exceeded 0.6, indicating substantial agreement, whereas that for the pixel-based methods was less than 0.2, showing poor fit. Pixel-based method showed lowest accuracy because speckles are observed overall in the result map (Fig. 5 A).

Accuracy by species, shown by the values of UA and PA, also varied greatly (Table 2). ZM, which accounted for the most significant percentage of the study area, showed the highest accuracy for every method. However, PA of the pixel-based classification of ZM (0.421) was much lower than UA (0.781), indicating overestimation. ZJ showed low accuracy in the pixel-based and object-based classifications (0.020-0.403), but higher in DNN (> 0.5). ZJ mainly misclassified to ZM and NV in the pixel-based and object-based classification, and these methods could hardly discriminate seagrass species. DNN similarly misclassified ZJ to ZM and NV, but a relatively small extent. GA showed lowest accuracy for almost all ~~the~~ methods. The UA was highest for the object-based classification (0.733) for BA, while the PA was higher for the pixel-based classification (0.625). NV showed higher UA than PA for all methods, indicating underestimation mainly due to misclassification to ZM.

Discussion

This study shows that the mapping method based on the combined use of UAV photography and DNN-based image-to-image translation is more accurate than ~~the~~ conventional methods, especially on species-by-species identification of seagrass and seaweed species in a multispecific seagrass bed.

Previous studies have attempted to discriminate species of seagrass and macroalgae using satellite and aerial images (e.g., Phinn et al., (2008), Kovacs et al., (2018)). Although comparisons should be made with caution due to the difference in sites and methods, the accuracy in our study (OA: 0.818) outperforms those by other studies (OA: 0.23 and 0.28 for Phinn et al., (2008), and 0.64~0.69 for Kovacs et al., (2018)). The grain size of our seagrass bed map (8.13 mm/pix) was much higher than these previous studies (2.4 and 4 m/pix for Phinn et al., (2008), and 2~30 m/pix for Kovacs et al., (2018)), indicating that the spatial resolution was a key factor for successful classification of different plants in multispecific seagrass meadows. In this study, however, spatial extent was small (7,884 m^2), covering only 0.035 % of the seagrass beds in Saroma-ko Lagoon (22.5 km^2 in 2015, Hokkaido Aquaculture Promotion Cooperation 2015). Linear extrapolation indicates that it would take us more than 1,000 hours (i.e. > 40 days) to cover the whole seagrass bed by this method, which is not practical considering the labor and seasonal changes in the seagrass bed.

To increase the accuracy of seagrass bed mapping, previous studies have used hyperspectral sensors aboard on satellites and aircraft for species discrimination. This is because seagrasses and seaweed species can be discriminated with different spectral reflectance as well as terrestrial plants (Fyfe, 2003). Although light-weight hyperspectral sensors that can be mounted on UAVs have been developed, they have not been widely used yet because they require complex pre- and post-flight operations for analysis (Adão et al., 2017). Furthermore, it may be difficult even with hyperspectral sensors to discriminate closely related congeneric species of seagrass which have similar characteristics of leaf color. Due to this limitation, it is more effective to develop a new discrimination method using information other than color in the visible band with UAV mounted RGB sensors, which are already used in the field of seagrass research. In this study, mapping based on the conventional classification methods using RGB showed very low overall accuracy, and especially for discriminating ZJ which were misclassified to ZM and NV. On the other hand, the mapping based on DNN showed higher accuracy than the conventional methods. This highlights the advantage of the DNN method, in which the computer can extract and use much more information than just color information from the UAV images (Albawi, Mohammed & Al-Zawi, 2017). Therefore, image data with limited spectral information can be analyzed in a more sophisticated way by applying DNN.

In contrast to discrimination of different *Zostera* species, results of green algae (GA) classification were in low accuracy in all ~~the~~ methods. When comparing GA visually in the orthoimage of the training and validation areas, the giant patch of green algae in the validation area has a bright green color that is not seen in the training area ~~(Fig. 6).~~ This is due to the difference in the species composition of the GA, which is mainly *Cladophora* sp. in this brighter clump, and mainly darker *Chaetomorpha crassa* in the rest. Since these green algae were mixed in the patch, it was not easy to separate them into different classes by visual interpretation. The training data did not sufficiently cover the variability of GA, which may be the reason for the low accuracy. This indicates that even when we use DNNs, there is a limit to their versatility, and performance varies by different seagrass and seaweed species. Higher generalizability will be possible by increasing the variety of training data that sufficiently cover the variability in the study area.

It has been reported in previous studies that the salt-pepper phenomenon, defined as individual pixels classified differently from their neighbors (Yu et al., 2006). reduces the classification accuracy of pixel-based classification for high-resolution images (Feng, Liu & Gong, 2015). The salt-pepper phenomenon is caused by the internal variability within a classification class that appears as noise in the classification results. In this study, the salt-pepper phenomenon was also observed, and it is one of the factors causing the low accuracy of pixel-based classification ~~(Fig. 5)~~. On the other hand, the results of object-based classification showed that segmentation suppressed the salt-pepper phenomenon ~~(Fig. 5),~~ making it a more suitable method for high-resolution images. Result of DNN also showed no sand-pepper phenomenon.

In this study, we used a ground-truth image produced by visual interpretation for training data to secure the amount of training data. The machine learning algorithm used in this study, SVM, is known to be more sensitive to the quality of training data than its size (Mountrakis et al., 2011). Therefore, the training data is prepared to represent each class in the training area, and we do not need the entire ground-truth image. On the other hand, the importance of the size of training data is confirmed for DNN by the fact that data argumentation is common to improve algorithm in previous studies (Mikolajczyk & Grochowski, 2018). Therefore, the DNN method is inferior in terms of the time and effort required to prepare the training data. In this study, it took only a few hours to prepare the training data for pixel-based and object-based classification, but it took several days for the DNN method. In addition, the ground-truth images created by visual interpretation contained some errors when validated by the ground-truth data (Table 2). Currently, there are examples of underwater photo datasets available for seagrass detection (Reus et al., 2018), but there are no available datasets with labeled aerial seagrass images. Therefore, researchers applying DNNs will need to start by creating a dataset by themselves. However, it is still worth considering the application of DNN because it is expected to achieve highly accurate mapping. Acquiring and creating ground-truth data with quality and quantity is a future challenge.

Conclusions

This study reports the result of a case study which apply UAVs and deep learning technique at multispecific seagrass bed. Image-to-image translation based on deep neural network could discriminate seagrass species and macroalgae, and show higher accuracy than conventional classification methods. UAVs enable easier acquisition of high spatial and temporal resolution data that was previously difficult to obtain by other remote sensing devices. DNN is especially useful when we can obtain high-resolution images by UAVs with conventional cameras with limited spectral range. Some challenges remain, such as limitation in covering wide areas for the mapping, and in labors for preparing ground truth data. Nevertheless, UAV detailed mapping at coastal area enables scientists further biological research of submerged vegetation based on spatial information.

Acknowledgements

We thank Mizuho Namba and Minako Abe Ito for their help with field surveys, and K. Sakaguchi and other members of the Aquaculture Fishery Cooperative of Saroma Lake for their logistical support during our field work in Saroma. We highly appreciate Minako Abe Ito for improving the manuscript.

Funding

Funding was provided by Saroma-ko Aquaculture Cooperation, Japan Society for the Promotion of Science (Kakenhi 16H01792) and Environmental Restoration and Conservation Agency (Environment Research and Technology Development Fund, S-15 Predicting and Assessing Natural Capital and Ecosystem Services (PANCES). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

・Satoru Tahara conducted the fieldwork, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

・Kenji Sudo conducted the fieldwork, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

・Takehisa Yamakita conceived and designed the research, authored or reviewed drafts of the paper, and approved the final draft.

· Masahiro Nakaoka conceived and designed the research, conducted the fieldwork, authored or reviewed drafts of the paper, and approved the final draft.

References

Adão T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, Sousa J. 2017. Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry. Remote Sensing 9:1110. DOI: 10.3390/rs9111110.

Albawi S, Mohammed TA, Al-Zawi S. 2017. Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). IEEE, 1–6. DOI: 10.1109/ICEngTechnol.2017.8308186.

Biodiversity Center of Japan. The Report for the 7th Natural Environmental Survey Shallow Marine Eco- system Survey (Aquatic Vegetation Survey) [Internet]. Yamanashi, Japan; 2008. Available: http://www. biodic.go.jp/reports2/6th/6_moba19/6_moba19.pdf

Calleja F, Galván C, Silió-Calzada A, Juanes JA, Ondiviela B. 2017. Long-term analysis of Zostera noltei: A retrospective approach for understanding seagrasses' dynamics. Marine Environmental Research 130:93–105. DOI: 10.1016/j.marenvres.2017.07.017.

Chayhard S, Manthachitra V, Nualchawee K, Buranapratheprat A. 2018. Application of unmanned aerial vehicle to estimate seagrass biomass in Kung Kraben Bay, Chanthaburi province, Thailand. International Journal of Agricultural Technology 14:1107–1114.

Coffer MM, Schaeffer BA, Zimmerman RC, Hill V, Li J, Islam KA, Whitman PJ. 2020. Performance across WorldView-2 and RapidEye for reproducible seagrass mapping. Remote Sensing of Environment 250:112036. DOI: 10.1016/j.rse.2020.112036.

Cohen J. 1960. A coefficient of agreement for nominal scales. Educational And Psychological Measurement XX:37–46.

Dat Pham T, Xia J, Thang Ha N, Tien Bui D, Nhu Le N, Tekeuchi W. 2019. A review of remote sensing approaches for monitoring blue carbon ecosystems: Mangroves, sea grasses and salt marshes during 2010–2018. Sensors (Switzerland) 19. DOI: 10.3390/s19081933.

Diesing M, Mitchell P, Stephens D. 2016. Image-based seabed classification: what can we learn from terrestrial remote sensing? ICES Journal of Marine Science: Journal du Conseil 73:2425–2441. DOI: 10.1093/icesjms/fsw118.

Duarte CM. 1991. Allometric scaling of seagrass form and productivity. Marine Ecology Progress Series 77:289–300.

Duffy JP, Pratt L, Anderson K, Land PE, Shutler JD. 2018. Spatial assessment of intertidal seagrass meadows using optical imaging systems and a lightweight drone. Estuarine, Coastal and Shelf Science 200:169–180. DOI: 10.1016/j.ecss.2017.11.001.

Feng Q, Liu J, Gong J. 2015. UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis. Remote Sensing 7:1074–1094. DOI: 10.3390/rs70101074.

Fyfe SK. 2003. Spatial and temporal variation in spectral reflectance: Are seagrass species spectrally distinct? Limnology and Oceanography 48:464–479. DOI: 10.4319/lo.2003.48.1_part_2.0464.

Gao Y, Mas J. 2008. A comparison of the performance of pixel based and object based classifications over images with various spatial resolutions. Online journal of earth sciences 2:27–35.

Goodfellow I. 2016. NIPS 2016 Tutorial: Generative Adversarial Networks.

Gumusay MU, Bakirman T, Tuney Kizilkaya I, Aykut NO. 2019. A review of seagrass detection, mapping and monitoring applications using acoustic systems. European Journal of Remote Sensing 52:1–29. DOI: 10.1080/22797254.2018.1544838.

Hobley B, Arosio R, French G, Bremner J, Dolphin T, Mackiewicz M. 2021. Semi-Supervised Segmentation for Coastal Monitoring Seagrass Using RPA Imagery. Remote Sensing 13:1741. DOI: 10.3390/rs13091741.

Hokkaido Aquaculture Promotion Cooperation (2015) Report on seagrass bed distribution in Saroma-ko Lagoon. Hokkaido Aquaculture Promotion Cooperation, 84pp.

Isola P, Zhu JY, Zhou T, Efros AA. 2017. Image-to-image translation with conditional adversarial networks. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017-Janua:5967–5976. DOI: 10.1109/CVPR.2017.632.

Kendrick GA, Hegge BJ, Wyllie A, Davidson A, Lord DA. 2000. Changes in seagrass cover on Success and Parmelia Banks, Western Australia between 1965 and 1995. Estuarine, Coastal and Shelf Science 50:341–353. DOI: 10.1006/ecss.1999.0569.

Knudby A, Nordlund L. 2011. Remote sensing of seagrasses in a patchy multi-species environment. International Journal of Remote Sensing 32:2227–2244. DOI: 10.1080/01431161003692057.

Kovacs E, Roelfsema C, Lyons M, Zhao S, Phinn S. 2018. Seagrass habitat mapping: how do Landsat 8 OLI, Sentinel-2, ZY-3A, and Worldview-3 perform? Remote Sensing Letters 9:686–695. DOI: 10.1080/2150704X.2018.1468101.

Kumar A, Cooper C, Remillard CM, Ghosh S, Haney A, Braun F, Conner Z, Page B, Boyd K, Wilde S, Mishra DR. 2019. Spatiotemporal monitoring of hydrilla [Hydrilla verticillata (L. f.) Royle] to aid management actions. Weed Technology 33:518–529. DOI: 10.1017/wet.2019.13.

Landis JR, Koch GG. 1977. The Measurement of Observer Agreement for Categorical Data. Biometrics 33:159. DOI: 10.2307/2529310.

Lyons MB, Phinn SR, Roelfsema CM. 2012. Long term land cover and seagrass mapping using Landsat and object-based image analysis from 1972 to 2010 in the coastal environment of South East Queensland, Australia. ISPRS Journal of Photogrammetry and Remote Sensing 71:34–46. DOI: 10.1016/j.isprsjprs.2012.05.002.

Mikolajczyk A, Grochowski M. 2018. Data augmentation for improving deep learning in image classification problem. In: 2018 International Interdisciplinary PhD Workshop (IIPhDW). IEEE, 117–122. DOI: 10.1109/IIPHDW.2018.8388338.

Mountrakis G, Im J, Ogole C. 2011. Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing 66:247–259. DOI: 10.1016/j.isprsjprs.2010.11.001.

Mtwana Nordlund L, Koch EW, Barbier EB, Creed JC. 2016. Seagrass Ecosystem Services and Their Variability across Genera and Geographical Regions. PLOS ONE 11:e0163091. DOI: 10.1371/journal.pone.0163091.

Nahirnick NK, Reshitnyk L, Campbell M, Hessing‐Lewis M, Costa M, Yakimishyn J, Lee L. 2019. Mapping with confidence; delineating seagrass habitats using Unoccupied Aerial Systems (UAS). Remote Sensing in Ecology and Conservation 5:121‐135. DOI: 10.1002/rse2.98.

Nowak MM, Dziób K, Bogawski P. 2019. Unmanned Aerial Vehicles (UAVs) in environmental biology: a review. European Journal of Ecology 4:56–74. DOI: 10.2478/eje-2018-0012.

Phinn S, Roelfsema C, Dekker A, Brando V, Anstee J. 2008. Mapping seagrass species, cover and biomass in shallow waters: An assessment of satellite multi-spectral and airborne hyper-spectral imaging systems in Moreton Bay (Australia). Remote Sensing of Environment 112:3413–3425. DOI: 10.1016/j.rse.2007.09.017.

Pottier A, Catry T, Trégarot E, Maréchal J-P, Fayad V, David G, Sidi Cheikh M, Failler P. 2021. Mapping coastal marine ecosystems of the National Park of Banc d'Arguin (PNBA) in Mauritania using Sentinel-2 imagery. International Journal of Applied Earth Observation and Geoinformation 102:102419. DOI: 10.1016/j.jag.2021.102419.

Reus G, Möller T, Jager J, Schultz ST, Kruschel C, Hasenauer J, Wolff V, Fricke-Neuderth K. 2018. Looking for seagrass: Deep learning for visual coverage estimation. 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans, OCEANS - Kobe 2018:2–7. DOI: 10.1109/OCEANSKOBE.2018.8559302.

Roca G, Alcoverro T, Krause-Jensen D, Balsby TJS, Van Katwijk MM, Marbà N, Santos R, Arthur R, Mascaró O, Fernández-Torquemada Y, Pérez M, Duarte CM, Romero J. 2016. Response of seagrass indicators to shifts in environmental stressors: A global review and management synthesis. Ecological Indicators 63:310–323. DOI: 10.1016/j.ecolind.2015.12.007.

Román A, Tovar-Sánchez A, Olivé I, Navarro G. 2021. Using a UAV-Mounted Multispectral Camera for the Monitoring of Marine Macrophytes. Frontiers in Marine Science 8:1–12. DOI: 10.3389/fmars.2021.722698.

Sherwood ET, Greening HS, Johansson JOR, Kaufman K, Raulerson GE. 2017. Tampa Bay (Florida, USA): Documenting Seagrass Recovery since the 1980's and Reviewing the Benefits. Southeastern Geographer 57:294–319. DOI: 10.1353/sgo.2017.0026.

Short F, Carruthers T, Dennison W, Waycott M. 2007. Global seagrass distribution and diversity: A bioregional model. Journal of Experimental Marine Biology and Ecology 350:3–20. DOI: 10.1016/j.jembe.2007.06.012.

Short FT, Koch EW, Creed JC, Magalhães KM, Fernandez E, Gaeckle JL. 2006. SeagrassNet monitoring across the Americas: Case studies of seagrass decline. Marine Ecology 27:277–289. DOI: 10.1111/j.1439-0485.2006.00095.x.

Short FT, Wyllie-Echeverria S. 1996. Natural and human-induced disturbance of seagrasses. Environmental Conservation 23:17–27. DOI: 10.1017/s0376892900038212.

Sudo K, Quiros TEAL, Prathep A, Luong C Van, Lin H-J, Bujang JS, Ooi JLS, Fortes MD, Zakaria MH, Yaakub SM, Tan YM, Huang X, Nakaoka M. 2021. Distribution, Temporal Change, and Conservation Status of Tropical Seagrass Beds in Southeast Asia: 2000–2020. Frontiers in Marine Science 8:1–11. DOI: 10.3389/fmars.2021.637722.

Tomasko DA, Corbett CA, Greening HS, Raulerson GE. 2005. Spatial and temporal variation in seagrass coverage in Southwest Florida: Assessing the relative effects of anthropogenic nutrient load reductions and rainfall in four contiguous estuaries. Marine Pollution Bulletin 50:797–805. DOI: 10.1016/j.marpolbul.2005.02.010.

Traore BB, Kamsu-Foguem B, Tangara F. 2018. Deep convolution neural network for image recognition. Ecological Informatics 48:257–268. DOI: 10.1016/j.ecoinf.2018.10.002.

Verhoeven G, Sevara C, Karel W, Ressl C, Doneus M, Briese C. 2013. Undistorting the Past: New Techniques for Orthorectification of Archaeological Aerial Frame Imagery. In: 31–67. DOI: 10.1007/978-3-319-01784-6_3.

Waycott M, Duarte CM, Carruthers TJB, Orth RJ, Dennison WC, Olyarnik S, Calladine A, Fourqurean JW, Heck KL, Hughes AR, Kendrick GA, Kenworthy WJ, Short FT, Williams SL. 2009. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proceedings of the National Academy of Sciences 106:12377–12381. DOI: 10.1073/pnas.0905620106.

Xu SS, Xu SS, Zhou Y, Yue S, Zhang X, Gu R, Zhang Y, Qiao Y, Liu M. 2021. Long-term changes in the unique and largest seagrass meadows in the Bohai sea (China) using satellite (1974–2019) and sonar data: Implication for conservation and restorationc. Remote Sensing 13:1–16. DOI: 10.3390/rs13050856.

Yamakita T, Sodeyama F, Whanpetch N, Watanabe K, Nakaoka M. 2019. Application of deep learning techniques for determining the spatial extent and classification of seagrass beds, Trang, Thailand. Botanica Marina 62:291–307. DOI: 10.1515/bot-2018-0017.

Yamakita T, Watanabe K, Nakaoka M. 2011. Asynchronous local dynamics contributes to stability of a seagrass bed in Tokyo Bay. Ecography 34:519–528. DOI: 10.1111/j.1600-0587.2010.06490.x.

Yu Q, Gong P, Clinton N, Biging G, Kelly M, Schirokauer D. 2006. Object-based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery. Photogrammetric Engineering & Remote Sensing 72:799–811.

Zoffoli ML, Gernez P, Godet L, Peters S, Oiry S, Barillé L. 2021. Decadal increase in the ecological status of a North-Atlantic intertidal seagrass meadow observed with multi-mission satellite time-series. Ecological Indicators 130:108033. DOI: 10.1016/j.ecolind.2021.108033.

Zoffoli ML, Gernez P, Rosa P, Le Bris A, Brando VE, Barillé A-L, Harin N, Peters S, Poser K, Spaias L, Peralta G, Barillé L. 2020. Sentinel-2 remote sensing of Zostera noltei-dominated intertidal seagrass meadows. Remote Sensing of Environment 251:112020. DOI: 10.1016/j.rse.2020.112020.

Figure 1

Methodology workflow of this study

Parallelograms, rectangles and arrows represent input/output data, data processes and data flows, respectively

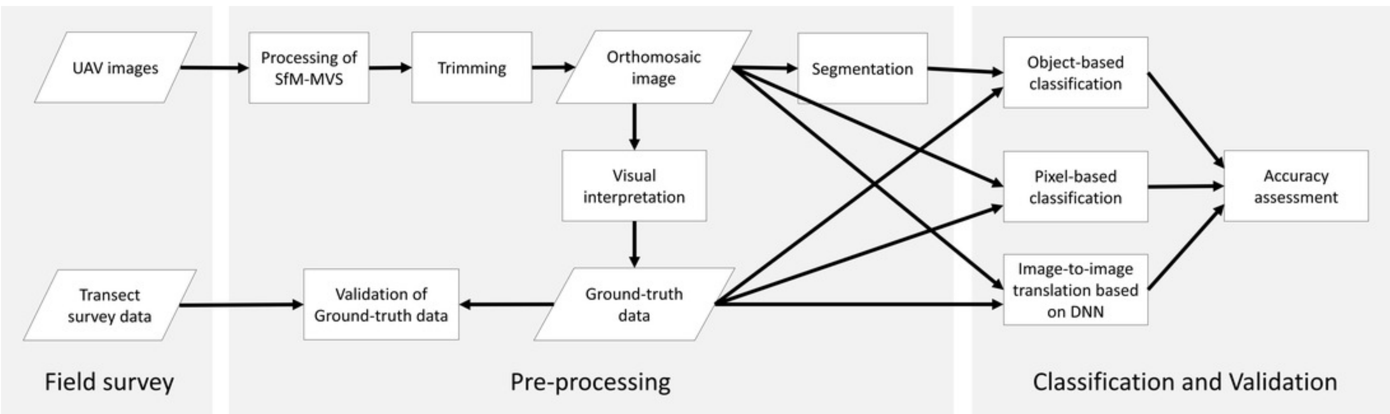

Figure 2

Study site

(A) Study site is located at Saroma-ko Lagoon in eastern Hokkaido, Japan. (B) Black point indicates the location of this study, black triangles show the channels connecting Saroma-ko Lagoon to the Sea of Okhotsk. (C) Seagrass bed extent along the eastern shore of Saroma-ko. The UAV flight area and cropped area are shown as a red rectangle, transect line is shown as a white solid line.

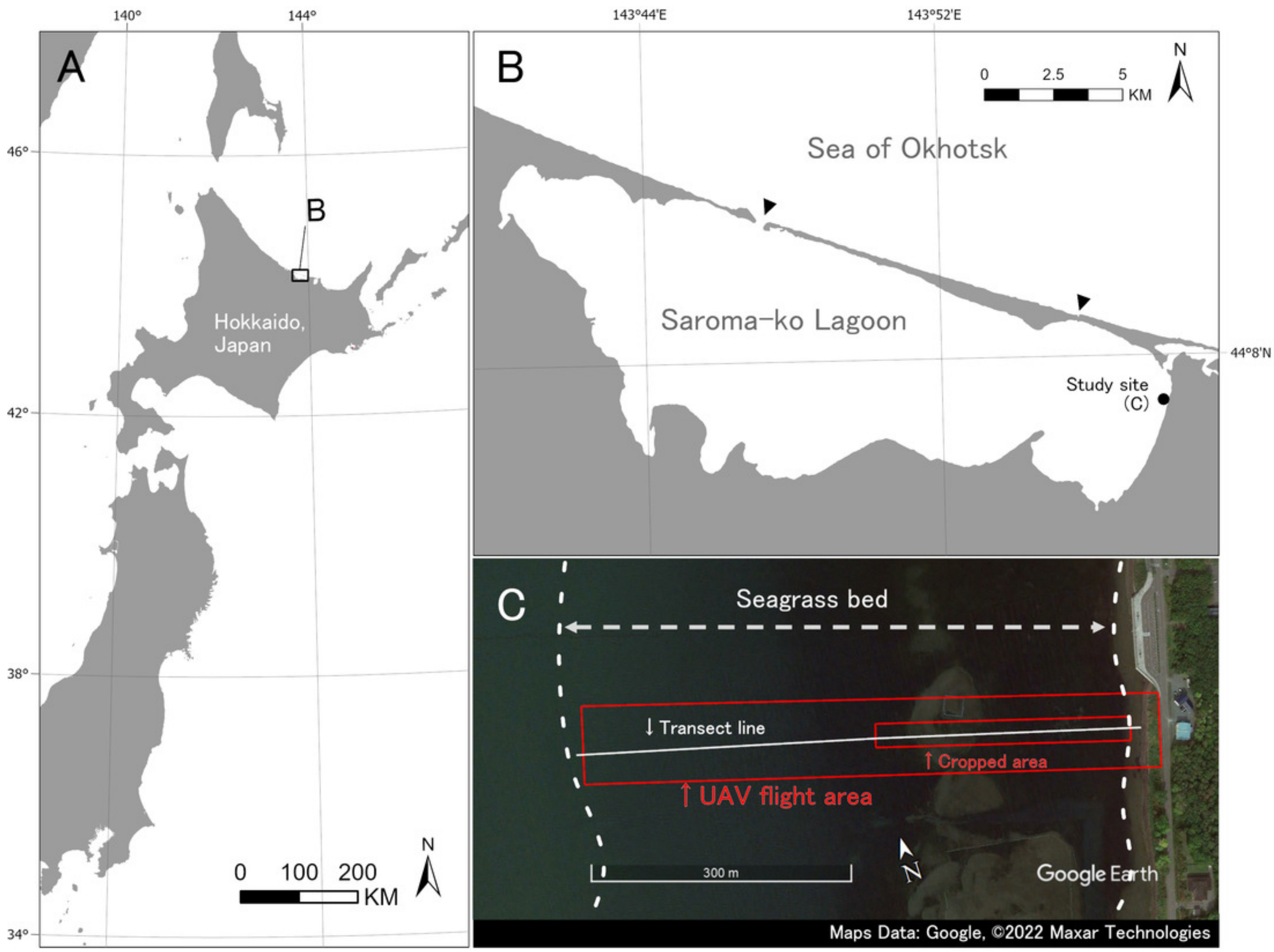

Figure 3

Example of training data which contain only *Zostera marina* (ZM)

Pixels of ground-truth area assigned to ZM (left) is re-assigned to three subclasses (right; ZM1, ZM2 and ZM3) by posterization.

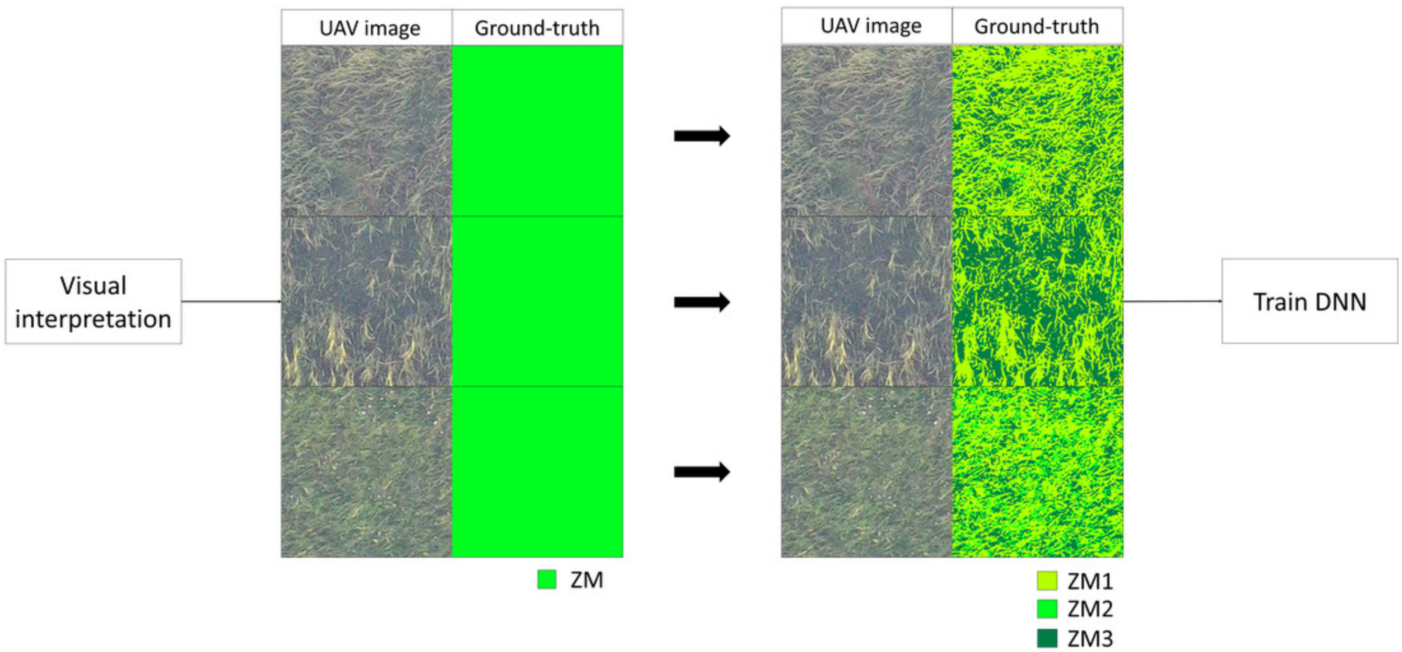

Figure 4

Comparisons of the orthomosaic image (A) and the ground-truth image (B)

The ground-truth image was produced by visual interpretation. A white solid line is a boundary between training area (upper) and validation area (lower) for the different mapping methods. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

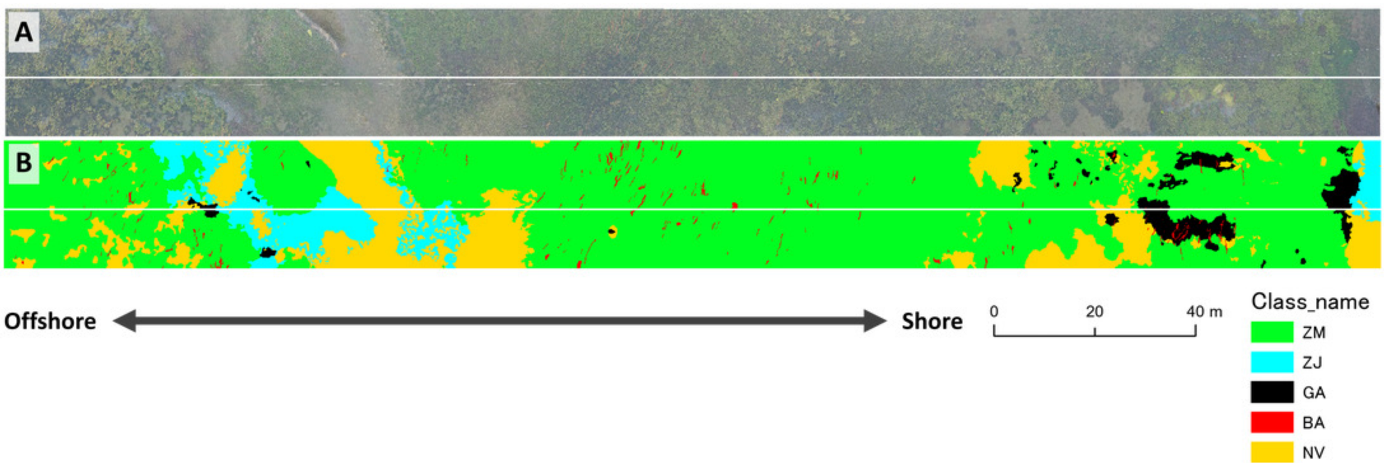

Figure 5

Result of mapping by the three different methods

Maps were produced from the validation area of orthoimage by pixel-based (A), object-based (B) and DNN (C) methods. The ground-truth data is shown in (D). Salt-and-pepper phenomena (speckles noise) was found by the pixel-based classification. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

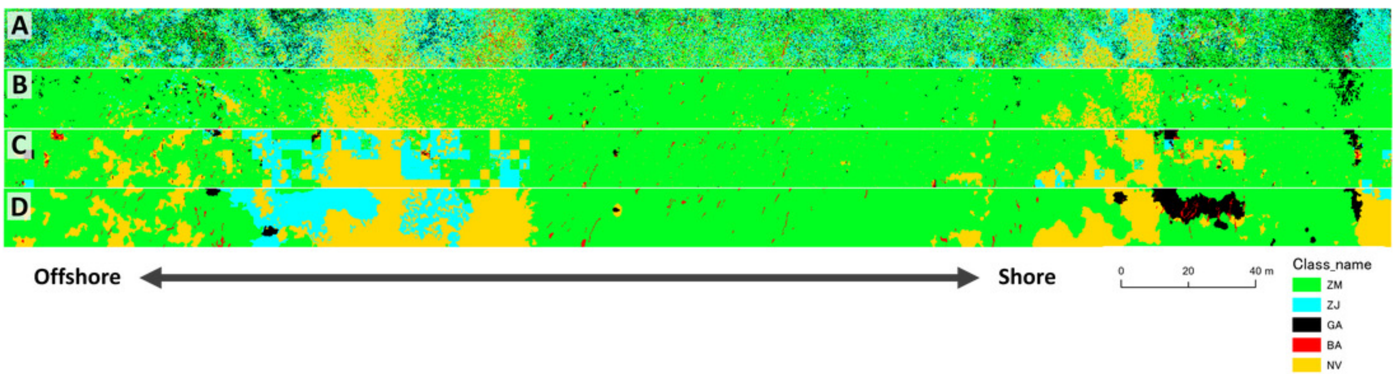

Figure 6

Comparison of GA (green algae) which was not mapped correctly at the validation area

UAV orthoimage (A), pixel-based classification (B), object-based classification (C), DNN (D) and the ground-truth (E). In the DNN map, the most GA was misclassified as NV. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

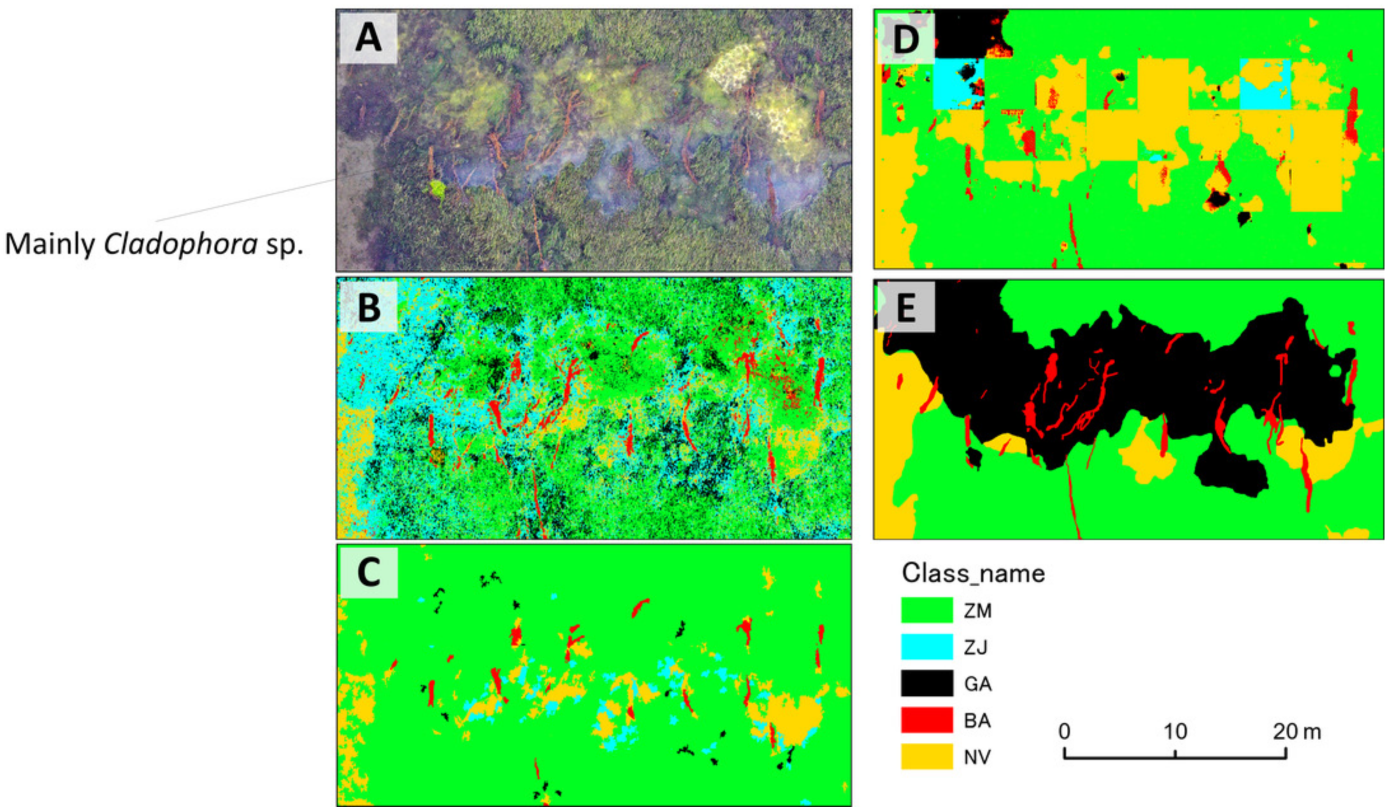

Table 1(on next page)

Confusion matrix evaluating accuracy of visual interpretation based on field data

UA: User Accuracy, PA: Producer Accuracy, OA: Overall Accuracy, K: Kappa coefficient. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

Table 1:

Confusion matrix evaluating accuracy of visual interpretation based on field data.

UA: User Accuracy, PA: Producer Accuracy, OA: Overall Accuracy, K: Kappa coefficient. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

		Field data							UA
		ZM	ZJ	GA	BA	RA	SD	total	
Visual interpretation	ZM	18	2	1	0	2	1	24	0.750
	ZJ	0	7	0	0	0	0	7	1.000
	GA	0	0	3	0	0	0	3	1.000
	BA	0	0	0	1	0	0	1	1.000
	RA	0	0	0	0	0	0	0	NA
	SD	0	2	0	0	1	4	7	0.571
	total	18	11	4	1	3	5	42	
PA		1.000	0.636	0.750	1.000	0.000	0.800		
OA		0.786							
K		0.687							

Table 2 (on next page)

Confusion matrices evaluating accuracy of different mapping methods (A: pixel-based, B: object-based, C: DNN) based on the ground-truth data.

UA: User Accuracy, PA: Producer Accuracy, OA: Overall Accuracy, K: Kappa coefficient. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation

Table 2. Confusion matrices evaluating accuracy of different mapping methods (A: pixel-based, B: object-based, C: DNN) based on the ground-truth data. UA: User Accuracy, PA: Producer Accuracy, OA: Overall Accuracy, K: Kappa coefficient. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

A

		Ground-truth						UA
		ZM	ZJ	GA	BA	NV	Total	
Map	ZM	1364	83	43	3	253	1746	0.781
	ZJ	912	178	43	5	280	1418	0.126
	GA	550	79	36	2	32	699	0.052
	BA	33	2	1	20	39	95	0.211
	NV	380	100	23	2	537	1042	0.515
	Total	3239	442	146	32	1141	5000	
PA		0.421	0.403	0.247	0.625	0.471		
OA		0.427						
K		0.178						

B

		Ground-truth						UA
		ZM	ZJ	GA	BA	NV	Total	
Map	ZM	3137	387	124	14	691	4353	0.721
	ZJ	16	9	3	1	23	52	0.173
	GA	37	5	12	0	2	56	0.214
	BA	3	0	0	11	1	15	0.733
	NV	46	41	7	6	424	524	0.809
	Total	3239	442	146	32	1141	5000	
PA		0.969	0.020	0.082	0.344	0.372		
OA		0.719						
K		0.315						

C

		Ground-truth						UA
		ZM	ZJ	GA	BA	NV	Total	
Map	ZM	3160	145	45	9	312	3671	0.861
	ZJ	11	225	6	0	149	391	0.575
	GA	4	6	17	0	6	33	0.515
	BA	6	0	10	16	3	35	0.457
	NV	58	66	68	7	671	870	0.771
	Total	3239	442	146	32	1141	5000	
PA		0.976	0.509	0.116	0.500	0.588		
OA		0.818						
K		0.618						