Species level mapping of a seagrass bed using an unmanned aerial vehicle and deep learning technique

Satoru Tahara [1], **Kenji Sudo** [2, 3], **Takehisa Yamakita** [4], **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] Japan Fisheries Research and Education Agency, Fisheries Technology Institute, Hatsukaichi, Hiroshima, Japan

[4] Marine Biodiversity and Environmental Assessment Research Center (BioEnv), Research Institute for Global Change (RIGC), Japan Agency for Marine Earth Science and Technology, Yokosuka, Kanagawa, Japan

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

Background. Seagrass beds are essential habitats in coastal ecosystems, providing valuable ecosystem services, but are threatened by various climate change and human activities. Seagrass monitoring by remote sensing have been conducted over past decades using satellite and aerial images, which have low resolution to analyze changes in the composition of different seagrass species in the meadows. 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 the 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.

Methods. We set five benthic 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).

Results. 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.

Conclusion. 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.

1 Species level mapping of a seagrass bed using an unmanned aerial vehicle
2 and deep learning technique
3

5 Satoru Tahara[1], Kenji Sudo[2,3], Takehisa Yamakita[4], Masahiro Nakaoka[2]
6
7 [1]Graduate School of Environmental Science, Hokkaido University, Sapporo, Hokkaido, Japan
8 [2]Akkeshi Marine Station, Field Science Center for Northern Biosphere, Hokkaido University,
9 Akkeshi, Hokkaido, Japan
10 [3]Fisheries Technology Institute, Japan Fisheries Research and Education Agency, Hatsukaichi,
11 Hiroshima, Japan
12 [4]Marine Biodiversity and Environmental Assessment Research Center (BioEnv), Research
13 Institute for Global Change (RIGC), Japan Agency for Marine Earth Science and Technology,
14 Yokosuka, Kanagawa, Japan
15
16 Corresponding Author:
17 Masahiro Nakaoka[2]
18 Akkeshi Marine Station, Field Science Center for Northern Biosphere, Hokkaido University,
19 Akkeshi, Hokkaido, Japan
20 Email address: nakaoka@fsc.hokudai.ac.jp

22
23
24

Abstract

Background. Seagrass beds are essential habitats in coastal ecosystems, providing valuable ecosystem services, but are threatened by various climate change and human activities. Seagrass monitoring by remote sensing have been conducted over past decades using satellite and aerial images, which have low resolution to analyze changes in the composition of different seagrass species in the meadows. 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 the 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.

Methods. We set five benthic 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).

Results. 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.

Conclusion. 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 beds composed of single or multiple species. While seagrass beds play an essential role in providing valuable ecosystem services, they are declining in many parts of the world due to natural and human-induced disturbances (Short & Wyllie-Eciieverria, 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 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) and 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

66   with less effort than field surveys. In addition, it is possible to analyze long-term temporal
67   changes by using aerial photographs (Yamakita, Watanabe & Nakaoka, 2011). While many
68   results have also been reported using satellite data for long-term monitoring (Lyons, Phinn &
69   Roelfsema, 2012; Calleja et al., 2017; Zoffoli et al., 2020; Xu et al., 2021), several limitations
70   have been pointed out for traditional optical remote sensing. The biggest problem is the
71   resolution. The most commonly used satellite data, the Landsat series, provides data over a wide
72   area at a low cost but has a spatial resolution of 30 m which is too low compared to detailed fine-
73   scale information obtained by in-situ field surveys. Phinn et al. (2008) has reported that higher
74   spatial and spectral resolutions are needed for more accurate detailed mapping. Studies using
75   commercial high-resolution satellite images such as WorldView-2 and RapidEye have reported
76   high mapping accuracy (Coffer et al., 2020). However, these commercial satellite images are too
77   expensive for long-term, broad-scale monitoring.
78
79   In recent years, UAVs (Unmanned Aerial Vehicles, or drones) have been increasingly used in
80   field research due to some advantages compared with conventional remote sensing (Nowak et
81   al., 2019). High spatial resolution data are available by low altitude UAV flights. Frequent flight
82   is possible because the no-cloud sky is unnecessary, like satellite, and operation cost is low. It is
83   also possible to adjust the survey time and day, which is impossible with satellites in a fixed
84   orbit. In seagrass research, UAVs have been used for detailed bed mapping (Duffy et al., 2018;
85   Nahirnick et al., 2019; Hobley et al., 2021). Nonetheless, most of these studies mapped seagrass
86   beds consisting of only a single species or conducted mapping without species discrimination.
87
88   Seagrasses have different morphologies and life histories depending on the species (Duarte,
89   1991), and when they live nearby, mapping them by species is necessary to obtain more accurate
90   information such as estimating biomass (Knudby & Nordlund, 2011). Different species provide
91   different ecosystem services (Mtwana Nordlund et al., 2016) and respond differently to changes
92   in the environment (Roca et al., 2016). Thus, developments of detailed methods that can
93   discriminate different seagrass species are promising not only for more effective monitoring of
94   seagrass beds, but also for monitoring and managing invasive species (Kumar et al., 2019).
95
96   Few studies performed species discrimination of seagrasses with UAV images. Román et al.
97   (2021) showed that seagrass bed mapping, including seagrass discrimination, can be performed
98   with high accuracy using a UAV-mounted ten bands multispectral camera and automatic
99   classification based on machine learning algorithms. Chayhard et al. (2018) showed that visual
100   interpretation could be applied to classify seagrass species with different morphology, such as
101   long leaves type (*Enhalus acoroides*) and short leaves type (*Halodule pinifolia* and *H. uninervis*),
102   even using the RGB images taken by UAVs. The camera installed in the consumer-grade UAV is
103   an RGB sensor, and the use of a multispectral camera is costly. Therefore, there is a need for
104   developing methods for seagrass species discrimination using image data with limited spectral
105   resolution but high spatial resolution.

106
107 In general, spatial distribution mapping of seagrass beds by optical remote sensing is carried out
108 using classification algorithms (Diesing et al., 2016). Classification algorithms classify the image
109 into several classes such as seagrass, bare sand, and macroalgae by computer. They can be
110 divided into supervised classification and unsupervised classification depending on whether
111 training data are used or not. In supervised classification, which uses ground-truth data obtained
112 from field surveys as training data, there are two types of classification: (1) pixel-based
113 classification which classifies each pixel, and (2) object-based classification which classifies
114 each object by grouping similarly colored neighboring pixels. It has been reported that object-
115 based classification provides higher accuracy for high spatial resolution images than pixel-based
116 classification (Gao & Mas, 2008). These classification methods have been used to analyze
117 optical remote sensing data based only on limited image information such as the color, object
118 shape, and size. On the other hand, in ultra-high-resolution UAV images, more features are
119 available, such as the pattern, texture, and location of the objects in the image. A deep neural
120 network (DNN) can automatically extract these various features using a convolutional neural
121 network (CNN), the basic network used for DNN image processing (Traore et al., 2018).
122 Seagrass mappings using UAV images and DNN have been conducted in recent years, but they
123 are limited to single species or discriminating seagrass from macroalgae (Hobley et al., 2021,
124 Jeon et al., 2021). There has been a report of successful species classification when used with
125 underwater images (Noman.,2021), but no reports with aerial images. To our knowledge, it is a
126 first study using DNN based methods for species level seagrass mapping by remote sensing. The
127 image-to-image translation, used in this study, is one of the applications of DNN. This model is
128 trained with supervised data for transforming the input image into a corresponding output image
129 using the extracted features (Isola et al., 2017). It can be used for semantic segmentation of input
130 images and has also been applied to seagrass bed mapping by remote sensing (Yamakita et al.,
131 2019).
132
133 This study aimed to use UAV images and image analysis techniques to create a detailed
134 multispecific seagrass map. The study site was set in a seagrass bed of Saroma-ko Lagoon in
135 northeastern Japan where several seagrass and seaweed species are mixed. We got RGB images
136 by consumer-grade UAV and created a benthic map including the following plant taxa; (1)
137 eelgrass *Zostera marina*, (2) dwarf eelgrass *Z. japonica*, (3) green algae (*Chaetomorpha crassa*,
138 *Cladophora* sp.), and (4) brown algae (*Cystoseira hakodatensis*). The accuracies of mapping
139 were compared among three methods, (1) conventional pixel-based supervised classification, (2)
140 object-based supervised classification, and (3) image-to-image translation based on DNN
141 method.
142
143 **Materials & Methods**

144   In this study, we first undertook UAV photography and transect surveys in the field to create
145   reference data, then conducted image analysis in the laboratory. The overall workflow is shown
146   in Fig. 1.
147
148   Fieldwork
149   Fieldwork was carried out on July 9, 2019 at Saroma-ko Lagoon in eastern Hokkaido, Japan
150   (Fig. 2). Saroma-ko Lagoon is a brackish lagoon of about 152 km^2 and is connected to the Sea of
151   Okhotsk by two channels, one about 300 m in width and another 50 m. The maximum depth of
152   the lagoon is 19.6 m. Three species of seagrasses (*Zostera japonica*, *Z. marina*, and *Z.*
153   *caespitosa*) occur along the intertidal and shallow subtidal zones of the lagoon (Biodiversity
154   Center of Japan, 2008). This study was conducted in a seagrass bed at the eastern coast of the
155   lagoon (Fig. 2).
156
157   The transect survey and UAV photography were conducted during a low tide. In the transect
158   survey, a transect line was set perpendicular to the shoreline from the shallowest end in the east
159   to the deepest part of the bed in the west until no seagrass appeared (about 600 m offshore). A
160   total of 86 quadrats of 0.25 m^2 were placed haphazardly along the transect to cover all present
161   seagrasses and macroalgae along the transect, and species and cover were recorded. Surveys
162   were conducted by wading, snorkeling, and SCUBA diving.
163
164   UAV photography was conducted from shore using a quadcopter Mavic2 pro (DJI Co. Ltd). The
165   flight area was set at 580 m offshore and 90 m wide, including a measuring tape used for the
166   transect. We took the images with the RGB sensor camera equipped with the Mavic2 pro at a
167   nadir angle. The flight was automated using DroneDeploy (DroneDeploy Co. Ltd.).
168   DroneDeploy enables automatic flight and photography by specifying the flight area, altitude,
169   and overlap rate (front and side) between images. To ensure sufficient spatial resolution for
170   seagrass species identification and to enable orthorectification, we used the setting for
171   DroneDeploy as follows: altitude 30 m, front overlap 80 % and side overlap 70 %. The camera
172   settings were set before the shooting and were not changed (aperture: f/2.8, shutter speed: 1/400
173   s, ISO: 200, and white balance: Auto).
174
175   Image pre-processing
176   The captured UAV images were orthorectified using the SfM-MVS processing through the
177   software Metashape ver. 1.7.1 (Agisoft Co. Ltd.). Through SfM-MVS processing, we can
178   produce an orthoimage from overlapped images (Verhoeven et al., 2013). Then, the images were
179   cropped for subsequent analyses. The orthoimage was first converted to a benthic cover map by
180   visual interpretation. Interpreters were the authors who conducted the transect survey and
181   confirmed that vegetation in the study area is uniform along with water depth. Thus, we could
182   distinguish each mapping class by visual inspection without putting quadrats, and applied this
183   method to prepare enough reference data outside of transect survey area for supervised mapping

184 methods. This method was used in a previous study where seagrass areas were recognized based
185 on images (Yamakita et al., 2019).

186

187 As a result of the transect survey, three species of seagrass (*Z. marina*, *Z. japonica*, *Z.*
188 *caespitosa*), green algae (*Chaetomorpha crassa*, *Cladophora* sp.), brown algae (*Cystoseira*
189 *hakodatensis*), and red algae (Ceramiaceae gen spp.) were observed. Three seagrass species were
190 continuously mixed and the dominant species changed with water depth; *Z. japonica* (intertidal),
191 *Z. marina* (shallower subtidal), and *Z. caespitosa* (deeper subtidal). *Zostera caespitosa* was
192 difficult to distinguish from *Z. marina* without observing the belowground part, so the area
193 offshore of 300 m from the shoreline where *Z. caespitosa* occurred was cropped and excluded
194 from subsequent analysis of orthoimage. This cropping resulted in a total area of 7,884 m^2, 291
195 m along the depth axis and 27 m horizontally to the depth axis. As for macroalgae, red algae
196 were found only in a limited area and were not distinguishable from other vegetation by the
197 naked eye, so they were excluded from the classification. Green algae were combined into one
198 class because it was difficult to distinguish the two species.

199

200 These resulted in five benthic classes in this study (*Z. marina* (ZM), *Z. japonica* (ZJ), green algae
201 (GA), brown algae (BA), and no vegetation (NV)). RGB values for each class had non-normal,
202 wide-range and overlapped distribution (Fig. 3). Based on these, especially for wide red and
203 green values of ZM and the red value of BA, it was predicted that classification based on color
204 information would be inaccurate. In fact, during visual interpretation, identification by color
205 alone was difficult. However, the patterns of the objects comprising the classes were visible as
206 texture. The interpreters used this information and hand-traced the boundaries of each class on an
207 image editing software, Paint. NET ver.4.2.16 (dotPDN LLC.). The inside of the boundaries was
208 painted and mapped according to class. The orthoimages showing each class are shown in Fig. 4.
209 ZM, ZJ and GA were similarly green, but brightness and visual texture differed. The brightness
210 of ZM ranged widely depending on whether they were submerged or not. We could confirm the
211 stripe pattern of ZM leaves on the orthoimage. Because ZJ leaves were smaller and sparser than
212 ZM, the texture of ZJ was smooth, and color was mixed with the ground under vegetation. The
213 texture of GA was also smooth, but GA was dense and covered with the ground completely. GA
214 was red and string-shaped, thus easily distinguishable. NV was grey and had a white dot pattern
215 of shells.

216

217 This study used the maps created by visual interpretation as ground-truth images for training and
218 accuracy verification data. To examine the credibility of the visual interpretation, we compared
219 the ground-truth images with the data obtained from the transect survey. For the comparison, the
220 location of each quadrat was first identified on the orthoimage based on the measurement tape
221 used for the transect installation, and the dominant vegetation classes (ZM, ZJ, GA, BA) were
222 examined. Next, the area corresponding to the quadrat area was cropped from the ground-truth
223 image. The dominant taxonomic classes were examined in the same way and compared with the

224 results of the transect survey. In all cases, however, if the coverage of the dominant class was
225 less than 10%, the no vegetation class (ND) was considered the dominant class.
226
227 Mapping method comparison
228 Mapping by visual interpretation is highly accurate, but requires extensive labor. This study
229 compared three mapping methods (pixel-, object-based classification and image-to-image
230 translation based on DNN) to find a more efficient and reproducible method. All of them are
231 supervised methods, which means that by training the computer using training data, mapping can
232 be done automatically for the rest of the data. In this study, we trained each method using the
233 ground-truth image by visual interpretation. About half of the orthoimage (54%) was used as a
234 training area and the rest (46%) as a validation area, from which accuracy assessment was
235 conducted for each method.
236
237 a). Conventional mapping (Pixel-based and object-based classifications)
238 Pixel-based and object-based classifications are standard methods for remote sensing images
239 (Dat Pham et al., 2019). These are supervised classifications in which data in some areas are
240 used as training data to classify data in other areas. In this study, the training data were created
241 on ArcGIS pro ver. 2.8.1 (Esri Co. Ltd.). Pixel-based classification classifies each pixel, while
242 object-based classification classifies each object. An object is a collection of similarly colored
243 neighboring pixels created by the segmentation of the input image. For segmentation, three
244 parameters were adjusted until the object became an appropriate size (Spectral detail: 20, Spatial
245 detail: 5, Minimum segment size: 500 pixels). Spectral detail and spatial detail range from 1 to
246 20, with higher values used for finer classifications.
247
248 The classification parameters were RGB value of pixels for the pixel-based classification, and
249 color and shape of objects for the object-based classification. Parameters describing color and
250 shape of objects are converged color, mean digital number, standard deviation, count of pixels,
251 compactness and rectangularity. These six parameters are optionally available for classification
252 on ArcGIS pro Spatial Analysis Tool.
253
254 In this study, the algorithm used for classification was support vector machine (SVM), which
255 was used in seagrass mapping and reported to be sufficiently accurate (Pottier et al., 2021). SVM
256 is not sensitive to training data size and does not assume the probability distribution of the data
257 (Mountrakis, Im & Ogole, 2011). The training data were polygons created from a ground-truth
258 image by uniformly selecting a representative area of each specific class. Pixel-based and object-
259 based classifications were fed different training polygons for higher classification accuracy. The
260 area (number) of training data for each class (ZM, ZJ, GA, BA, and NV) was 75.6 m^2 (9), 12.7
261 m^2 (10), 4.04 m^2 (5), 0.442 m^2 (7) and 21.6 m^2 (12) for the pixel-based classification and 75.6 m^2
262 (9), 12.8 m^2 (11), 4.31 m^2 (7), 0.669 m^2 (9) and 15.5 m^2 (14) for the object-based classification.
263
264 b). Image translation based on deep learning (pix2pix)

265   Pix2pix is an image-to-image translation model based on conditional generative adversarial
266   networks (cGANs) (Isola et al., 2017). cGANs are the application of CNN and have two
267   networks: generator and discriminator. The generator transforms the input image, and the
268   discriminator classifies translated image as fake or real by comparing it with the ground-truth
269   image. The generator and discriminator compete with each other, and the generator comes to
270   transform the image into a more realistic one. This model can also be used for remote sensing
271   mapping by translating images to classified images and showed higher accuracy than other deep
272   learning models (Isola et al., 2017). Pix2pix has been applied to various examples, including
273   seagrass mapping for black-and-white aerial photography (Yamakita et al., 2019).
274
275   The translation process in pix2pix requires the size of the input image to be 256 x 256 pixels.
276   Therefore, the orthoimage was sliced to an appropriate size without overlap beforehand. Each
277   slice can include one or more classes. After slicing the orthoimage, the numbers of training and
278   validation data were 980 and 840. In general, DNNs are trained more robustly with increasing
279   training data. To prepare a sufficient amount of training data, data augmentation has been carried
280   out in previous studies (Mikolajczyk & Grochowski, 2018, Yamato et al, 2021). Therefore, we
281   added flipped copies to augment the training data. Horizontal, vertical, and simultaneous
282   horizontal and vertical flipped copies of the original training data were added. With this
283   procedure, the number of training data (256 x 256 pixels) increased from 980 to 3920.
284
285   GANs-based networks often suffer from a problem called mode collapse (Goodfellow, 2016).
286   This occurs when the training data contain a lot of similar ground-truth images. In such cases, the
287   translated image by the network would also result in similar images. In the study area, the
288   percentage of the ZM area is high, and most of ground-truth data of the training data are
289   dominated by ZM only, which can cause mode collapse. We divided the training data ZM into
290   three subclasses to solve this problem. We reduced colors in the orthoimage of the ZM area to
291   three by posterization and assigned a subclass to each of them. This prevented the
292   homogenization of the ground truth image (Fig. 5).
293
294   Accuracy assessment
295   Accuracy assessment was performed by comparing the mapping results of each method in the
296   validation area with the ground truth data. Five thousand random points were extracted in the
297   validation area, and a confusion matrix was created for each resulting map. The confusion matrix
298   was used to calculate the overall accuracy (OA) and Kappa coefficient (K) for all classes and the
299   user accuracy (UA) and producer accuracy (PA) for individual classes. OA represents the ratio of
300   the pixel classified correctly. K is a statistic value that expresses the degree of agreement
301   between data, taking into account coincidence (Cohen, 1960). K = 0 means that the degree of
302   agreement is equal to that obtained by chance, and positive values indicate a degree of agreement
303   greater than chance, with the maximum value of 1. In general, the relationship between K and
304   strength of agreement is < 0.00: poor, 0.00-0.20: slight, 0.21-0.40: fair, 0.41-0.60: moderate, and
305   0.61-0.80: substantial (Landis & Koch, 1977). UA is the ratio of each class assigned by the

306   correctly classified mapping, and PA is the ratio of each class assigned by ground truth that is
307   correctly classified.

308

309   Results

310   Image pre-processing

311   The flight time of the UAV photography was 22 minutes, and 406 out of 534 taken images were
312   used to orthorectification. The remaining 128 images were taken in deep water where the
313   seagrass was submerged entirely, and they did not show any features such as seagrass leaves,
314   rocks, artificial objects. Due to this, the computer couldn't detect any matching points across the
315   images which were necessary for creating orthoimage.

316

317   The spatial resolution of the created orthoimage was 8.13 mm/pix (Fig. 6 A). After cropping
318   orthoimage for analysis, the number of quadrats for the transect survey included in the image
319   was 42. Among them, those dominated by *Z. marina*, *Z. japonica*, green algae, brown algae, red
320   algae and no vegetation were 18, 11, 4, 1, 3, and 5, respectively. In the ground-truth image, a part
321   of *Z. japonica* (4/11) and green algae (1/4) were misclassified as adjacent *Z. marina* or no
322   vegetation. Similarly, the red algae that did not appear in the ground-truth images were classified
323   as *Z. marina* or no vegetation (Table 1). Overall accuracy and Kappa of visual interpretation data
324   (Fig. 6 B) validated by the field data were 0.786 and 0.687, respectively.

325

326   Mapping and accuracy assessment

327   The results of mapping generally agreed among the three different methods (Fig. 7). The study
328   site was overall dominated by *Z. marina* (ZM). *Z. japonica* (ZJ) occurred mostly at the western
329   part with the large gap (no vegetation, or NV). Gaps are also observed in the eastern part of the
330   site. Green algae (GA) mostly occurred at the eastern part, whereas brown algae (BA) only in
331   small patches scattered along the whole site. Some misidentifications were observed in these
332   classification methods (Fig. 7). For example, small gaps (no vegetation) in deeper parts were not
333   identified by the pixel-based and object-based methods. In contrast, GA in the shallower parts
334   were overestimated by the pixel-based method. When compared visually, the DNN showed the
335   closest result to the ground-truth. Salt-and-pepper phenomena (speckles noise) was found by the
336   pixel-based classification and made the map difficult to read. Most ZJ in object-based
337   classification were undetected and misclassified into ZM. The DNN also misclassified a giant
338   GA patch at shallow and some NV areas into other classes.

339

340   Accuracy assessment showed that the values of OA and K were highest for DNN, followed by
341   the object-based, and the pixel-based methods (Table 2). K value for DNN exceeded 0.6,
342   indicating substantial agreement, whereas that for the pixel-based methods was less than 0.2,
343   showing poor fit. Pixel-based method showed lowest accuracy, because speckles are observed
344   overall in the result map (Fig. 7 A).

345

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

357

358 **Discussion**

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

363

364 Previous studies have attempted to discriminate species of seagrass and macroalgae using
365 satellite and aerial images (e.g., Phinn et al., (2008), Kovacs et al., (2018)). Although
366 comparisons should be made with caution due to the difference in sites and methods, the
367 accuracy in our study (OA: 0.818) outperforms those by other studies (OA: 0.23 and 0.28 for
368 Phinn et al., (2008), and 0.64~0.69 for Kovacs et al., (2018)). The grain size of our seagrass bed
369 map (8.13 mm/pix) was much higher than these previous studies (2.4 and 4 m/pix for Phinn et
370 al., (2008), and 2~30 m/pix for Kovacs et al., (2018)), indicating that the spatial resolution was a
371 key factor for successful classification of different plants in multispecific seagrass meadows. In
372 this study, however, spatial extent was small (7,884 m^2), covering only 0.035 % of the seagrass
373 beds in Saroma-ko Lagoon (22.5 km^2 in 2015, Hokkaido Aquaculture Promotion Cooperation
374 2015). Linear extrapolation indicates that it would take us more than 1,000 hours (i.e. > 40 days)
375 to cover the whole seagrass bed by this method, which is not practical considering the labor and
376 seasonal changes in the seagrass bed.

377

378 To increase the accuracy of seagrass bed mapping, previous studies have used hyperspectral
379 sensors aboard on satellites and aircraft for species discrimination. This is because seagrasses
380 and seaweed species can be discriminated with different spectral reflectance as well as terrestrial
381 plants (Fyfe, 2003). Although light-weight hyperspectral sensors that can be mounted on UAVs
382 have been developed, they have not been widely used yet because they require complex pre- and
383 post-flight operations for analysis (Adão et al., 2017). Furthermore, it may be difficult even with
384 hyperspectral sensors to discriminate closely related congeneric species of seagrass which have
385 similar characteristics of leaf color. Due to this limitation, it is more effective to develop a new

386 discrimination method using information other than color in the visible band with UAV mounted
387 RGB sensors, which are already used in the field of seagrass research. This study used high-
388 resolution orthoimage for mapping, in which the surface pattern of each class can be identified as
389 well as color. However, mapping based on the conventional classification methods using RGB
390 showed very low overall accuracy, and especially for discriminating ZJ which were misclassified
391 to ZM and NV. This suggests that the pixel and segment colors and shape information used as
392 classification parameters were overlapped among the classes. On the other hand, the mapping
393 based on DNN showed higher accuracy than the conventional methods. This highlights the
394 advantage of the DNN method, in which the computer can extract and use much more
395 information than just color and shape information from the UAV images (Albawi, Mohammed &
396 Al-Zawi, 2017). In fact, ZM and ZJ differ in the presence or absence of a stripe pattern, and the
397 DNN might have identified them based on this information. Therefore, image data with limited
398 spectral information can be analyzed in a more sophisticated way by applying DNN.
399

400 In contrast to discrimination of different *Zostera* species, results of green algae (GA)
401 classification were in low accuracy in all methods. When comparing GA visually in the
402 orthoimage of the training and validation areas, the giant patch of green algae in the validation
403 area has a bright green color that is not seen in the training area. This is due to the difference in
404 the species composition of the GA, which is mainly *Cladophora* sp. in this brighter clump, and
405 mainly darker *Chaetomorpha crassa* in the rest. Since these green algae were mixed in the patch,
406 it was not easy to separate them into different classes by visual interpretation. The training data
407 did not sufficiently cover the variability of GA, which may be the reason for the low accuracy.
408 This indicates that even when we use DNNs, there is a limit to their versatility, and performance
409 varies by different seagrass and seaweed species. Higher generalizability will be possible by
410 increasing the variety of training data that sufficiently cover the variability in the study area. BA,
411 which had the smallest area in the training data, also had the lowest UA among the classes in the
412 DNN method. The relatively high UA for the object-based classification suggests that the
413 amount of data is insufficient for training the DNN rather than spectral overlap. Therefore, a
414 possible method to improve accuracy is to provide more training data as for GA.
415

416 It has been reported in previous studies that the salt-pepper phenomenon, defined as individual
417 pixels classified differently from their neighbors (Yu et al., 2006), reduces the classification
418 accuracy of pixel-based classification for high-resolution images (Feng, Liu & Gong, 2015). The
419 salt-pepper phenomenon is caused by the internal variability within a classification class that
420 appears as noise in the classification results. In this study, the salt-pepper phenomenon was also
421 observed, and it is one of the factors causing the low accuracy of pixel-based classification . On
422 the other hand, the results of object-based classification showed that segmentation suppressed the
423 salt-pepper phenomenon, making it a more suitable method for high-resolution images. Result of
424 DNN also showed no sand-pepper phenomenon.
425

426 In this study, we used a ground-truth image produced by visual interpretation for training data to
427 secure the amount of training data. The machine learning algorithm used in this study, SVM, is
428 known to be more sensitive to the quality of training data than its size (Mountrakis et al., 2011).
429 Therefore, the training data is prepared to represent each class in the training area, and we do not
430 need the entire ground-truth image. On the other hand, the importance of the size of training data
431 is confirmed for DNN by the fact that data augmentation is common to improve algorithms in
432 previous studies (Mikolajczyk & Grochowski, 2018). Therefore, the DNN method is inferior in
433 terms of the time and effort required to prepare the training data. In this study, it took only a few
434 hours to prepare the training data for pixel-based and object-based classification, but it took
435 several days for the DNN method. In terms of computational effort, it took about one night for
436 training DNN using a computer equip high-performance GPU, but several minutes for SVM. In
437 addition, the ground-truth images created by visual interpretation contained some errors when
438 validated by the ground-truth data. Currently, there are examples of underwater photo datasets
439 available for seagrass detection (Reus et al., 2018), but there are no available datasets with
440 labeled aerial seagrass images. Therefore, researchers applying DNNs will need to start by
441 creating a dataset by themselves. However, it is still worth considering the application of DNN
442 because it is expected to achieve highly accurate mapping. Acquiring and creating ground-truth
443 data with quality and quantity are future challenges. The training data proportions are biased
444 toward ZM in this study, thus we would have to be cautious in applying trained network to sites
445 with different vegetation proportions. Investigating the generality of the method is also a future
446 work.
447

Conclusions

449 This study reports the result of a case study which applied UAV and machine learning
450 techniques including deep learning at multispecific seagrass bed. UAV enables easier acquisition
451 of high spatial resolution data that was previously difficult to obtain by other remote sensing
452 devices. Pixel-based classification is not suitable for mapping due to the salt-pepper
453 phenomenon. Image-to-image translation based on DNN could discriminate seagrass species and
454 macroalgae, and show higher accuracy than conventional classification methods. Our result
455 indicates that DNN is especially useful when we can obtain high-resolution images with
456 conventional cameras with limited spectral range. Some challenges remain, such as limitation in
457 covering wide areas for the mapping, and in labors for preparing ground truth data. Nevertheless,
458 UAV detailed mapping at coastal area enables scientists further biological research of submerged
459 vegetation based on spatial information.
460
461

Acknowledgements

463 We thank Mizuho Namba and Minako Abe Ito for their help with field surveys, and K.
464 Sakaguchi and other members of the Aquaculture Fishery Cooperative of Saroma Lake for their

465 logistical support during our field work in Saroma. We highly appreciate Minako Abe Ito for
466 improving the manuscript.
467

Funding

469 Funding was provided by Saroma-ko Aquaculture Cooperation, Japan Society for the Promotion
470 of Science (Kakenhi 16H01792) and Environmental Restoration and Conservation Agency
471 (Environment Research and Technology Development Fund, S-15 Predicting and Assessing
472 Natural Capital and Ecosystem Services (PANCES). The funders had no role in study design,
473 data collection and analysis, decision to publish, or preparation of the manuscript.
474

Competing Interests

476 The authors declare there are no competing interests.
477

Author Contributions

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

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

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

485 · Masahiro Nakaoka conceived and designed the research, conducted the fieldwork, authored or
486 reviewed drafts of the paper, and approved the final draft.
487

References

489 Adão T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, Sousa J. 2017. Hyperspectral Imaging:
490   A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and
491   Forestry. Remote Sensing 9:1110. DOI: 10.3390/rs9111110.
492 Albawi S, Mohammed TA, Al-Zawi S. 2017. Understanding of a convolutional neural network.
493   In: 2017 International Conference on Engineering and Technology (ICET). IEEE, 1–6.
494   DOI: 10.1109/ICEngTechnol.2017.8308186.
495 Biodiversity Center of Japan. The Report for the 7th Natural Environmental Survey Shallow
496   Marine Eco- system Survey (Aquatic Vegetation Survey) [Internet]. Yamanashi, Japan;
497   2008. Available: http://www. biodic.go.jp/reports2/6th/6_moba19/6_moba19.pdf
498 Calleja F, Galván C, Silió-Calzada A, Juanes JA, Ondiviela B. 2017. Long-term analysis of
499   Zostera noltei: A retrospective approach for understanding seagrasses' dynamics. Marine
500   Environmental Research 130:93–105. DOI: 10.1016/j.marenvres.2017.07.017.
501 Chayhard S, Manthachitra V, Nualchawee K, Buranapratheprat A. 2018. Application of
502   unmanned aerial vehicle to estimate seagrass biomass in Kung Kraben Bay, Chanthaburi
503   province, Thailand. International Journal of Agricultural Technology 14:1107–1114.

504 Coffer MM, Schaeffer BA, Zimmerman RC, Hill V, Li J, Islam KA, Whitman PJ. 2020.
505   Performance across WorldView-2 and RapidEye for reproducible seagrass mapping.
506   Remote Sensing of Environment 250:112036. DOI: 10.1016/j.rse.2020.112036.
507 Cohen J. 1960. A coefficient of agreement for nominal scales. Educational And Psychological
508   Measurement XX:37–46.
509 Dat Pham T, Xia J, Thang Ha N, Tien Bui D, Nhu Le N, Tekeuchi W. 2019. A review of remote
510   sensing approaches for monitoring blue carbon ecosystems: Mangroves, sea grasses and
511   salt marshes during 2010–2018. Sensors (Switzerland) 19. DOI: 10.3390/s19081933.
512 Diesing M, Mitchell P, Stephens D. 2016. Image-based seabed classification: what can we learn
513   from terrestrial remote sensing? ICES Journal of Marine Science: Journal du Conseil
514   73:2425–2441. DOI: 10.1093/icesjms/fsw118.
515 Duarte CM. 1991. Allometric scaling of seagrass form and productivity. Marine Ecology
516   Progress Series 77:289–300.
517 Duffy JP, Pratt L, Anderson K, Land PE, Shutler JD. 2018. Spatial assessment of intertidal
518   seagrass meadows using optical imaging systems and a lightweight drone. Estuarine,
519   Coastal and Shelf Science 200:169–180. DOI: 10.1016/j.ecss.2017.11.001.
520 Feng Q, Liu J, Gong J. 2015. UAV Remote Sensing for Urban Vegetation Mapping Using
521   Random Forest and Texture Analysis. Remote Sensing 7:1074–1094. DOI:
522   10.3390/rs70101074.
523 Fyfe SK. 2003. Spatial and temporal variation in spectral reflectance: Are seagrass species
524   spectrally distinct? Limnology and Oceanography 48:464–479. DOI:
525   10.4319/lo.2003.48.1_part_2.0464.
526 Gao Y, Mas J. 2008. A comparison of the performance of pixel based and object based
527   classifications over images with various spatial resolutions. Online journal of earth
528   sciences 2:27–35.
529 Goodfellow I. 2016. NIPS 2016 Tutorial: Generative Adversarial Networks.
530 Gumusay MU, Bakirman T, Tuney Kizilkaya I, Aykut NO. 2019. A review of seagrass detection,
531   mapping and monitoring applications using acoustic systems. European Journal of Remote
532   Sensing 52:1–29. DOI: 10.1080/22797254.2018.1544838.
533 Hobley B, Arosio R, French G, Bremner J, Dolphin T, Mackiewicz M. 2021. Semi-Supervised
534   Segmentation for Coastal Monitoring Seagrass Using RPA Imagery. Remote Sensing
535   13:1741. DOI: 10.3390/rs13091741.
536 Hokkaido Aquaculture Promotion Cooperation (2015) Report on seagrass bed distribution in
537   Saroma-ko Lagoon. Hokkaido Aquaculture Promotion Cooperation, 84pp.
538 Isola P, Zhu JY, Zhou T, Efros AA. 2017. Image-to-image translation with conditional
539   adversarial networks. Proceedings - 30th IEEE Conference on Computer Vision and
540   Pattern Recognition, CVPR 2017 2017-Janua:5967–5976. DOI: 10.1109/CVPR.2017.632.
541 Jeon E, Kim S, Park S, Kwak J, Choi I. 2021. Semantic segmentation of seagrass habitat from
542   drone imagery based on deep learning: A comparative study. Ecological Informatics
543   66:101430. DOI: 10.1016/j.ecoinf.2021.101430.

544 Kendrick GA, Hegge BJ, Wyllie A, Davidson A, Lord DA. 2000. Changes in seagrass cover on
545   Success and Parmelia Banks, Western Australia between 1965 and 1995. Estuarine,
546   Coastal and Shelf Science 50:341–353. DOI: 10.1006/ecss.1999.0569.
547 Knudby A, Nordlund L. 2011. Remote sensing of seagrasses in a patchy multi-species
548   environment. International Journal of Remote Sensing 32:2227–2244. DOI:
549   10.1080/01431161003692057.
550 Kovacs E, Roelfsema C, Lyons M, Zhao S, Phinn S. 2018. Seagrass habitat mapping: how do
551   Landsat 8 OLI, Sentinel-2, ZY-3A, and Worldview-3 perform? Remote Sensing Letters
552   9:686–695. DOI: 10.1080/2150704X.2018.1468101.
553 Kumar A, Cooper C, Remillard CM, Ghosh S, Haney A, Braun F, Conner Z, Page B, Boyd K,
554   Wilde S, Mishra DR. 2019. Spatiotemporal monitoring of hydrilla [Hydrilla verticillata (L.
555   f.) Royle] to aid management actions. Weed Technology 33:518–529. DOI:
556   10.1017/wet.2019.13.
557 Landis JR, Koch GG. 1977. The Measurement of Observer Agreement for Categorical Data.
558   Biometrics 33:159. DOI: 10.2307/2529310.
559 Lyons MB, Phinn SR, Roelfsema CM. 2012. Long term land cover and seagrass mapping using
560   Landsat and object-based image analysis from 1972 to 2010 in the coastal environment of
561   South East Queensland, Australia. ISPRS Journal of Photogrammetry and Remote Sensing
562   71:34–46. DOI: 10.1016/j.isprsjprs.2012.05.002.
563 Mikolajczyk A, Grochowski M. 2018. Data augmentation for improving deep learning in image
564   classification problem. In: 2018 International Interdisciplinary PhD Workshop (IIPhDW).
565   IEEE, 117–122. DOI: 10.1109/IIPHDW.2018.8388338.
566 Mountrakis G, Im J, Ogole C. 2011. Support vector machines in remote sensing: A review.
567   ISPRS Journal of Photogrammetry and Remote Sensing 66:247–259. DOI:
568   10.1016/j.isprsjprs.2010.11.001.
569 Mtwana Nordlund L, Koch EW, Barbier EB, Creed JC. 2016. Seagrass Ecosystem Services and
570   Their Variability across Genera and Geographical Regions. PLOS ONE 11:e0163091.
571   DOI: 10.1371/journal.pone.0163091.
572 Nahirnick NK, Reshitnyk L, Campbell M, Hessing‐Lewis M, Costa M, Yakimishyn J, Lee L.
573   2019. Mapping with confidence; delineating seagrass habitats using Unoccupied Aerial
574   Systems (UAS). Remote Sensing in Ecology and Conservation 5:121‐135. DOI:
575   10.1002/rse2.98.
576 Noman MK, Islam SMS, Abu-Khalaf J, Lavery P. 2021. Multi-species Seagrass Detection Using
577   Semi-supervised Learning. In: 2021 36th International Conference on Image and Vision
578   Computing New Zealand (IVCNZ). IEEE, 1–6. DOI:
579   10.1109/IVCNZ54163.2021.9653222.Nowak MM, Dziób K, Bogawski P. 2019.
580   Unmanned Aerial Vehicles (UAVs) in environmental biology: a review. European Journal
581   of Ecology 4:56–74. DOI: 10.2478/eje-2018-0012.
582 Phinn S, Roelfsema C, Dekker A, Brando V, Anstee J. 2008. Mapping seagrass species, cover
583   and biomass in shallow waters: An assessment of satellite multi-spectral and airborne

584   hyper-spectral imaging systems in Moreton Bay (Australia). Remote Sensing of
585   Environment 112:3413–3425. DOI: 10.1016/j.rse.2007.09.017.
586 Pottier A, Catry T, Trégarot E, Maréchal J-P, Fayad V, David G, Sidi Cheikh M, Failler P. 2021.
587   Mapping coastal marine ecosystems of the National Park of Banc d'Arguin (PNBA) in
588   Mauritania using Sentinel-2 imagery. International Journal of Applied Earth Observation
589   and Geoinformation 102:102419. DOI: 10.1016/j.jag.2021.102419.Reus G, Möller T, Jager
590   J, Schultz ST, Kruschel C, Hasenauer J, Wolff V, Fricke-Neuderth K. 2018. Looking for
591   seagrass: Deep learning for visual coverage estimation. 2018 OCEANS - MTS/IEEE Kobe
592   Techno-Oceans, OCEANS - Kobe 2018:2–7. DOI:
593   10.1109/OCEANSKOBE.2018.8559302.
594 Roca G, Alcoverro T, Krause-Jensen D, Balsby TJS, Van Katwijk MM, Marbà N, Santos R,
595   Arthur R, Mascaró O, Fernández-Torquemada Y, Pérez M, Duarte CM, Romero J. 2016.
596   Response of seagrass indicators to shifts in environmental stressors: A global review and
597   management synthesis. Ecological Indicators 63:310–323. DOI:
598   10.1016/j.ecolind.2015.12.007.
599 Román A, Tovar-Sánchez A, Olivé I, Navarro G. 2021. Using a UAV-Mounted Multispectral
600   Camera for the Monitoring of Marine Macrophytes. Frontiers in Marine Science 8:1–12.
601   DOI: 10.3389/fmars.2021.722698.
602 Sherwood ET, Greening HS, Johansson JOR, Kaufman K, Raulerson GE. 2017. Tampa Bay
603   (Florida, USA): Documenting Seagrass Recovery since the 1980's and Reviewing the
604   Benefits. Southeastern Geographer 57:294–319. DOI: 10.1353/sgo.2017.0026.
605 Short F, Carruthers T, Dennison W, Waycott M. 2007. Global seagrass distribution and diversity:
606   A bioregional model. Journal of Experimental Marine Biology and Ecology 350:3–20.
607   DOI: 10.1016/j.jembe.2007.06.012.
608 Short FT, Koch EW, Creed JC, Magalhães KM, Fernandez E, Gaeckle JL. 2006. SeagrassNet
609   monitoring across the Americas: Case studies of seagrass decline. Marine Ecology 27:277–
610   289. DOI: 10.1111/j.1439-0485.2006.00095.x.
611 Short FT, Wyllie-Eciieverria S. 1996. Natural and human-induced disturbance of seagrasses.
612   Environmental Conservation 23:17–27. DOI: 10.1017/s0376892900038212.
613 Sudo K, Quiros TEAL, Prathep A, Luong C Van, Lin H-J, Bujang JS, Ooi JLS, Fortes MD,
614   Zakaria MH, Yaakub SM, Tan YM, Huang X, Nakaoka M. 2021. Distribution, Temporal
615   Change, and Conservation Status of Tropical Seagrass Beds in Southeast Asia: 2000–2020.
616   Frontiers in Marine Science 8:1–11. DOI: 10.3389/fmars.2021.637722.
617 Tomasko DA, Corbett CA, Greening HS, Raulerson GE. 2005. Spatial and temporal variation in
618   seagrass coverage in Southwest Florida: Assessing the relative effects of anthropogenic
619   nutrient load reductions and rainfall in four contiguous estuaries. Marine Pollution Bulletin
620   50:797–805. DOI: 10.1016/j.marpolbul.2005.02.010.
621 Traore BB, Kamsu-Foguem B, Tangara F. 2018. Deep convolution neural network for image
622   recognition. Ecological Informatics 48:257–268. DOI: 10.1016/j.ecoinf.2018.10.002.
623 Verhoeven G, Sevara C, Karel W, Ressl C, Doneus M, Briese C. 2013. Undistorting the Past:
624   New Techniques for Orthorectification of Archaeological Aerial Frame Imagery. In: 31–67.
625   DOI: 10.1007/978-3-319-01784-6_3.

626 Waycott M, Duarte CM, Carruthers TJB, Orth RJ, Dennison WC, Olyarnik S, Calladine A,
627   Fourqurean JW, Heck KL, Hughes AR, Kendrick GA, Kenworthy WJ, Short FT, Williams
628   SL. 2009. Accelerating loss of seagrasses across the globe threatens coastal ecosystems.
629   Proceedings of the National Academy of Sciences 106:12377–12381. DOI:
630   10.1073/pnas.0905620106.
631 Xu SS, Xu SS, Zhou Y, Yue S, Zhang X, Gu R, Zhang Y, Qiao Y, Liu M. 2021. Long-term
632   changes in the unique and largest seagrass meadows in the Bohai sea (China) using
633   satellite (1974–2019) and sonar data: Implication for conservation and restorationc.
634   Remote Sensing 13:1–16. DOI: 10.3390/rs13050856.
635 Yamakita T, Sodeyama F, Whanpetch N, Watanabe K, Nakaoka M. 2019. Application of deep
636   learning techniques for determining the spatial extent and classification of seagrass beds,
637   Trang, Thailand. Botanica Marina 62:291–307. DOI: 10.1515/bot-2018-0017.
638 Yamakita T, Watanabe K, Nakaoka M. 2011. Asynchronous local dynamics contributes to
639   stability of a seagrass bed in Tokyo Bay. Ecography 34:519–528. DOI: 10.1111/j.1600-
640   0587.2010.06490.x.
641 Yamato C, Ichikawa K, Arai N, Tanaka K, Nishiyama T, Kittiwattanawong K. 2021. Deep
642   neural networks based automated extraction of dugong feeding trails from UAV images in
643   the intertidal seagrass beds. PLoS ONE 16(8): e0255586. DOI:
644   10.1371/journal.pone.0255586.
645 Yu Q, Gong P, Clinton N, Biging G, Kelly M, Schirokauer D. 2006. Object-based Detailed
646   Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery.
647   Photogrammetric Engineering & Remote Sensing 72:799–811.
648 Zoffoli ML, Gernez P, Godet L, Peters S, Oiry S, Barillé L. 2021. Decadal increase in the
649   ecological status of a North-Atlantic intertidal seagrass meadow observed with multi-
650   mission satellite time-series. Ecological Indicators 130:108033. DOI:
651   10.1016/j.ecolind.2021.108033.
652 Zoffoli ML, Gernez P, Rosa P, Le Bris A, Brando VE, Barillé A-L, Harin N, Peters S, Poser K,
653   Spaias L, Peralta G, Barillé L. 2020. Sentinel-2 remote sensing of *Zostera noltei*-
654   dominated intertidal seagrass meadows. Remote Sensing of Environment 251:112020.
655   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.

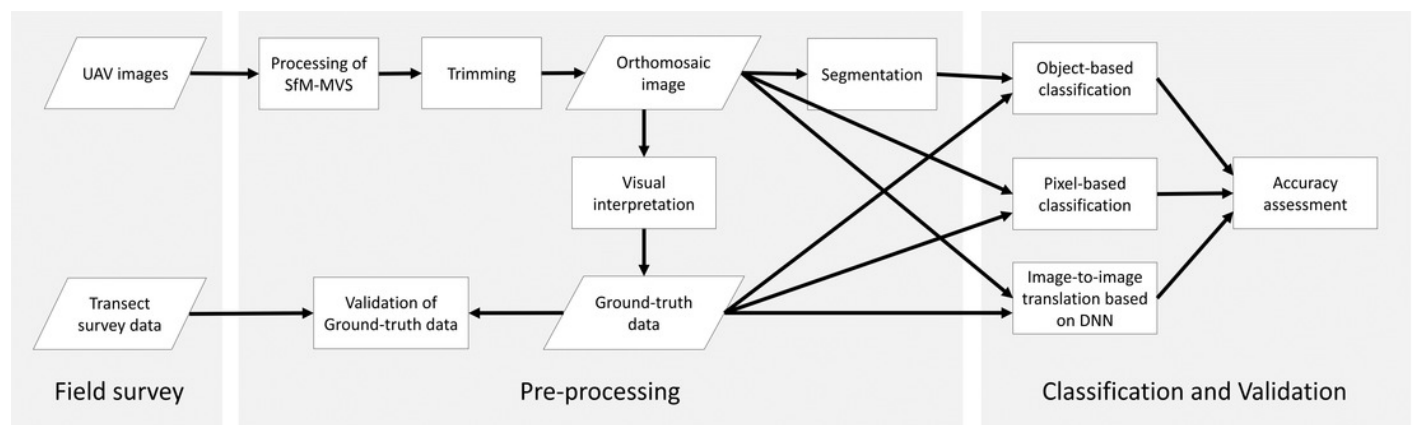



Figure 2

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.

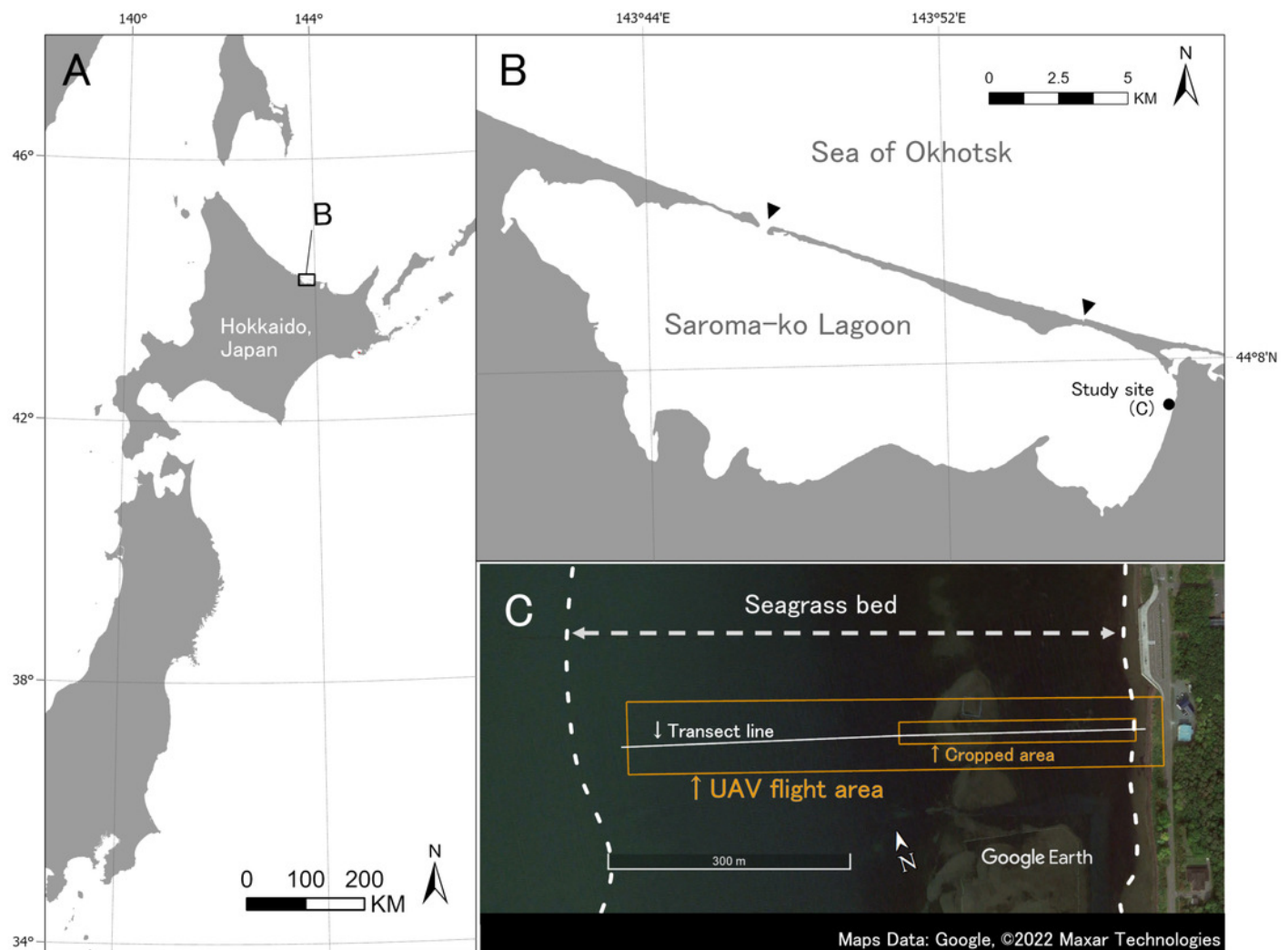



Figure 3

Frequencies of RGB values for each benthic class

The histograms show the frequencies of pixel values by color and class. The solid gray vertical line is the average value. Overall, there is a large overlap among classes. The red value in BA are relatively higher.

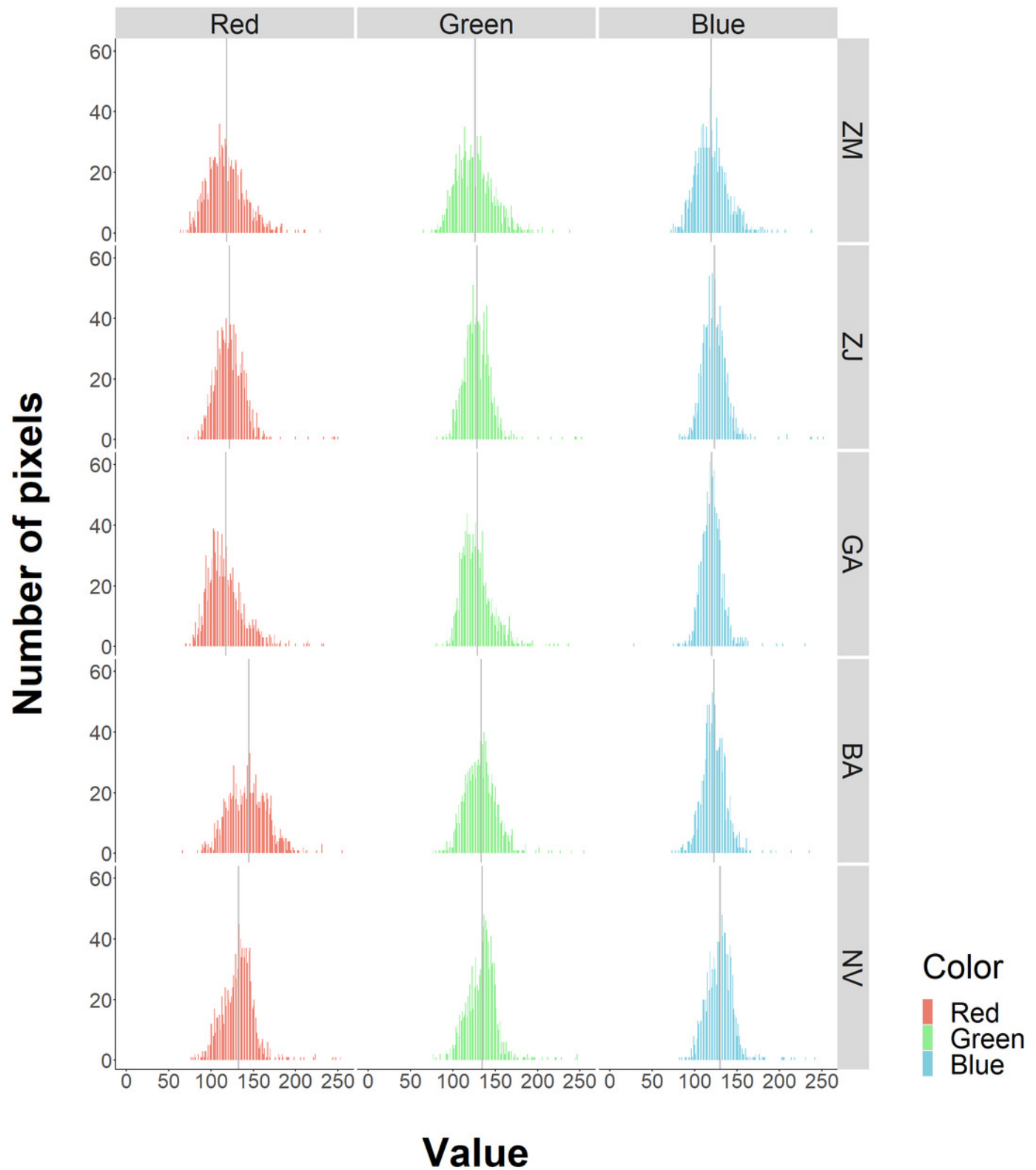



Figure 4

Orthoimage of each benthic class

A part of the orthoimage of each class is shown. (A): *Z. marina*, (B): *Z. japonica*, (C): green algae, (D): brown algae, and (E): no vegetation. Each benthic class has the unique visual appearance, which enable our visual interpretation.

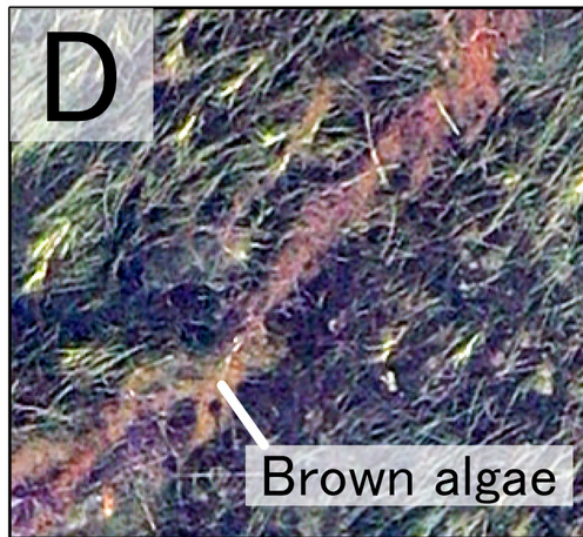
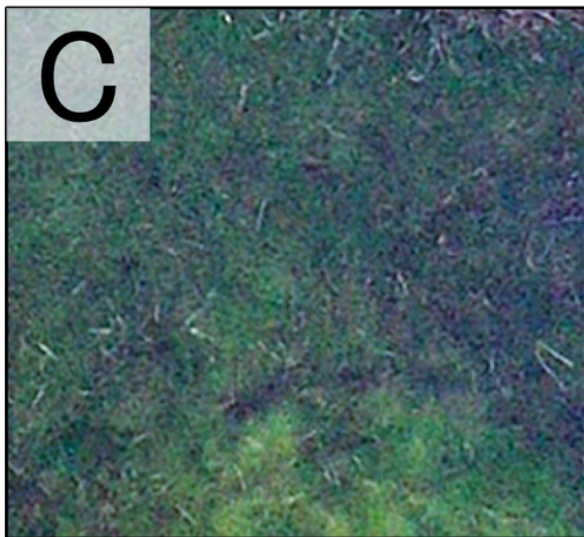
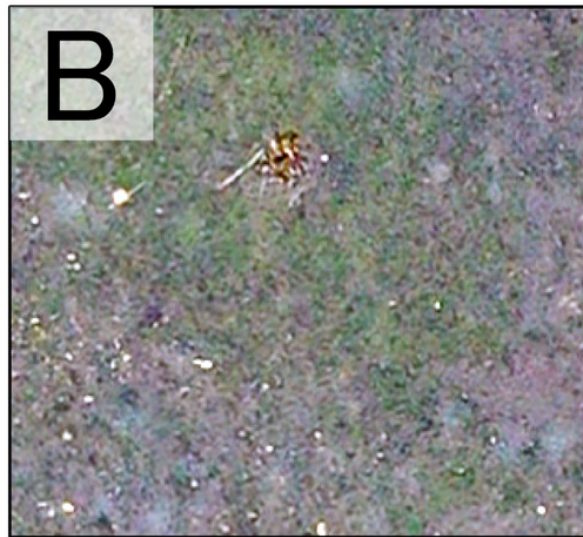
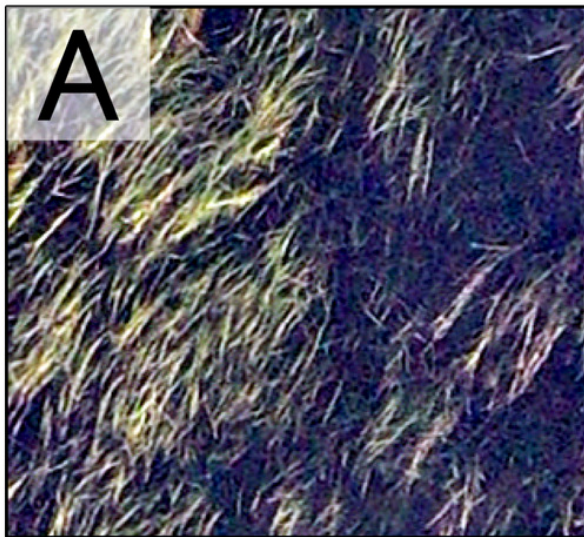
Brown algae

0   0.5   1m

Figure 5

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.

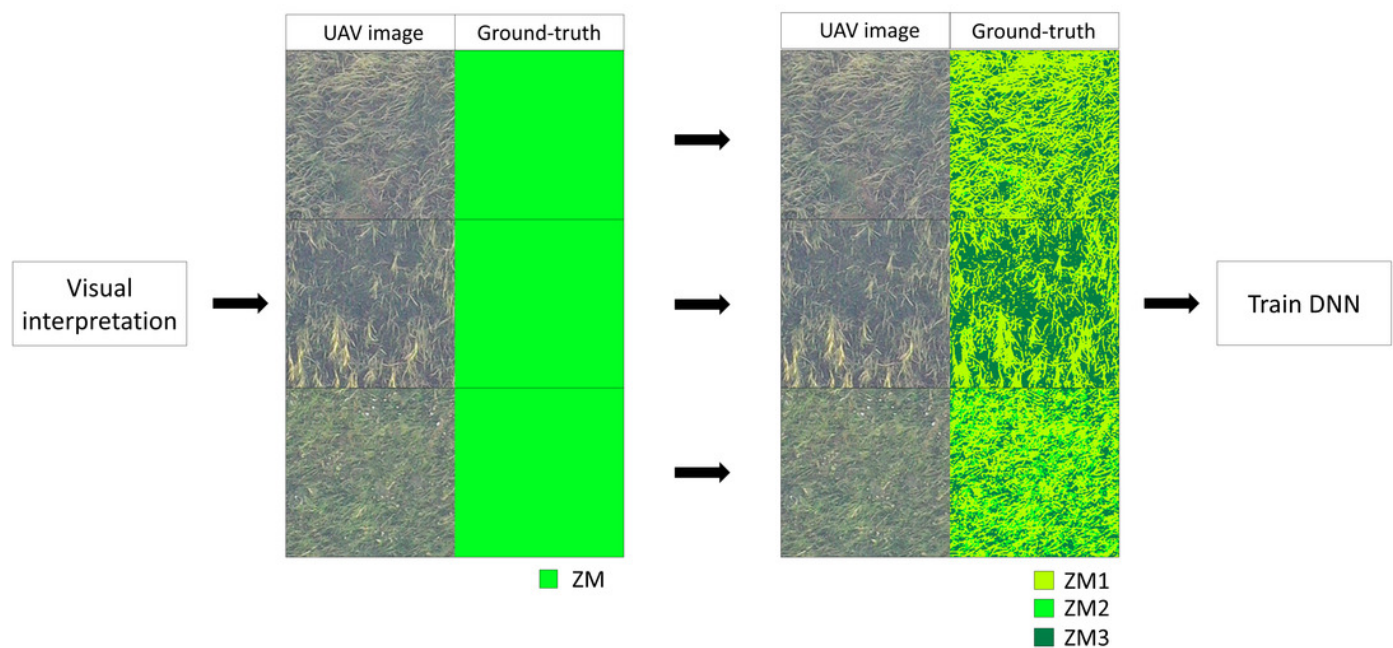



Figure 6

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.

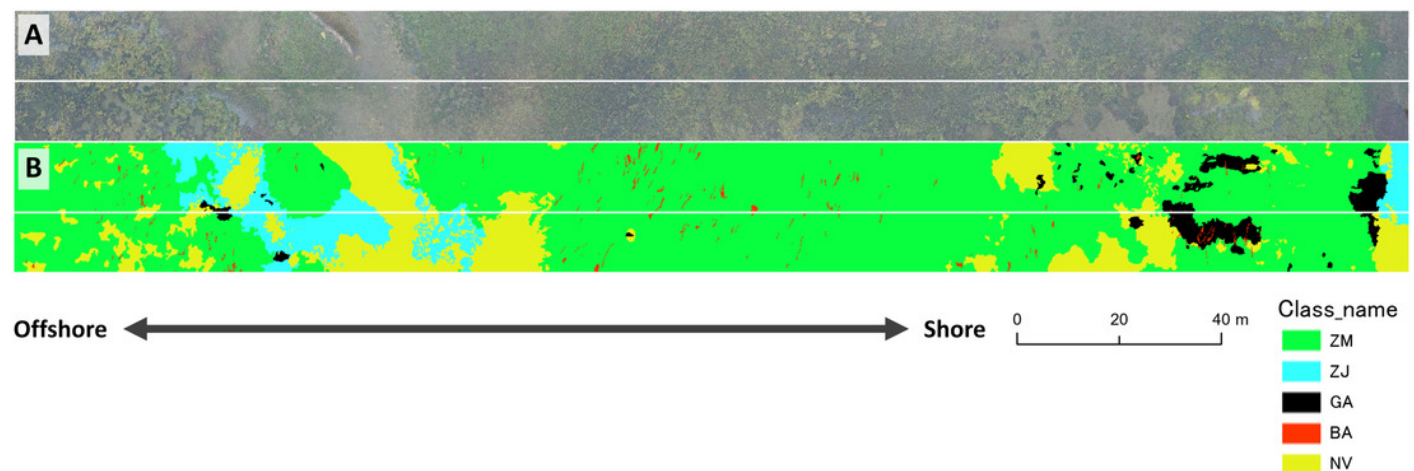
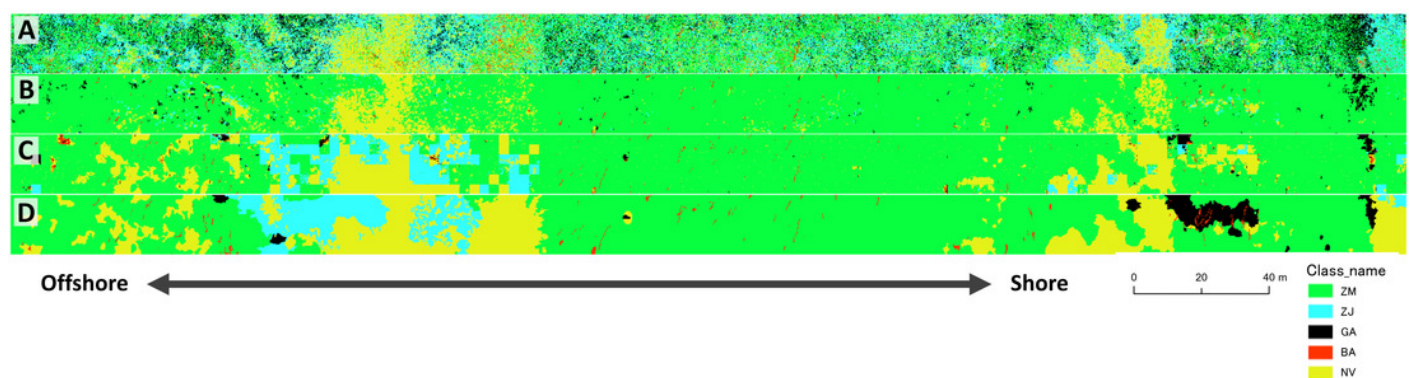



Figure 7

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. ZM: *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

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.

1 **Table 1:**

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

3 UA: User Accuracy, PA: Producer Accuracy, OA: Overall Accuracy, K: Kappa coefficient. ZM:

4 *Zostera marina*, ZJ: *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

		Field data							UA
		ZM	ZJ	GA	BA	RA	SD	total	
	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
Visual interpretation	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							

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

1 **Table 2. Confusion matrices evaluating accuracy of different mapping methods (A: pixel-**

2 **based, B: object-based, C: DNN) based on the ground-truth data.** UA: User Accuracy, PA:

3 Producer Accuracy, OA: Overall Accuracy, K: Kappa coefficient. ZM: *Zostera marina*, ZJ:

4 *Zostera japonica*, GA: Green Algae, BA: Brown Alga, NV: No Vegetation.

7 A

		Ground-truth						UA
		ZM	ZJ	GA	BA	NV	Total	
	ZM	1364	83	43	3	253	1746	0.781
	ZJ	912	178	43	5	280	1418	0.126
Map	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						

17 B

		Ground-truth						UA
		ZM	ZJ	GA	BA	NV	Total	
	ZM	3137	387	124	14	691	4353	0.721
	ZJ	16	9	3	1	23	52	0.173
Map	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						

33 C

34

		Ground-truth						UA
		ZM	ZJ	GA	BA	NV	Total	
	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
Map	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						