Mapping Marine Macroalgae along the Norwegian Coast Using Hyperspectral UAV Imaging and Convolutional Nets for Semantic Segmentation

Skjelvareid, M. H., Rinde, E., Hancke, K., Blix, K. & Hoarau, G. G

Published in: IEEE International Geoscience and Remote Sensing Symposium proceedings
DOI: 10.1109/IGARSS52108.2023.10282809

Available online: 20 Okt 2023

Citation:
Skjelvareid, M. H., Rinde, E., Hancke, K., Blix, K. & Hoarau, G. G. (2023). Mapping Marine Macroalgae along the Norwegian Coast Using Hyperspectral UAV Imaging and Convolutional Nets for Semantic Segmentation. IEEE International Geoscience and Remote Sensing Symposium proceedings. doi: 10.1109/IGARSS52108.2023.10282809

MAPPING MARINE MACROALGAE ALONG THE NORWEGIAN COAST USING HYPERSPECTRAL UAV IMAGING AND CONVOLUTIONAL NETS FOR SEMANTIC SEGMENTATION

Martin H. Skjelvareid [1], *Eli Rinde* [2], *Kasper Hancke* [2], *Katalin Blix* [3], *Galice G. Hoarau* [4]

[1] Dep. of Comp. Sci. and Comp. Eng., UiT The Arctic University of Norway, Norway
[2] Department of Marine Biology, Norwegian Institute for Water Research, Norway
[3] Department of Physics and Technology, UiT The Arctic University of Norway, Norway
[4] Faculty of Biosciences and Aquaculture, Nord University, Norway

ABSTRACT

Marine macroalgae form underwater "blue forests" with several important functions. Hyperspectral imaging from unmanned aerial vehicles provides a rich set of spectral and spatial data that can be used to map the distribution of such macroalgae. Results from a study using 81 annotated hyperspectral images from the Norwegian coast are presented. A U-net convolutional network was used for classification, and accuracies for all macroalgae classes were above 90%, indicating the potential of the method as an accurate tool for blue forest monitoring.

1. INTRODUCTION

Marine macroalgae have several important functions: Capturing carbon and contributing to primary production, forming structures that are habitats for a plethora of marine animals, and providing food for both animals and humans[1]. In many regions across the world, the "blue forests" are changing rapidly due to climate change and human activity. The current methods for measuring such changes are mainly based on manual sampling (e.g. diving or using a "drop camera"), and imaging from satellites or airplanes.

Satellite images generally have a fairly coarse spatial resolution, and imaging from airplanes is costly and requires detailed planning. Unmanned aerial vehicles (UAVs) have recently been introduced as a low-cost and flexible imaging platform for remote sensing. In this paper we present a study where blue forests were imaged using a hyperspectral camera mounted on a UAV. Hyperspectral cameras capture images with hundreds of spectral channels, and thus provide more detailed information than e.g. RGB imagery.

Maerl beds are a type of seabed formed by unattached nodules of coralline algae. Maerl beds are habitats with a high diversity and abundance of marine species, but are threatened in large parts of Europe [2]. In Norway, their status is unknown due to lack of data. The data presented in this paper was collected in an area with a high abundance of maerl, to study whether maerl beds can be identified using remote sensing from UAVs or satellites. The area also has a high abundance of kelp, which is one of the most important blue forest species.

Classifying the individual pixels of a hyperspectral image is a semantic segmentation task. Convolutional neural nets (CNNs), and particularly the encoder-decoder architectures first introduced with the U-Net[3], have demonstrated very high performance on semantic segmentation tasks. A convolutional network combines the values of neighboring pixels, and thus uses both spectral and spatial (textural) features for classification. In this paper we show that a relatively small U-net can yield good accuracy for macroalgae classification.

2. METHODOLOGY

2.1. Imaging and annotation

The data presented in this paper was collected on the island of Søla (65.68°N, 11.71°E), Norway. Shallow waters (<15m) close to the island were imaged using a Resonon Pika L hyperspectral camera mounted on a DJI Matrice 600 Pro UAV, shown in Fig. 1. The hyperspectral camera has 300 spectral channels, spanning a wavelength range of 400-1000 nm. A spectrometer measuring downwelling irradiance was mounted on top of the UAV, and the irradiance measurements were used in post-processing to convert radiance to remote sensing reflectance, $R_{rs}(\lambda)$. The UAV was flown at 50m altitude, yielding a ground sampling distance of 4 cm. On the day of the imaging (August 23. 2023), the weather was good, with cloudless skies and low wind, yielding high quality images with a low amount of light reflected from the water surface.

Ground truth was collected using several methods: In shallow areas (< 5 meter), the seafloor was imaged from an autonomous surface vehicle, from a boat, and while snorkeling, while in deeper areas (> 5 meters), a "drop camera", a small ROV, or diving was used. The ground truth informa-

Fig. 1: DJI Matrice 600 Pro drone carrying gimbal with Resonon Pika L hyperspectral camera.

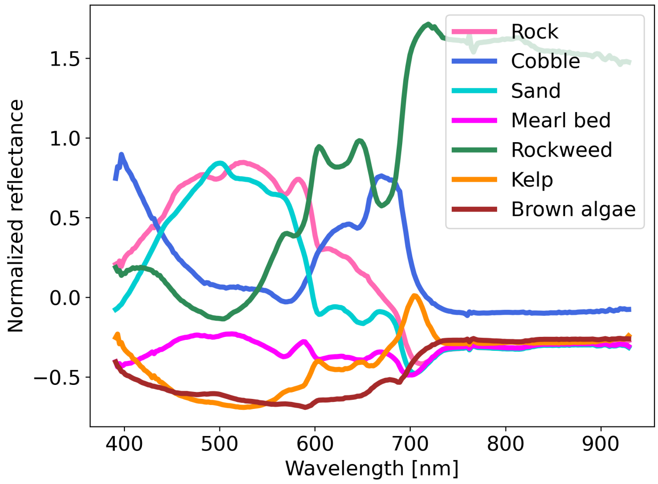

Fig. 2: Mean spectra (normalized) for each class.

tion was used to annotate the hyperspectral images, using the dominant vegetation species or nature type as classes. Seven habitat classes were defined:

1. **Rock**: Underwater bedrock.
2. **Cobble**: Large stones, mostly along shoreline.
3. **Sand**: Mainly coralline sand with high albedo.
4. **Maerl bed**: Bed with unattached nodules of coralline algae.
5. **Rockweed**: Macroalgae in the *Fucaceae* family, growing in the intertidal zone
6. **Kelp**: *Laminaria hyperborea* and *Laminaria digitata*, growing below the intertidal zone .
7. **Brown algae**: All other macroalgae not included in classes 5 and 6. Mainly *Halidrys siliquosa*, *Chorda filum*, *Desmarestia aculeata* and *Ectocarpus siliculosus*.

Example annotations are shown in Fig. 5. The classes could potentially have been split into more specific classes, e.g. at species level. However, species frequently intermingle, and accurate species-level annotations are often not possible. Using a lower number of classes also results in a wider range of annotated examples per class.

2.2. Data preprocessing and machine learning

200 GB of raw hyperspectral data were collected in total. In order to reduce the amount of data to a manageable level, the images were transformed using principal component analysis (PCA), reducing 300 spectral channels to 8 components representing 98,6 % of the variance in the dataset. Before PCA, spectra were normalized by subtracting the mean and dividing by the standard deviation. The mean (normalized) spectra for each class are shown in Fig. 2.

A U-net encoder-decoder convolutional network implemented using Tensorflow and Keras [4] was used to classify the PCA images. The structure of the network is shown in Fig. 3. The input image is fed into an augmentation layer (only used during training) which randomly flips the image horizontally or vertically. The data then flows through a number of convolution layers, with spatial downsampling and upsampling performed using standard and transpose 2D convolutions, respectively. Skip connections concatenate the data from downsampling and upsampling layers operating at the same spatial resolutions. The implementation is similar to that of the generator in the Tensorflow pix2pix tutorial [5]. The depth and size of the network was kept fairly small, with 435 000 trainable parameters. Deeper networks with a U-net structure were tested, but gave no improvement in accuracy.

71 images were used for training and 10 images were used for validation. Training images were split into 128x128 pixel ”tiles”, and only tiles with more than 5% annotated pixels were included, thus removing a significant amount of non-annotated data from the training pipeline. The trained U-net was used to predict classes for all pixels in the validation images, but only annotated pixels were used to calculate the confusion matrix.

3. RESULTS

The classification results are summarized in the confusion matrix in Fig. 4. The classification of algae (rockweed, kelp, brown algae) is fairly accurate, with over 90% accuracy for each class. The substrate classes (rock, cobble, sand, and maerl bed) have slightly lower accuracies (65-85%). There is also very low overlap between algae and substrate, i.e., misclassifications are mainly within the algae or substrate ”superclasses”.

Three example images are shown in Fig. 5. The first image demonstrates good performance in classifying rockweed, brown algae, cobble and rock. The second image shows accurate classification of kelp, but a maerl bed is misclassified as rock. The third image shows an interesting case of a maerl

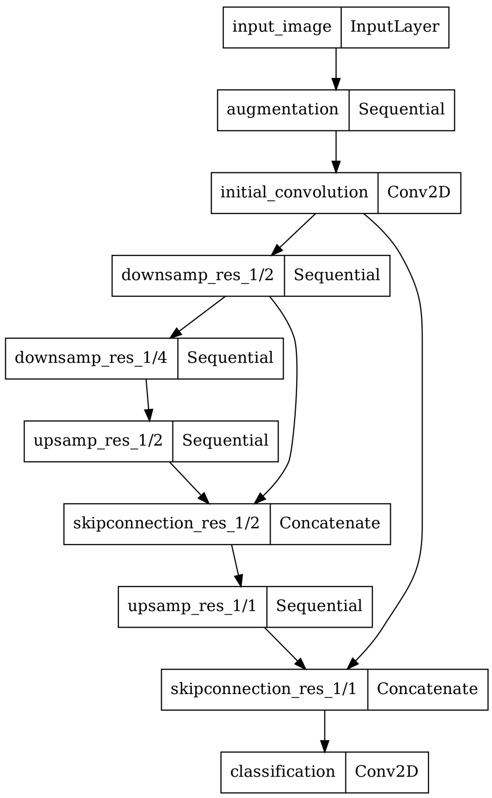

Fig. 3: Structure of convolutional net (U-net) .

bed covered by clusters of brown algae. The predicted classes show accurate separation between maerl and brown algae.

4. DISCUSSION AND CONCLUSION

The results show high classification accuracy for algae classes and lower accuracy for substrate classes. A possible explanation for this is that algae have more distinct features, both spectral and spatial, than substrates. Another potential contributing factor is that the algae often grow at distinct water depths, and that the effect of the water column on the spectrum becomes part of the spectral signature of each algae class. The substrates, on the other hand, are present at all water depths, and even though there may be spectral differences between the substrate classes, the varying water depth introduces "noise" in the spectra. Combining bathymetry (depth) data with hyperspectral data may improve the classification accuracy for all classes.

The high accuracies achieved for macroalgae mapping demonstrate that UAV-based remote sensing can be a useful tool for monitoring blue forests. UAV hyperspectral imaging is complex and costly compared to RGB imaging, but the datasets collected can be used to develop more specialized and lightweight sensors that can mounted on small UAVs.

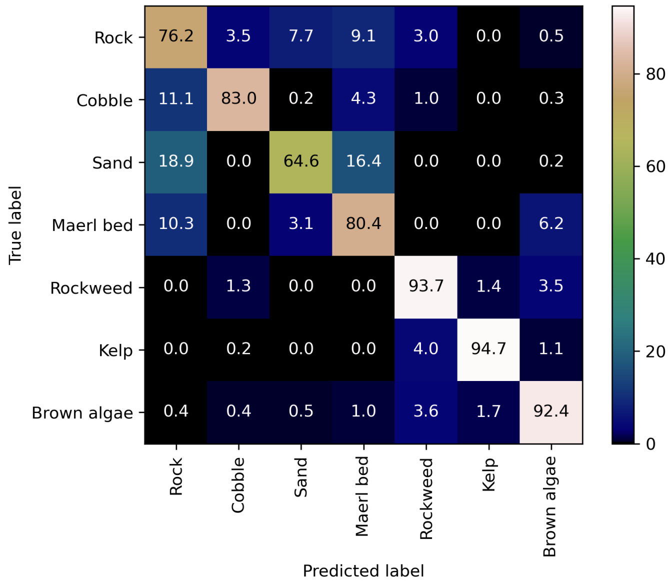

Fig. 4: Confusion matrix for classification of validation images. For each class (true label), the numbers (in percent) indicate how the predictions are distributed across classes.

5. ACKNOWLEDGMENT

This work is part of the MASSIMAL project, funded by the Research Council of Norway (project number 301317).

6. REFERENCES

[1] Helene Frigstad et al., *Blue Carbon–climate adaptation, CO2 uptake and sequestration of carbon in Nordic blue forests: Results from the Nordic Blue Carbon Project*, Nordic Council of Ministers, 2021.

[2] OSPAR, "OSPAR Status Assesment 2019 - Maerl beds," 2019, https://www.ospar.org/work-areas/bdc/species-habitats/list-of-threatened-declining-species-habitats/habitats/maerl-beds.

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Lecture Notes in Computer Science*, vol. 9351, pp. 234–241, 2015.

[4] François Chollet et al., "Keras," 2015, https://keras.io.

[5] Tensorflow, "Pix2pix: Image-to-image translation with a conditional gan," 2023, https://www.tensorflow.org/tutorials/generative/pix2pix.

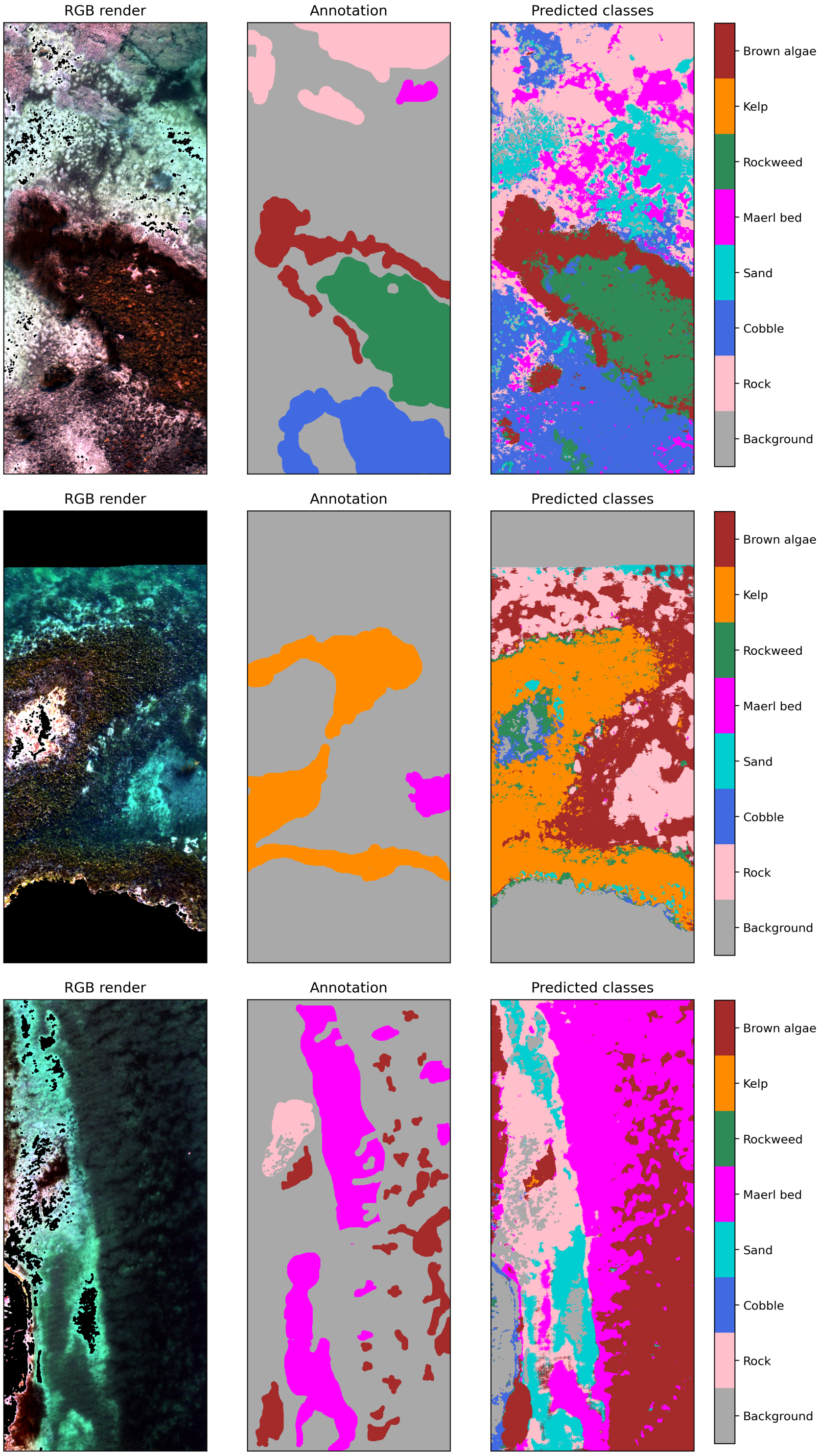

Fig. 5: Example images. RGB images (left) were made from hyperspectral image bands corresponding to 640, 550 and 460 nm. Saturated pixels are black. Each image band was independently "percentile stretched" (2 to 98 %). Annotations (middle) show manually drawn classifications. Non-annotated areas are shown as gray ("Background"). Predicted classes (right) show predictions from U-net.