An Autoencoder Model of Bathymetry and Multibeam Echosound Backscatter

Shakhboz Abdulazizov
Faculty of Computer Science
Dalhousie University
Halifax, Canada
s.abdulazizov@dal.ca

Thomas Trappenberg
Faculty of Computer Science
Dalhousie University
Halifax, Canada
tt@cs.dal.ca

Scott C. Lowe
Faculty of Computer Science
Dalhousie University
Halifax, Canada
scott.lowe@dal.ca

***Abstract*—Benthic habitat mapping is crucial to monitor and understand ongoing changes to the ocean environment caused by humanity and preserve fragile ocean ecosystems. Benthic habitats can be identified precisely from underwater images, but these images can not be collected at sufficient scale to build a habitat map on their own. Meanwhile, large-scale surveys can be conducted with multi-beam echosound that can collect both bathymetry and backscatter data, however this data is hard for humans to interpret. Consequently, benthic habitat maps are currently using simple linear models to classify habitats using hand-picked features from the echosound data. We aim to improve benthic habitat mapping by training a model to classify the habitat from the underlying backscatter and bathymetry maps. Towards this end, in this paper we show build an autoencoder model of both bathymetry and backscatter that can extract high-level features from the echosound data. In the future, this model will be used for habitat classification.**

***Index Terms*—benthic habitat, backscatter, bathymetry, underwater imagery, convolutional autoencoder**

I. Introduction

A benthic habitat is an underwater environment on the seafloor that sustains a particular community of plants and animals. The mapping of benthic habitats is important to discern changes in the distribution of habitats and understand ongoing changes to ocean environments, especially those driven by human activities, and protect fragile ocean ecosystems.

Benthic habitat maps are created using a variety of data sources. The most precise data is visible-light images of the seafloor, collected by divers and Autonomous Underwater Vehicles (AUV) equipped with cameras. Marine scientists can examine these photographs to identify the habitat type. However, these dives are highly localized and cannot be applied at scale. Meanwhile, large-scale surveys can be conducted with techniques such as multi-beam echosound (MBES), collected from aboard a ship traversing the ocean surface. MBES can collect both bathymetry and backscatter data. Bathymetry is the underwater equivalent of topography and comprises the depth of the ocean floor, derived from the latency of the echosound return signal. Backscatter is the intensity of the echosound return, which provides information about the hardness of the seafloor and can be used to differentiate between different types of substrate.

Currently, benthic habitat maps are created by extracting hand-crafted features from the backscatter such as depth, slope, aspect, and rugosity, and fitting a simple linear model to generalize from the classifications at ground-truthing camera sites to the rest of the survey [4]. Developments in the field of computer vision have demonstrated that end-to-end differentiable deep learning models, which learn to extract features from the raw data, greatly exceed the performance of models utilizing hand-crafted features. We intend to apply modern computer vision techniques to this problem in order to extract more data pertinent to benthic habitat classification from the maps of bathymetry and backscatter. Our aim is to build a model which can improve the accuracy of benthic habitat maps, and minimize the need for costly camera deployments so these maps can be generated at scale.

II. Related Work

Kostylev et al. [1] collected benthic imagery using a drop camera and inferred habitat types from this imagery were related to the terrain complexity, depth, water current, and backscatter to produce a habitat map. This method is limited by the slow rate of data acquisition and sparsity of collected sample using a drop camera. Benthic imaging AUVs equipped with advanced navigation solutions collect georeferenced imagery as they run their survey paths, thereby greatly increasing the samples collected. The abundance of data collected by AUVs enables the use of data-driven machine learning models.

Marsh and Brown [2] trained an unsupervised bathymetry and backscatter classifier using self-organizing maps. Although this classifier groups similar bathymetry and backscatter areas, it lacks corresponding habitat classes requiring further classification. Ahsan et al. [3] use Gaussian Mixture Models (GMMs) to predict the habitat class from bathymetry, by first extracting the morphological features of rugosity, slope, and aspect. Shields et al. [4] uses a denoising autoencoder for feature extraction from the bathymetry rather than extracting morphological features. This allows more information in the feature space than manually collected.

In this paper, we build on the techniques of Shields et. al. [4] by incorporating backscatter data alongside bathymetry data, and extract features from these using an autoencoder. We anticipate that the learnt representation will be useful for habitat classification.

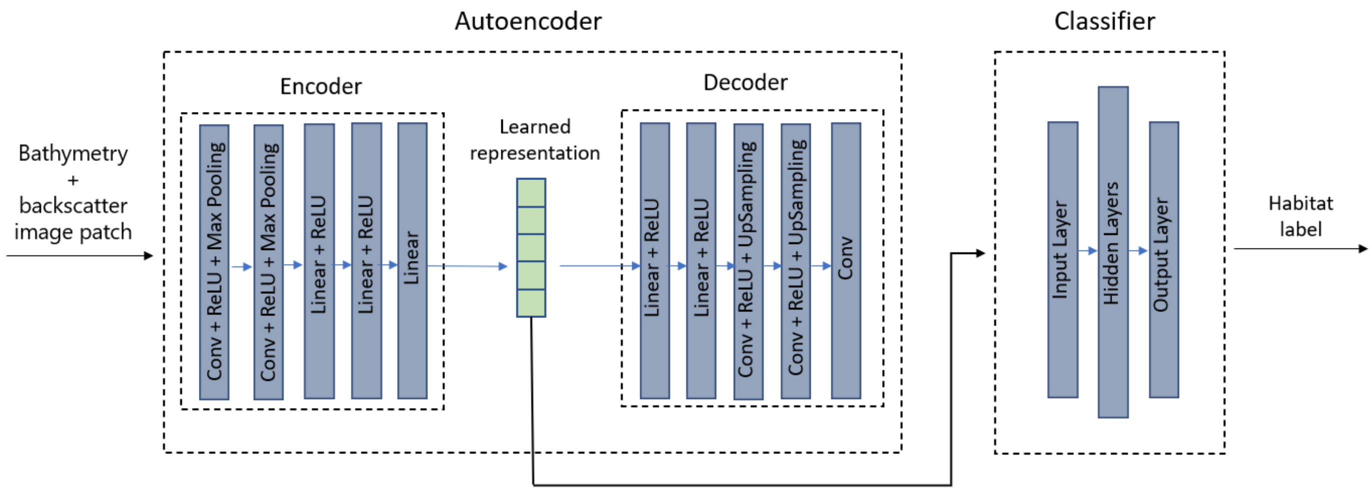

Fig. 1. Model architecture.

III. Methods

Bathymetry and backscatter data are strongly correlated to habitat class as they carry information about depth, slope, aspect, and rugosity that are the main factors affecting benthic habitat [4]. In order to extract useful features from acoustic data, we trained a convolutional autoencoder on the bathymetry and backscatter data, which were presented to the network simultaneously as a two-channel image. An autoencoder has two components: an encoder, which processes the input and produces a compressed representation of it; and a decoder that reconstructs the original input from the compressed representation. The learnt encoding can then be used for downstream tasks, taking advantage of the features extracted by the encoder. In our case, this would be a benthic habitat classifier.

We implemented an autoencoder in Keras, the architecture for which is illustrated in Fig. 1. Our encoder has two 2D convolutional layers, with 128 and 64 filters respectively, a kernel size of 3, and “same” padding. Both convolutional layers are followed by rectified linear unit (ReLU) activation and max-pooling with pool size (2,2), stride 2, the output of which is flattened. Following this, two fully connected linear layers are used with 64 units each and ReLU activation. There is one more fully connected linear layer with 32 units to constitute the latent space. The decoder has a mirrored network structure, with an additional convolutional layer with 2 filters to output the reconstructed patch.

For training the autoencoder, we used the southeast Tasmania survey that was carried out by Geoscience Australia in 2008 and 2009 with corresponding bathymetry [5] and backscatter [6] data. Missing data points within the survey grid were filled with linear interpolation. These can arise when neighboring transects of the ship during the survey are too far apart. Since inputs to our model are square and the shape of the survey is not rectilinear, we also interpolated data points outside of the convex hull of the survey using nearest-neighbor

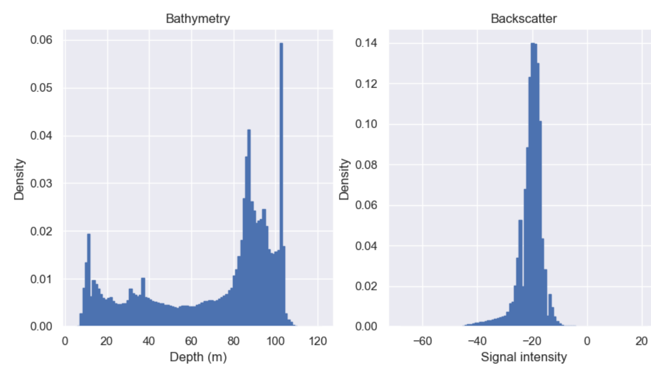

Fig. 2. Bathymetry and backscatter data histograms.

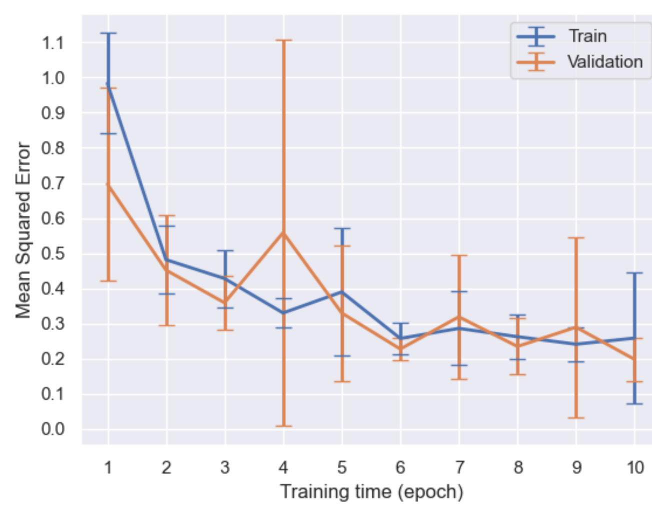

Fig. 3. Average and standard deviation of mean squared error on training and validation partitions across 10 random initialisations.

interpolation.The bathymetry grid has a resolution of $1.6\,\mathrm{m}$, whereas the backscatter has a resolution of $2.0\,\mathrm{m}$. We linearly

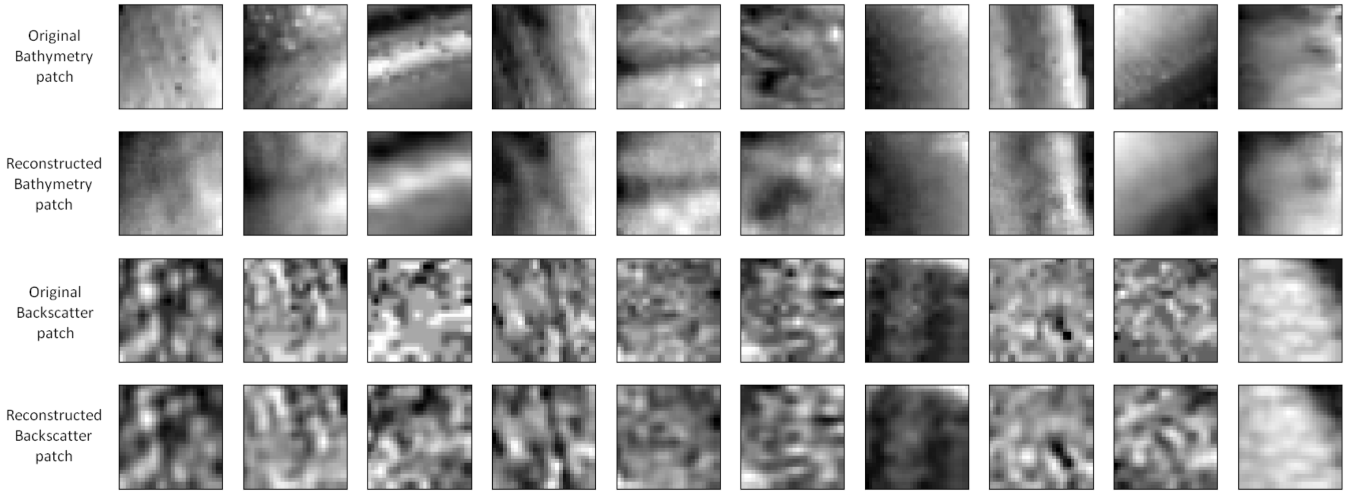

Fig. 4. Original and reconstructed bathymetry and backscatter patches for 10 samples from the test partition.

interpolated all $28\,\mathrm{m} \times 28\,\mathrm{m}$ sample patches of bathymetry and backscatter onto a new grid with a resolution of $1.0\,\mathrm{m}$. For our experiments, we extracted 11,010 random patches (10,000 for training, 1,000 for validation, 10 for test) of bathymetry and backscatter. The autoencoder was trained for 10 epochs with a batch size of 32 using the Adam [7] optimizer, learning rate 0.001, weight decay 1e-7. The loss function used was mean squared error (MSE).

IV. Results

During experiments it took around 500s to fill missing values inside and outside of the convex hull of the survey for each of bathymetry and backscatter on a local machine (CPU: Intel Core i7 7700HQ, RAM: 16GB-DDR4). As we are using the combination of bathymetry and backscatter data as a model input, we needed to make sure they both have the similar scale to guarantee that the autoencoder loss function (MSE) is affected by both data in the same level. From Fig. 2, it is clear they have different scales and require normalization. In order to normalize the input data in patch, we subtracted the mean of the patch, and divided by the average standard deviation over patches in the survey. The average standard deviation over patches was calculated by cutting 28×28 patches from whole survey data with 50% overlap. We trained the autoencoder 10 times; the average and standard deviation of the MSE loss over 10 epochs are shown in Fig. 3.

Exemplar reconstructed bathymetry and backscatter test patches are shown alongside the originals in Fig. 4. The reconstruction is accurate, indicating the autoencoder has learnt an appropriate embedding space.

V. Discussion

In this paper we extracted features from the bathymetry and backscatter data on a particular survey using a convolutional autoencoder. Our feature extraction appears to be working well, as reconstructed images are reasonably close to the originals (Fig. 4). However, our goal for this feature extraction is to learn an embedding space that is useful for the downstream habitat classification task. We will be able to better examine the utility of the features for this task once we have implemented the classifier component of the model.

Going forward, we are acquiring more bathymetry and backscatter surveys from around the globe in order to train a more powerful embedding model which can provide a suitable representation of the MBES data across a diversity of benthic habitats. We intend to use these embeddings to train a classification model which can map from the learned MBES representations to the habitat class labels.

References

[1] Vladimir Kostylev et al. “Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs.’ In: Marine Ecology Progress Series 219 (2001), pp. 121–137. DOI: 10.3354/meps219121.

[2] Ivor Marsh and Colin Brown. “Neural network classification of multibeam backscatter and bathymetry data from Stanton Bank (Area IV).” In: Applied Acoustics 70.10 (2009), pp. 1269–1276. ISSN: 0003682X. DOI: 10.1016/j.apacoust.2008.07.012.

[3] Nasir Ahsan, Stefan B. Williams, and Oscar Pizarro. “Robust broad-scale benthic habitat mapping when training data is scarce”. In: OCEANS 2012 MTS/IEEE Yeosu: The Living Ocean and Coast - Diversity of Resources and Sustainable Activities. Yeosu, South Korea, 2012. DOI: 10.1109/OCEANS-Yeosu.2012.6263540.

[4] Shields, J. et al. “Towards Adaptive Benthic Habitat Mapping.” 2020 IEEE International Conference on Robotics and Automation (ICRA) (2020): 9263-9270.

[5] Spinoccia, M. 2011. Bathymetry grids of south east Tasmania Shelf. Geoscience Australia, Canberra. DOI: 10.4225/25/53D9AD27B343F.

[6] Siwabessy, J. 2011. Backscatter grids of South East Tasmania Shelf. Geoscience Australia, Canberra. http://pid.geoscience.gov.au/dataset/ga/72000.

[7] Kingma, Diederik P. and Jimmy Ba. “Adam: A Method for Stochastic Optimization.” CoRR abs/1412.6980 (2015). arXiv: 1412.6980.