LIMNOLOGY
and
OCEANOGRAPHY: METHODS

Limnol. Oceanogr.: Methods 2026, e70037
Published 2026. This article is a U.S. Government work and is in the public domain in the USA. *Limnology and Oceanography: Methods* published by Wiley Periodicals LLC on behalf of Association for the Sciences of Limnology and Oceanography.
doi: 10.1002/lom3.70037

NEW METHODS

Seabed classification in the Gulf of Alaska from acoustic surveys using deep learning

Karuna Agarwal,[1] Christopher N. Rooper,[2] Kresimir Williams [1]*

[1]Midwater Assessment and Conservation Engineering Program, Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, Seattle, Washington, USA; [2]Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, British Columbia, Canada

Abstract

High-resolution mapping of seafloor habitats has wide applications for fisheries, conservation efforts, offshore infrastructure planning, mineral extraction, and scientific modeling. This study leverages existing widespread single-beam acoustic data and machine learning to create habitat maps for the Gulf of Alaska at a spatial resolution of less than 30 m. A convolutional neural network was trained using concurrent underwater images as the ground truth. To extract habitat type from the images, a semantic segmentation based on a random forest model was used to classify the interspersed substrate types. For the acoustic data, two models were trained: a full model using five transducer frequencies, 18, 38, 70, 120, and 200 kHz, and a reduced two-frequency model using data at 38 and 120 kHz. The acoustic inputs were normalized and scaled to reduce the effects of varying bottom depths and vessel speeds. Model outputs are seabed rugosity and proportions of substrates: sand/mud, gravel, cobble, boulder, and bedrock. Both models have similar performance, with errors of about 12% in compositions. Comparing results across different survey years shows an agreement of 95% at similar locations. The resulting substrate maps appear to correctly identify known geographic features in the Gulf of Alaska. Descriptions of co-occurrence of substrate types, their spatial distribution, and correlations in the study area are also provided.

Introduction

Seafloor mapping is important for a variety of ocean research and management activities. The mapping of marine habitats is critical for ecological studies (Swanborn et al. 2022), improves the understanding of species distributions (Laman et al. 2018; Pirtle et al. 2019), and areas that are essential for the survival of economically important fishes and other marine organisms (Rosenberg et al. 2000; Borland et al. 2021; Levin and Stunz 2005). High-resolution maps of seafloor characteristics are also important for studying benthic currents and their effects on submesoscale dynamics and mass transport, tsunami hazard assessments, marine energy and communication infrastructure planning, and mineral exploration (Lecours et al. 2016; Björk et al. 2018). However, in many remote locations such as the marine environment off the coast of Alaska, even basic bathymetric data are unavailable (Mayer et al. 2018; Zimmermann and Prescott 2023). Further, detailed maps of the types of sediments, relief, rock formations, and other important aspects of the habitat have incomplete and fragmentary coverage and thus are of limited utility for incorporating into large-scale studies. With a changing climate and increasing offshore infrastructure installations, regular monitoring of seafloor habitat becomes more salient (Stokesbury et al. 2024).

The North Pacific Ocean off Alaska supports large-scale fisheries, which are managed to ensure sustainable harvests into the future. One aspect of management is to understand fish distributions relative to seafloor habitats and to manage areas that are critical for fish populations, including foraging, spawning and nursery grounds (Pirtle et al. 2023; Gibson et al. 2022).

*Correspondence: kresimir.williams@noaa.gov

Associate editor: Xiao Hua Wang

These habitats are varied and extensive, covering an area of over 124,000 mile^2 in the Gulf of Alaska (GOA). Currently, the primary method of assessing fish abundance in this area is by conducting bottom trawl surveys (Siple et al. 2024) and to a lesser degree longline hook-and-line surveys (Siwicke and Malecha 2024). Trawl surveys can only be carried out in areas with suitable habitats, which include sandy and muddy bottoms with low relief (von Szalay and Somerton 2017). Detailed habitat mapping can benefit the planning of these surveys by precisely predicting areas where trawl sampling is less likely to be successful as well as informing the bias from sampling in limited areas.

Underwater acoustics have a long history of being used for seafloor classification, as the acoustic reflectivity and the interaction between sound waves and the seafloor is directly dependent on physical properties of the seafloor itself (Anderson et al. 2008). Acoustics-based approaches to seabed classification in Alaska include the use of single-beam and multibeam acoustic instruments and have been applied in a variety of research contexts, including survey trawlability (von Szalay and Somerton 2017; Weber et al. 2013; Stienessen et al. 2021; Pirtle et al. 2015). Several approaches have used multibeam acoustics, which provide potential for improved discrimination due to angle-dependent reflectivity and wider coverage; however, the acoustic instrumentation required for this is complex and not widely available. Separately, single-beam acoustic-based fisheries surveys for pelagic fish species have been carried out in these areas for decades, covering extensive portions of the shelf. These data represent a large untapped source for seafloor classification.

Large ecological datasets have benefited from recent developments in the application of machine learning (ML) to classification problems in science (Rubbens et al. 2023). In addition to handling statistical correspondence for large datasets, ML can also deal with the noise associated with the unknown underlying mechanisms. In particular, convolutional neural networks (CNNs), while initially applied to imagery, are being widely tested and adopted to more general problems. In this deep-learning scheme, feature extraction based on various learnt filters eliminates less relevant or correlated variables, either to improve performance or to gain explainability (Razavian et al. 2014; Jogin et al. 2018). The established ML process of training on a portion of the data followed by testing on additional unused observations provides a straightforward method to establish the precision and accuracy of the approach (Olden et al. 2008). More recently, CNNs are being applied for acoustics datasets consisting of acoustic returns over range and time, often represented as image-like echograms. For example, CNNs were used to classify the entire echogram (Senjaliya et al. 2024) as well as segmentation of the primary species from backscatter shapes of fish schools. (Brautaset et al. 2020; Zhang et al. 2024). CNNs also allow for adding auxiliary context channels, that improve prediction power; for example, including water depth (Ordonez et al. 2022; Slonimer et al. 2023).

ML applications in evaluating acoustic data have presented a possible pathway for leveraging historical survey data to develop a better understanding of seafloor habitat distributions. While different habitat types often result in visible differences in pattern and variability of the seafloor echoes, formulating specific algorithms to fully exploit these differences is challenging. Traditionally derived metrics from the seafloor echoes based on backscattering theory such as substrate roughness and hardness (Anderson et al. 2008) could be readily used; however, these quantities require additional statistical analyses to connect them to biologically meaningful habitat descriptors (Siwabessy et al. 1999). ML methods such as random forests and early neural networks have shown improved accuracy in mapping of sediment types based on derived quantities from bathymetry and multibeam acoustic-based in comparison to traditional methods (Stephens and Diesing 2014; Baker et al. 2019). Including the echosounder data in the water column in addition to the derived metrics has been shown to improve the benthic habitat mapping (Porskamp et al. 2022). CNNs are well suited to the proposed problem because the mechanistic linkages between the inputs (acoustic backscatter) and outputs (habitat classification) are inherently complex (Anderson et al. 2008). This study explores the potential of ML for extracting information that is more closely aligned with the potential uses for this data, including more detailed habitat structure and suitability of terrain for sampling using the bottom trawl directly from the acoustic backscatter.

Underwater imagery can provide high-resolution habitat descriptions but is limited in spatial coverage. Images have been leveraged as ground truth for supervised classification using multi-beam sonar data (Lucieer et al. 2013). In this study, we apply a CNN to acoustic data collected in the Gulf of Alaska, using a novel method of constructing standardized images from acoustic data utilizing multiple frequencies, paired with stereo-camera data for validation of habitat composition and seabed rugosity. We leverage a historical acoustic data set collected by the Alaska Fisheries Science Center, combined with synoptic high-resolution underwater imagery over a wide range of habitat to generate an ML model that outputs seafloor habitat compositions that are relevant for fish studies. We present two applications for the model outputs from this work in the field of fisheries management. First, the resulting habitat map is critical for species distribution maps, designation of essential fish habitats and defining species-specific habitat interaction models (Laman et al. 2018; Pirtle et al. 2023). Second, the mapping outputs may be applied for estimating trawlability, which can allow fisheries survey operators to more accurately assign effort while limiting anthropological impacts on rugged terrains and understand the survey domain (Zimmermann 2003).

Materials and procedures

Acoustic and image data collection

Acoustic and camera data were collected during fisheries surveys in the Gulf of Alaska (GOA) aboard the NOAA ship *Oscar Dyson* in 2017 and 2019. The main goal of the survey was to assess pelagic fish populations, primarily that of walleye pollock, *Gadus chalcogrammus* (Jones et al., 2022). During these surveys, the data were collected using a Simrad EK60 scientific echosounder operating at five frequencies (18, 38, 70, 120, and 200 kHz) with ping intervals close to 1 s. Acoustic instrumentation was calibrated by applying the standard approach (Simmonds et al. 1992). Briefly, this consists of suspending a metal sphere of known acoustic reflectivity below the vessel to establish correct sounder settings. Calibrations are generally conducted before and after the survey. Survey transects were perpendicular to the continental shelf edge, spanning depths of approximately 100–500 m. While the vessel was transecting at speeds generally between 11 and 12 knots, the present analysis has been applied for the entire dataset, including the times of trawling at 3–4 knots, camera deployment at speeds around 0.5–1.5 knots, and transits between transects at maximum speeds of 15 knots. The Differential Global Positioning System (Leica MX420) was used for geolocating acoustic data, providing positional accuracy of $\pm$ 5 m.

The acoustic-trawl fish abundance observations were conducted during daylight hours only; however, the acoustic data were collected continuously throughout the day. During the nighttime hours in 2017 and 2019, an underwater lowered stereo camera system (LSC) was deployed as part of a separate study comparing fish abundance and seafloor habitat. The camera deployments consisted of lowering the camera within view of the seafloor for 15 min as the vessel drifted or was actively operated at speeds between 0.5 and 1.5 knots. The camera system provided a live image feed and was dynamically controlled using a winch system to maintain a constant height from the seafloor of approximately 0.5–1 m. The system consisted of two machine vision color cameras (Chameleon3, Teledyne Flir LCC) with 2/3-in. CMOS sensors set at a resolution of 2048 $\times$ 1536 pixels (3.2 MP). The cameras were equipped with 8 mm lenses and a set of two custom-built high-intensity LED strobes. The cameras were offset by approximately 150 mm such that a left and right image were acquired to enable depth mapping for each scene. The acquired images were saved locally on the camera device at a rate of 1 Hz. At this acquisition rate and vessel speed, the successive images had a horizontal spatial separation of approximately 0.5 m. The system is also equipped with a pressure sensor (PA4LD, Keller Inc.) which was used to estimate depth. More details regarding the camera system and the untrawlable habitat study can be found in Jones et al. (2021). While the camera was being towed, the towing cable was at an angle of approximately 30 degrees from the vertical due to the drag on the camera system as the vessel moved or drifted over the seabed. As a result, the camera images represented seabed observations that were typically 1–2 min behind the acoustic data collected at the vessel at the same time (Fig. 1a). Therefore, the two datasets needed to be spatially unified as discussed later (section Assessment). For the present deep learning model, a total of 93 camera deployments from GOA-wide 2019 survey were used as the training set, and an additional 18 deployments from the same year, selected at random, were used as the model validation set.

Image analysis

Two analytical components were applied to the stereo image dataset. First, the seafloor surface rugosity was estimated by using automated depth mapping of the sea floor seen in the images. In this context, rugosity represents a measure of the small-scale (order of 0.1 vertical m per 1 horizontal m^2) variations in elevations along the bottom, enabling distinctions such as those between rippled sand bed and a flat one. Transient sand dunes can cause rugosity to be comparable to those of rocky habitats and can affect the observed acoustic backscatter (Ferrini and Flood 2006; Boggild et al. 2024). With estimates of rugosity in conjunction with composition, the habitat description can be enhanced. The stereographic imagery was calibrated using the standard checkerboard method described by Zhang (1999) and adapted for underwater applications by Williams et al. (2010). Before matching features across the left and right images, the images were converted to grayscale and the histograms equalized to improve contrast. Features on the left image were detected using the Shi-Tomasi method and matched to the right image using Lucas Kanade optical flow (Shi 1994). Only matches on the epipolar lines with acceptable disparity range were selected. The sandy habitats were prone to spurious matches due to the presence of suspended particles. In such cases, a cross-correlation based analysis was adopted. Due to changes in the camera view relative to the horizontal, the seafloor appeared tilted on the resulting depth maps. To remove this effect, a mean bed plane was estimated by fitting a plane through the detected 3D points. The rugosity was then calculated as the root mean square changes (RMSE) from the mean seafloor plane. More details about the processing can be found in the Supporting Information.

The second component of image data analysis aimed to determine the seafloor habitat composition of each image (only the left camera image was used). This was accomplished using a random forest model for semantic segmentation implemented with an open-source interactive software Ilastik (Version 1.4, Berg et al. 2019). Forty-six different features were extracted from the imagery, representing image intensities, gradients and texture, with Gaussian filters of standard deviations varying from 0.3 to 384 pixels. The model was trained for seven outputs: five corresponding to geological features of fine-grained sand/mud, coarse-grained gravel/pebble, cobble, boulders, bedrock; a "none" category that included regions which were either under- or over-exposed to the camera lights; and one corresponding to biological features such as fish and benthic invertebrates. The habitat classifications were adapted from a larger classification scheme presented by Greene et al. (1999). Rather than determining a single predominant habitat in each frame, semantic segmentation allows a proportional

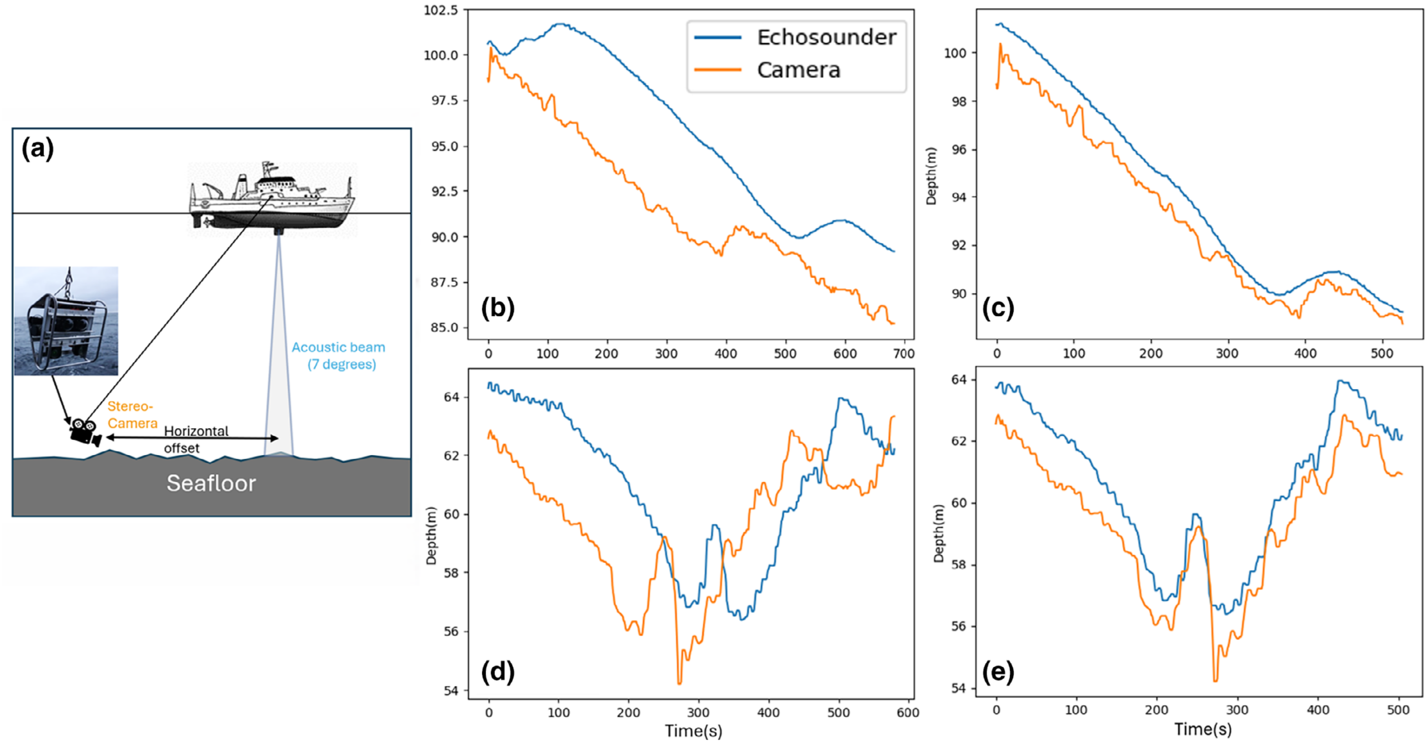

Fig. 1. (**a**) Underwater stereo imagery and acoustic survey setup; with the camera lagging behind the ship resulting in time of sampling mismatch (**b**, **d**) examples of the depth recorded by the camera and the shipboard echosounder, assuming a static lag based on a 30-degree angle of the camera cable; and (**c**, **e**) the respective matched depths using cross-correlation.

approach, with each pixel with a clear view of the seafloor assigned to a habitat type. It is to be noted that the pixel resolution changes with distance from the camera; thus, the learning is dependent on distance from the camera and pixels closer to the camera have greater representation. This effect also makes quantitative descriptions of grain size and distinctions between sand and mud, larger pebbles and smaller cobble difficult. To limit the effect of variations in the lighting conditions, all the images were color balanced before segmentation. The software was trained on a set of ~ 200 images, with out of bag errors of about 1%. Following training, the entire 2019 concurrent imaging data set consisting of 111,440 frames was processed. On a standard 32 GB RAM computer, the analysis of each 3.2 MP image took up to 3 min. Due to the high computing demand needed to classify the large volume of seafloor images, a cloud computing approach was undertaken. Using 20 instances of CPU processors, each running the Ilastik software on a batch of images, the time taken to obtain the classifications was significantly reduced. The resulting image habitat compositions, as well as rugosity, were then averaged over a spatial distance of 30 m along the seafloor (~ 50 frames, as determined by the vessel speed). Averaging across frames controlled for erroneous segmentation of a few pixels. This distance was determined to represent the scale of habitat patchiness from observations of image datasets. Since the data was evaluated for each frame, changes over a few meters can be represented by the moving average.

Matching image sampling with acoustic data

To create the machine learning training set, the acoustic data had to be paired with the spatially corresponding image ground truth. The time of arrival lag, or offset between the time at which the acoustic data and images of the same location were taken, varied with the vessel speed, camera depth, current speed and direction, for example, at greater depths the time lags could be larger. To estimate this dynamic offset, the changes in the bottom bathymetry obtained from the acoustic data and depth from the camera sensor were compared. Since the camera depth was controlled by a winch to maintain a constant altitude relative to the bottom, it closely followed any changes in seafloor depth. An initial estimate of the lag (Δt) based on 30 degrees angle of the camera cable from the vertical, the current bottom depth (D) and the ship speed (s) was calculated as $\Delta t = D/(s \tan 30°)$. An example of the depth signals based on this is illustrated in Fig. 1b,d. Then the phase mismatch between the camera depth and vessel detected bottom depth was estimated using a cross correlation of these signals. The cross-correlation coefficient peaked when the signals were aligned. (Fig. 1c,e). Note, the lag at each instance varied smoothly over time. Therefore, for each deployment, any outlier in lag estimates, defined as z-score over 2.5, was replaced with the mean of the neighboring lag estimates and the time series of lags were median filtered with a window size of five frames. For a few developments, the average cross-correlation coefficients were below 0.5, likely due to poor camera control or very little variance in depth. These deployments were not considered for analysis.

Acoustic data pre-processing

All the acoustic datafiles were processed using the PyEchoLab library for the Python programming language (Wall et al. 2018) to find the calibrated volumetric backscatter values (db re 1 m^2 m^{-3}) across the water column and the bottom depths. In addition, the acoustic data were pre-processed to reduce noise. For example, interference from the shipboard acoustic Doppler current profiler (ADCP), a separate instrument operated during some deployments and during transit, generated noise which appeared as streaks with regular frequencies on the 70 kHz and higher frequency echograms. This pattern was removed using filters applied to Fourier-transformed data in the frequency domain. Furthermore, the data were resampled for uniformity in the horizontal distance as well as number of bottom echoes. The bottom echo appears as a sharp gradient, which for varying bottom depths would be located at different ranges without normalization. Since the model is expected to learn composition based on differences in the echo appearance and not the bottom depth (bottom depth is weakly correlated with composition), the acoustic data inputs were standardized to an interval from the surface to the third bottom echo. Data from all five frequencies were then arranged sequentially as columns in a matrix for a set of seven evenly interpolated pings spanning over 45 m, resulting in a final matrix of acoustic return intensity or backscatter in units of sV (MacLennan et al. 2002). This matrix with dimensions of 1484 rows, corresponding to range, and 35 columns, corresponding to frequency and horizontal span, represented the acoustic data unit for training and analysis. By interpolating on a distance grid and not pings, the effect of varying vessel speeds was minimized. Since the equivalent acoustic beam width was seven degrees for four of the five frequencies, each ping received information over ~ 12 m diameter circle at a depth of 100 m. Therefore, at this depth, pings that are within 22.5 m distance of a ping all include information over the same 30 m spatial resolution, similar to the compositions from the image dataset. The analysis was performed for every ping (every 0.5–7.5 m), but composition output values represent a moving average over 30 m.

Deep learning model setup

Two deep learning models were trained using the described dataset: a full model with acoustic inputs that included all five acoustic data frequencies and provided composition and rugosity as outputs and a second reduced frequency model using only two frequencies. The reduced frequency model was used to evaluate incorporating data from non-survey vessels; for example, fishing vessels which typically operate echosounders with two frequencies (von Szalay and Somerton 2017). For this model, only the 38 and 120 kHz frequencies were used and the output consisted of substrate proportions only. Before training the full model, the rugosity was normalized by the maximum observed value. The model was implemented using PyTorch (Paszke et al. 2019) version 2.0.1 built in the Python programming language. The choice of representing the acoustic backscatter data as images was due to the inherent correlations across the frequency as well as horizontal space. CNN models have been shown to be good at pattern recognition and detecting image gradients, so a CNN architecture has been chosen to connect the changes in the acoustic signal over pings/frequency and depth to composition.

Figure 2 presents the model architecture, which consists of three layers of convolutions. Each of the convolutions was also followed by group normalization to limit impacts of batches (Wu and He 2018), rectified linear unit (ReLu) activation to deal with vanishing gradients (Lin and Shen 2018), and max-pooling with size 2×2 for dimensionality reduction (Akhtar and Ragavendran 2020). The resulting features from the convolutions were flattened to a vector to create three fully connected layers before the outputs. The first layer also included a 50% dropout (Wu and Gu 2015), which inactivates half the neurons randomly selected to avoid overfitting during training. The convolution-based architecture introduces a locality bias, which may be helpful when nearby values are related and present good generalization (Wang and Wu 2023). The dropout layer can aid with generalization without the locality bias. The acoustic beam spreading causes the acoustic signals to encompass wider areas with increasing depth. To ensure this behavior does not bias the learning, the second layer included the bottom depths at the current ping as an additional variable. Ordonez et al. (2022) found improved performance with a late inclusion of depth in a UNet architecture; therefore, seafloor depth was incorporated in the fully connected layer after dropout.

While training, the loss function minimized the Euclidean norm of the difference from the known composition and rugosity and the norm of deviation of the sum of the compositions from 100%. Weights were optimized using stochastic gradient descent, with a final learning rate of 0.02 and momentum 0.75. The learning rate was linearly increased with a start factor of 0.2 over 20 iterations. A batch size of 60 was used for the training. Since the composition changes were gradual over neighboring pings during the camera deployments, to avoid overfitting, the test datasets were from different deployments than the training dataset, chosen at random. Similar compositions were found in the training and test sets, except slightly more gravel/pebble representation in training. For both sets, the compositions contained a high proportion of sand/mud, with cobble and boulder being the least prevalent. The evolution of the loss is available in the

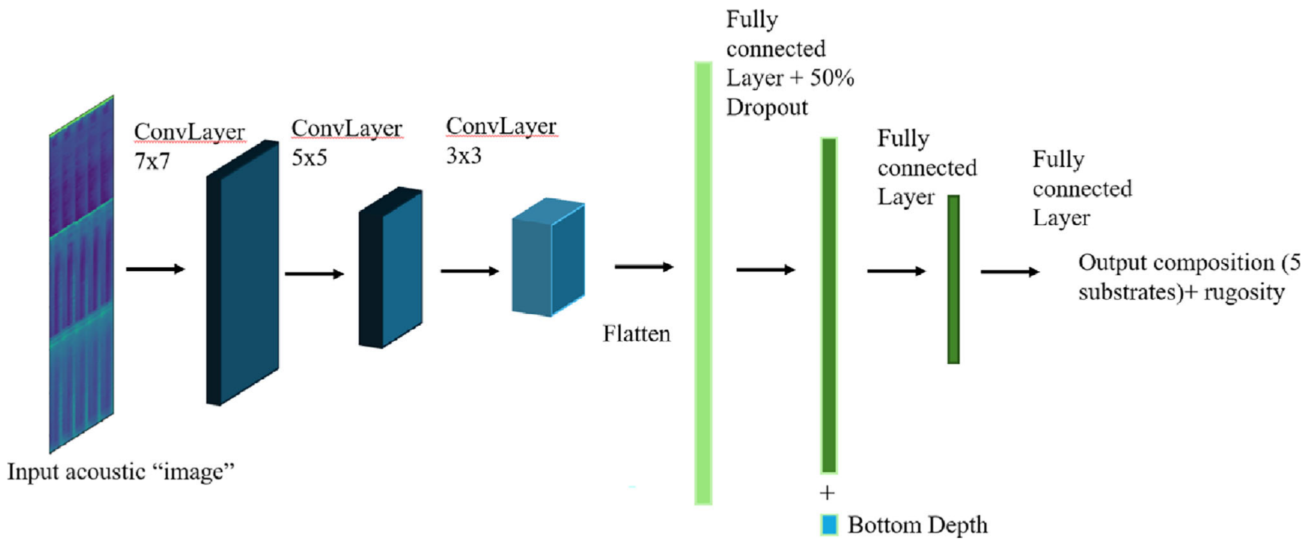

Fig. 2. Presently used neural network architecture, consisting of three layers of convolution and three fully connected layers. Each convolution layer also involves group normalization, ReLu activation, and max pooling.

Supporting Information. Training was interrupted at 11 epochs for the reduced frequency model and 19 epochs for the full model as the loss for the test dataset plateaued. Once trained, the acoustic data input generation and model execution was completed in under 1 ms. Therefore, the process can be executed in near real time, that is as soon as the ping that is over 23 m away is recorded. The results from the image classification and the application of the trained neural network to the entire survey acoustic dataset follow.

Assessment

Habitat classification from images

A representative benthic image with rocky bed and the respective three-dimensional reconstruction of the sea floor is presented in Fig. 3. The reconstruction captures the small-scale variations in the depth; that is, the rugosity over a few meters. For the present sample, consisting of cobble and boulders, the resulting rugosity value (RMSE) was 0.7 m. The image-based habitat description consists of rugosity as well as substrate composition. Figure 3 also shows example results of the semantic segmentation into substrate types for a sandy and a rocky habitat. The sandy habitat was uniformly identified, whereas the rocky habitat was primarily bedrock with some boulder, cobble, and gravel. Image-derived habitat descriptions were averaged over 30 m for each of the matched acoustic samples ("pings") before being used to train the ML model.

Habitat classification from acoustic data

Representative samples of the acoustic data inputs associated with sandy and rocky habitats are presented in Fig. 4a,b. The data inputs represent the whole water column, normalized from the surface to the third bottom echo. While the effects of seabed substrates should be isolated near the bottom echoes, the entire water column has been used to avoid unphysical gradients in the data. While the sharpest gradients are at the seabed, variations with frequency and horizontal location are also evident in Fig. 4a,b. This patterning with both axes was our rationale for the choice of the input data representation and network architecture. The mean acoustic signals at two frequencies for sandy (sand/mud exceeding 80%) and rocky (bedrock over 20%) regions across all training and test datasets are presented in Fig. 4c. Due to the depth normalization, the sample and average signals are slightly more stretched for the rocky habitat, which tends to be shallower regions. This average trend does not compromise the model's performance for deeper, rocky regions, as discussed later. However, certain patterns are still evident on average. For example, for rocky habitats, the second echo has more symmetric peaks and larger relative contribution from the higher frequencies, as compared to sandy habitats.

The accuracy of the trained model is evaluated based on the test set. The error in composition is 12% for both the full and two-frequency models, and 8% in rugosity for the full model. Due to the proportions summing up to 100%, misappropriating substrate composition by 1% is reflected in other substrate proportions; therefore, the error may not be the best metric for evaluating the models. Figure 5 presents model performance in terms of how well the most dominant substrates are identified. For the most part, the models can accurately identify sand/mud, pebble/gravel, and bedrock but are challenged by cobble and boulders, which mostly occur in small percentages. The precision, defined as correct identifications over all predictions, and the recall, defined as correct identifications out of all true labels, are also presented. For the intermediate grained pebble to boulder, the precision is better

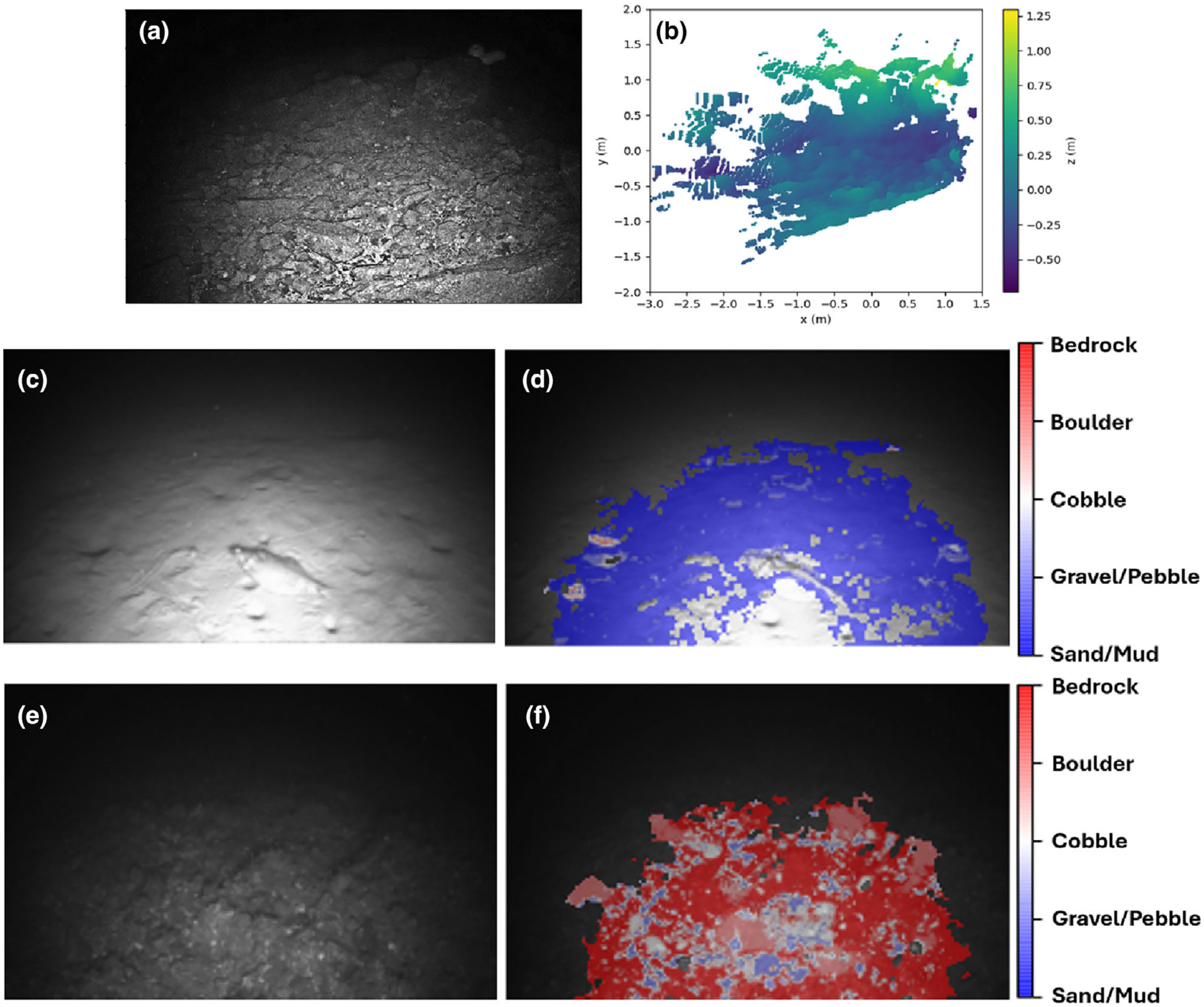

Fig. 3. (**a**) An example benthic image and (**b**) the respective stereographically reconstructed depth field. (**c**, **e**) Examples of sandy and rocky habitats with (**d**, **f**) the respective semantic segmentation into relevant substrate types.

than the recall, indicating there are significant cases of false negatives. The two-frequency model is comparable in its performance to the full model. As shown in Fig. 5, the frequency-dependent changes are gradual and so may be well captured by a single low and high frequency combination.

Figure 6 shows the GOA-wide map of proportion of sand/mud from the outputs of the full model on the 2017 and 2019 summer acoustic surveys. The map has similar characteristics compared to other sediment or trawlability studies (Bryan et al. 2023; Baker et al. 2019). The well-surveyed areas of the Shelikof Strait and Shumagin Islands that are known to be suitable for survey trawl operations show a high proportion of soft sediments. Most of the continental slopes are composed of harder seafloor. Figure 6b,c shows the sand/mud proportions along with the bathymetry to inform the sediment type for prominent topographical features around Kodiak Island. The presented bathymetry data are approximate as they are based on the NOAA acoustic survey bottom range estimates taken during summer and winter surveys of the area from 2013 to present. The shallower northern Shelikof Strait consists of primarily sandy habitat; however, the deeper western portion of the central strait is composed of rockier substrates. There is greater diversity in habitats south of Kodiak Island; however, the Barnabas and Chiniak Troughs and Marmot Bay all have high levels of soft sediments, which is consistent with the known geology of the area. While bathymetry is an important predictor, the GOA-wide correlations of depth with proportions of sand/mud and bedrock are just 18% and 24%, respectively. The rugosity outputs of the full model on the 2017 and 2019 summer acoustic surveys are presented in Supporting Information Fig. S4. Rugosity is well correlated with the proportions of boulders (71%) and bedrock (60%).

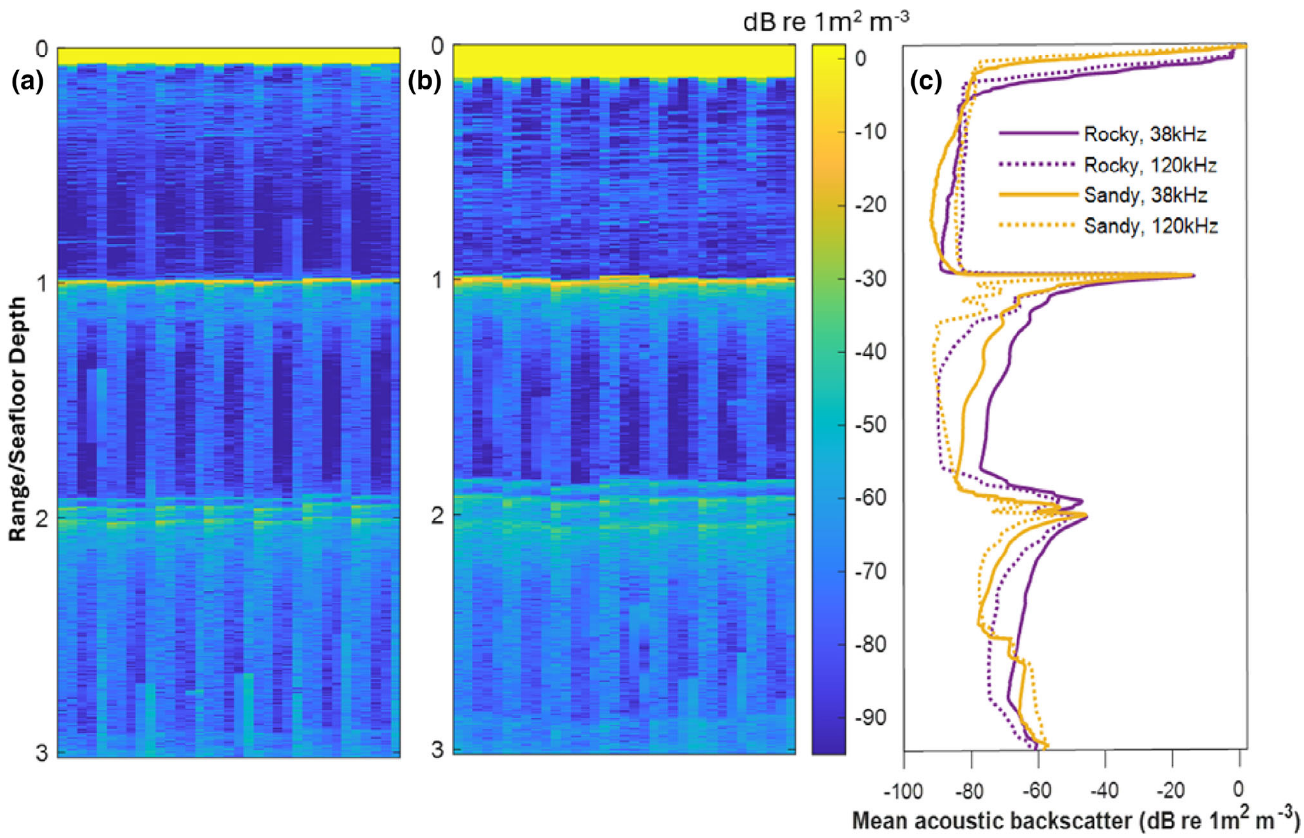

Fig. 4. Sample acoustic data inputs in the training dataset for (**a**) sandy habitat and (**b**) rocky habitat. The vertical axes represent the range relative to the bottom depth from the surface (0) to the third bottom echo and the horizontal axes represent the seven adjacent locations and five transducer frequencies. (**c**) The mean acoustic signal (for the training dataset) at the central ping locations where the proportion of sand/mud is over 80% (sandy) and where the proportion of bedrock is over 20% (rocky) at frequencies of 38 and 120 kHz.

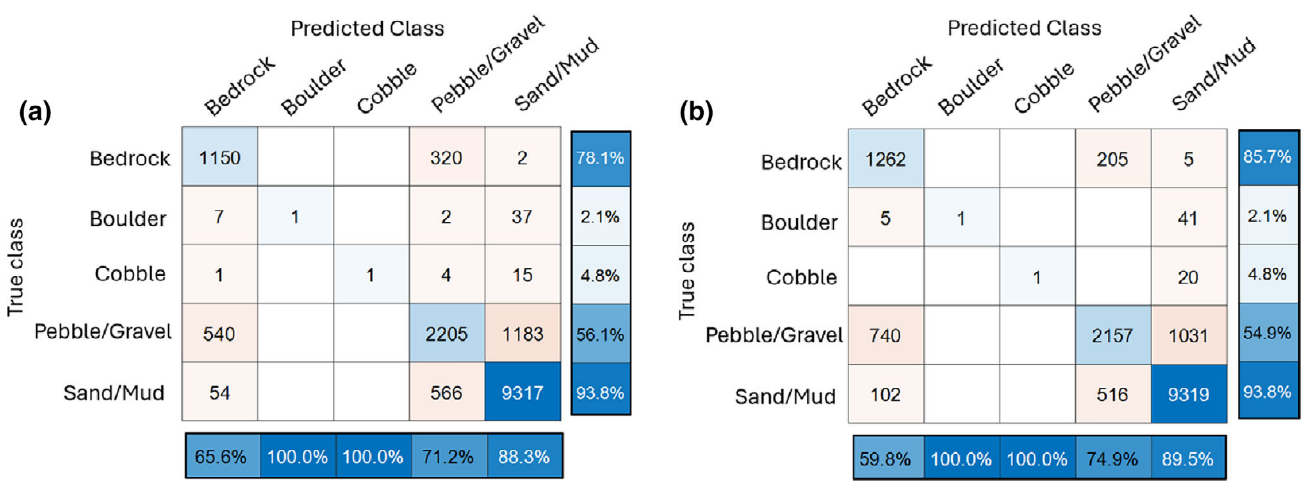

Fig. 5. The confusion matrix of the dominant substrate detected in the validation dataset for (**a**) the full model and (**b**) the reduced two-frequency model. The true classes are arranged in rows and the predictions as columns, and the class-specific precision rates at the bottom and recall rates to the right of the matrix. The darker blue color indicates better predictions.

To elucidate the spatial scales over which the habitat composition changes and compare with coarser scale maps, the spatial correlation of the composition is presented in Fig. 7a. The spatial correlations were calculated as the normalized covariance between the proportions of a substrate with the proportions of the same substrate up to a maximum

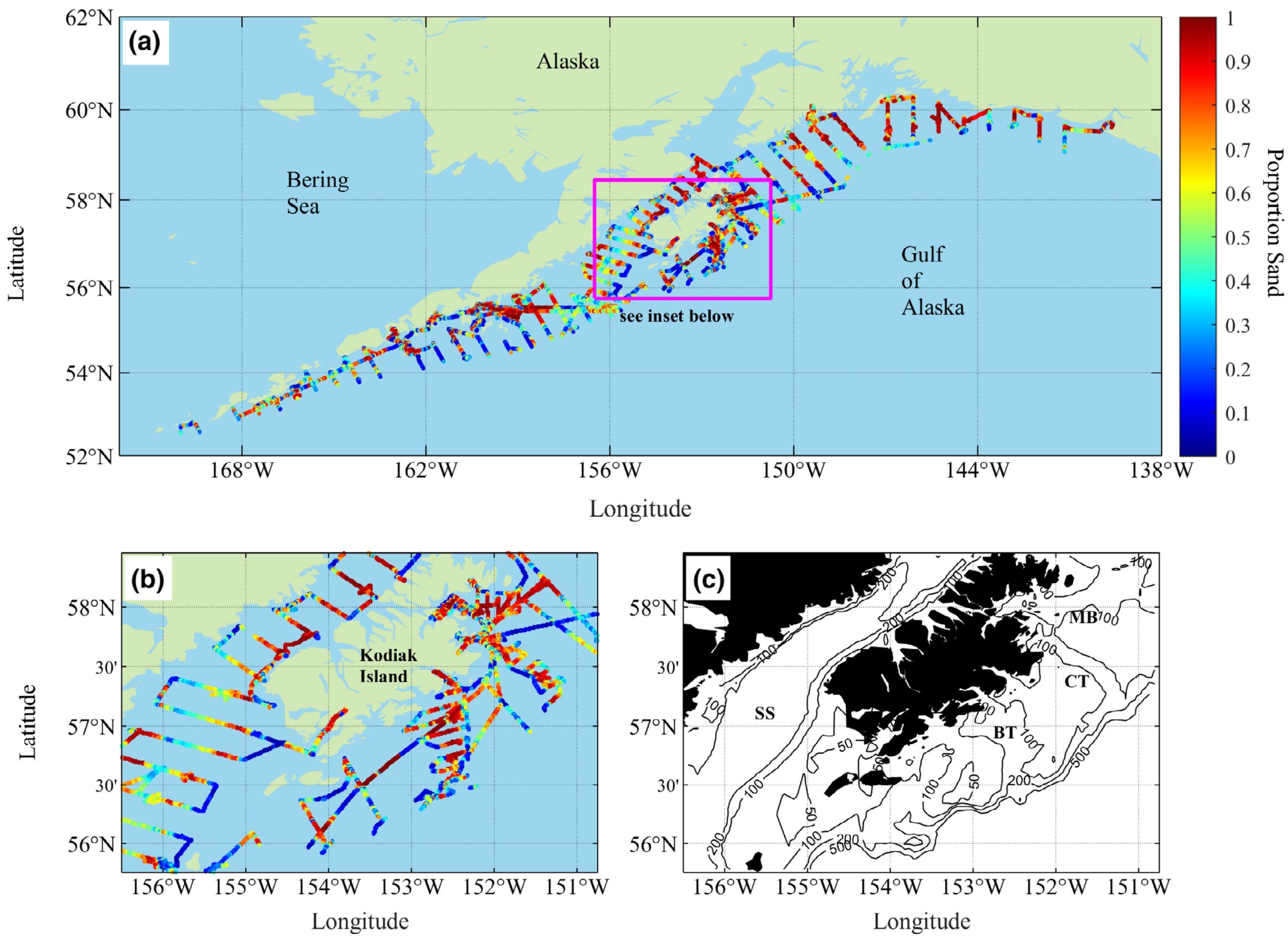

Fig. 6. (**a**) Proportion of sand/mud predicted by the ML model for 2017 and 2019 Gulf of Alaska summer surveys. (**b**) An inset shows detailed proportion of sand/mud around the Kodiak Island. (**c**) Bathymetry contours and relevant topographical features including the Shelikof Strait (SS), Barnabas Trough (BT), Chiniak Trough (CT), and Marmot Bay (MB).

geographic distance away, in increments of 10 m. Boulder and cobble habitat types were less frequent in the datasets and mostly co-occurred with bedrock or gravel respectively. Hence, subsequent analyses combine the proportions of boulders with bedrock and that of cobble with those of pebble/gravel. The softer sediments remain more correlated at larger distances, compared to the more variable bedrock. While the correlations are close to 95% at 50 m, they drop to as low as 60% at 1 km. The choice of 30 m spatial units in the analysis likely contributes to the high correlations close by; however, it is likely that any analysis performed for grids coarser than 1 km would not fully capture the habitat variability. To further assess the present model's dependence on the training dataset, and the applicability of the model to other acoustic datasets, the outputs from the 2019 survey were compared against those from the distance-weighted average in a vicinity of 50 m radius for the same locations visited by the acoustic survey in 2017. The compositions were correlated by 95%, similar to the expected values from Fig. 7a.

To further illustrate the co-occurrence of substrate types, Fig. 7b presents the ternary diagram for the distribution of the three class compositions. The ternary diagram constrains the possible proportions such that they sum up to be 100%, and the color scale is the logarithm of the number of cases. Sandy regions tend to have uniformly high proportions of sand/mud. Due to large swaths of uniform sandy habitat, a very strong peak in distributions is seen at >95% sand/mud cover. In contrast, habitats containing gravel/pebble, as well as those comprised of boulders/bedrock, are more diverse and therefore appear as the diffuse region on the logarithmic plot. For habitats that contain boulders and bedrock, the composition tends to be 25–60% bedrock and less than 10% sand/mud, with the remainder being the gravel/pebble class.

To compare the model outputs with other coarser resolution studies of seafloor habitat in the GOA, the substrate proportions were interpolated over a 1 km^2 grid by distance weighing. This grid-cell result was compared to Bryan et al. (2023), which pooled 5 different mapping layers from

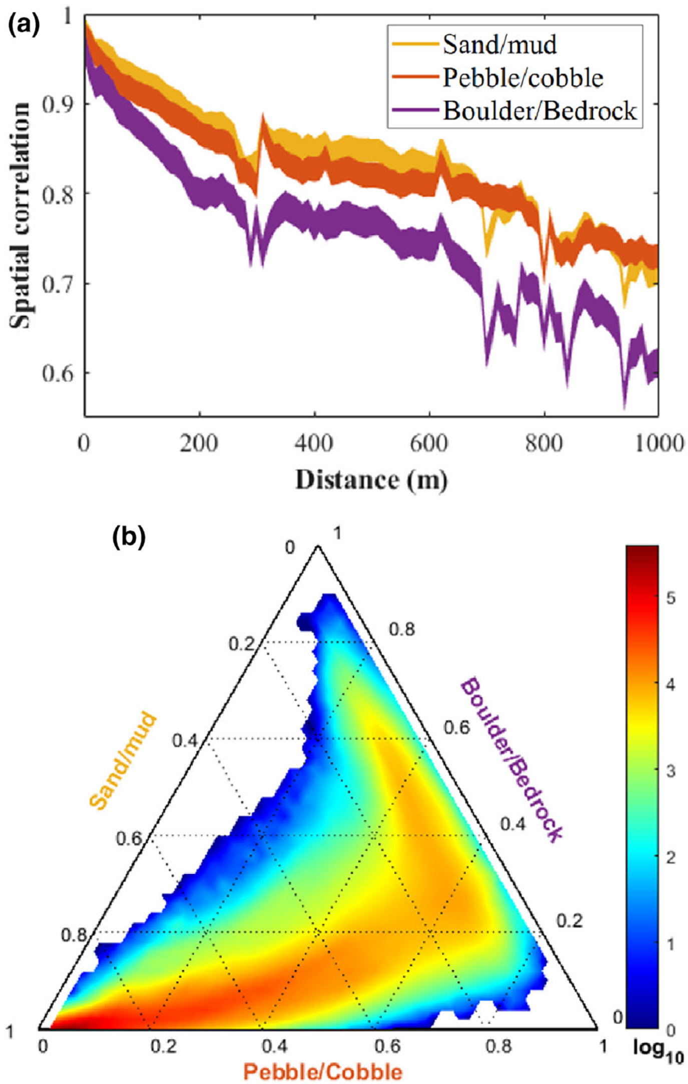

Fig. 7. (**a**) Spatial correlation of the proportion of substrates for sand/mud, pebble/cobble, and boulder/bedrock over a kilometer. The thickness prescribes the confidence levels based on bootstrapping 5 randomly selected sub-samples. (**b**) Ternary diagram for the distributions of substrate types, where the color scale represents the logarithmic occurrence of a given combination of substrate classifications.

different data sources to characterize each cell's trawlability, with a score of 5 denoting the cell is trawlable based on all the sources, and −5 untrawlable based on all. In their study, if 10% of data points within the 1 km^2 grid were untrawlable, the cell was defined as untrawlable. Trawlable habitat, characterized as values above 2 by Bryan et al. (2023), matched the regions with less than 10% bedrock cover from the current study for 75% of the cells. Similarly, the untrawlable habitat with values under −2 matched regions with greater than 10% of bedrock cover for 82% of the grid cells. Comparative maps are presented in the Supporting Information. Figure 8 shows composition maps and variance in the region around Kodiak Island for the 1 km^2 grid. In agreement with Fig. 7, large swaths of the tracks are dominated by sand/mud. The grid averaged proportions of pebble/cobble are highly correlated with that of boulder/bedrock, and inversely related to proportions of sand/mud. The bedrock proportions, where they exist, are highly variable. Unlike rocky habitats, in regions of high proportions of sand/mud, the variance in compositions is low.

Discussion

This paper presents the successful application of a ML model to identify bottom habitat, resulting in distributions of substrate cover in the Gulf of Alaska. The model results compare well with other habitat and trawlability maps from the region. The model also provides rugosity of the seafloor, which not only impacts the acoustic dead zone (Simmonds and MacLennan 2008) but is also an important habitat feature. A reduced frequency model was also implemented, which performed similarly to the full frequency model. For both models, the *F*-score, a combination of precision and recall, is around 91% and 71% for sand/mud and bedrock, respectively.

The concurrent underwater imagery data enabled the generation of ground truth compositions for the training dataset and for testing the model accuracy. While underwater imagery captures habitat variations, translation into composition and rugosity requires relying on another ML tool. For the present study, semantic segmentation through a random forest model allowed descriptions involving interspersed substrate types. However, this introduces some amount of uncertainty into the ground truth composition from images themselves. As each acoustic data input represented composition over several image frames, the uncertainty associated with misclassifying individual pixels is reduced due to cross-frame averaging. While manual validation of classification for images indicates very little bias in errors, behavior over the entire dataset with variations in deployment conditions and compositions may include minor bias. With the current training inputs, the model performance is acceptable; however, it may still benefit from additional data sources and transfer learning (Hosna et al. 2022) when deploying over non-represented regions.

The choice to use backscatter across the full water column and multiple frequencies was made to keep preprocessing of the data to a minimum and let the model derive the important features. While multibeam acoustics are usually considered the most suitable tool for habitat and seafloor mapping, we show here that more widely available single beam acoustics can have substantial utility when paired with ML. This project also provides an example of opportunistic use of existing datasets, where the original motivation behind the data collections did not include acoustic-based habitat mapping. The performance of the reduced frequency model being similar to the full frequency model (Fig. 5) is not completely surprising since other studies have been able to use the backscatter at two

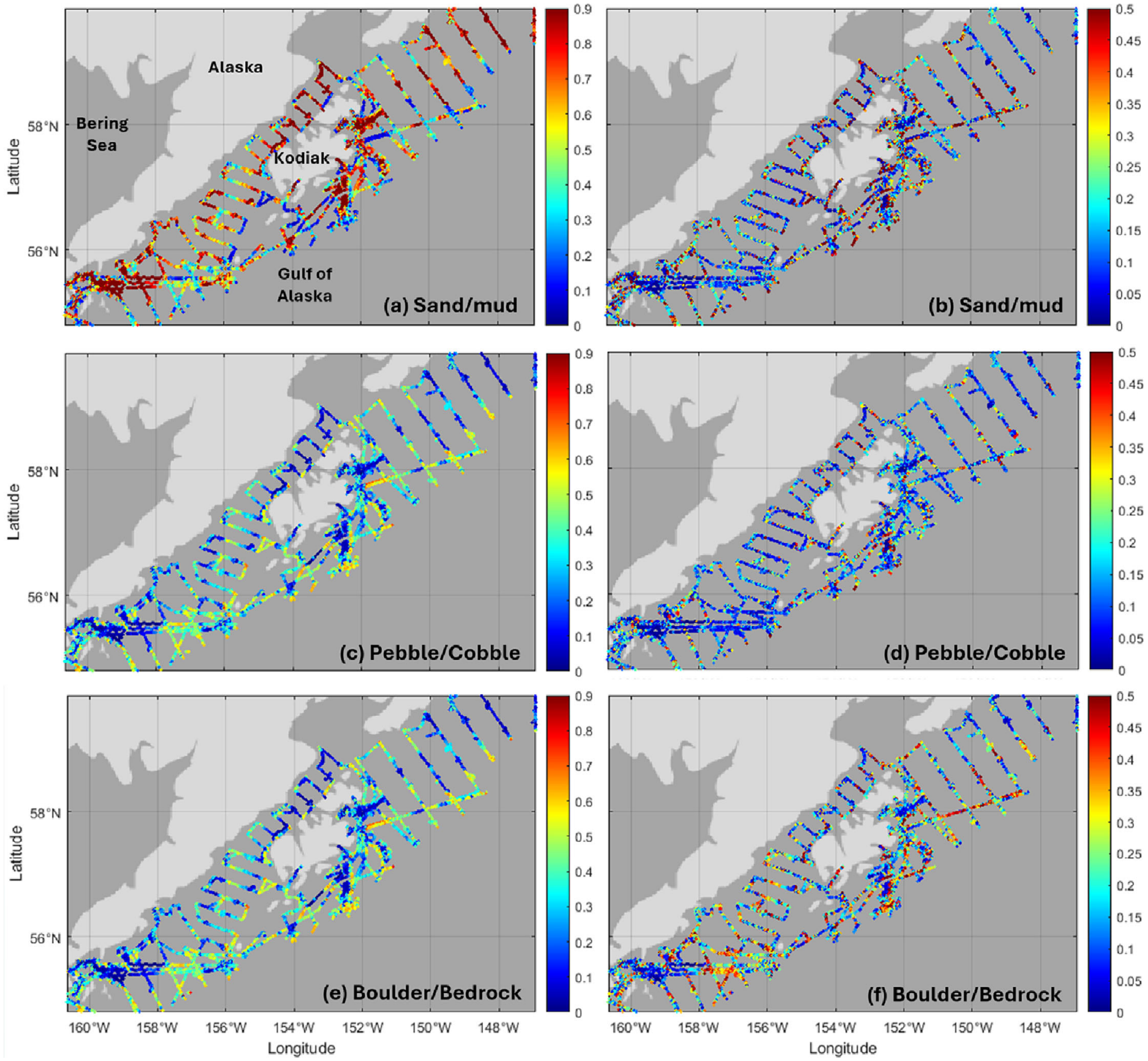

Fig. 8. Seabed composition maps of the Shelikof Strait and surrounding regions for 1 km^2 grids. (**a**, **c**, **e**) Distance-weighted average proportion of substrates, and (**b**, **d**, **f**) the variance in proportion of substrates within grids for (a, b) sand/mud, (c, d) pebble/cobble, and (e, f) boulder/bedrock.

frequencies, and the difference, as an indicator of scatterer size and shape (De Robertis et al. 2010). This result is important as it allows for more general application of the approach to a broader class of acoustic data. However, both the full and reduced frequency models struggle with classification of cobble and boulder due to limited representation as well as difficulty in differentiating those signals when compared to pebbles or bedrock. While the variations of sandy and rocky habitats are well-represented in training and test datasets, sufficient observations of mostly cobble habitat were not captured and therefore may be difficult for the model to detect. Still, averaging over several pings and frames improves the reliability of the data. Assimilation of data sources such as sediment grabs that quantify the grain sizes can help future models with differentiating between sand and mud or gravel and pebble, where such fine-scale habitat differences are important; for example, for Pacific sand lance studies (Baker et al. 2024).

The current ML model is relatively small (0.86 million parameters for the full model; Fig. 2), and allows real-time processing, which can be incorporated into survey operations. Having a small model as well as inclusion of a dropout layer minimizes the risk of overfitted and ungeneralizable models. The development of tools such as gradcam (Selvaraju et al. 2017) to parse out the learned features of the model is a very promising development and can be included in future models. Alternative input architecture such as treating the transducer frequencies as separate channels could be beneficial. Transformer networks for contextual learning have been successful for language models and may be able to learn associations when applied to acoustic data, particularly when large temporal sequences are available to model behavior (Måløy 2020). Hybrid CNN-transformer models such as CoATNet are also promising for improving generalizations, particularly for larger model capacities (Dai et al. 2021). On the other hand, to adapt the model for other cases with limited data, semi-supervised classification that relies on both clustering and classification may also be used (Choi et al. 2021). Inclusion of additional acoustic data into the inputs such as the phase angles of the acoustic backscatter and bathymetric slope may also improve performance. As the vessel moves, the transducers may no longer send and receive signals orthogonal to the seafloor, depending on the sea state, which could potentially bias the data (Jech et al. 2021). While ML models may learn the uncertainty in the data due to vessel motion, providing the measured roll, pitch and yaw as additional inputs to better account for motion is a possibility.

The present analysis utilizes existing survey data, which have been collected over several years. In most cases, the acoustic tracklines are slightly offset for each survey (Supporting Information Fig. S6), with extensive acoustic coverage of the Gulf of Alaska. The agreement of the results from two different surveys is encouraging. The direct inclusion of depth allows the model to resolve any correlations between habitat types and bottom depth; for example, bedrock is identified in both deep and shallow locations. The resulting substrate maps appear to correctly identify known geographic ocean features, such as the known sandy areas of the continental shelf troughs of Chiniak, Barnabas, and Marmot Bays (Fig. 6). The maps also appeared to show some less known habitat features, such as the harder substrates along the western portion of the central Shelikof Strait. For the GOA, softer sediments were more homogenous over larger areas compared to rocky habitat, which were more likely to co-occur with gravel (Fig. 7). The high-resolution maps also demonstrate the inability of coarser resolution maps to capture the fine-scale heterogeneity in compositions, particularly for rocky habitats. The description of co-occurrence of substrate types and their spatial distribution and correlations can inform development of methods for spatial interpolation to unsampled areas in GOA. Additional data may also be assimilated from other vessels, such as fishing vessels operating in the region. However, in addition to operating fewer frequencies, these data are not usually calibrated; thus, the model will have to be modified based on transfer learning and semi-supervised methods, referenced before. However, since the model is not just a function of absolute intensity thresholds but is dependent on gradients in the data, this approach may be successful.

The results of this study, and resulting mapping products, will have immediate applications for a variety of research activities in the Gulf of Alaska, such as essential fish habitat studies, species distribution modeling, as well as bottom trawl survey planning. The combination of high-quality habitat maps and localized species abundance provides the basic inputs needed to define essential fish habitats for effective management and conservation of commercially important fish species. Precise, high spatial resolution habitat maps would also improve species distribution models as it will allow for local variations in habitat composition to be more closely related to species density data. The ability to identify, with high confidence, previously unsampled areas that are not amenable to bottom trawl survey gear would allow more efficient effort allocation for bottom trawl surveys. Trawlability could also inform survey gear comparison models as a covariate, or to define and predict sampling frames for different gears. The methods and training sets described here might be directly applied to other regions with similar general substrate types, such as the Pacific west coast of the United States and possibly northeast Atlantic regions. Areas consisting of more complex habitat types would require additional or separate training sets, but the general approach of matching acoustics and image data would be applicable to a wide array of habitat identification and classification needs.

Author Contributions

Karuna Agarwal: Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review and editing. Kresimir Williams: Conceptualization, Funding acquisition, Project administration, Data curation, Writing – review and editing. Chris Rooper: Funding acquisition, Project administration, Writing – review and editing.

Acknowledgments

Funding for this project was provided by an Alaska Fisheries Science Center Essential Fish Habitat internal grant. We are greatly indebted to Pamela Goddard, Vanessa Lowe, and Kimberly Rand for image annotation, Rick Towler for camera technology development, Sandy Parker-Stetter for administrative assistance, and the Midwater Assessment and Conservation Engineering Program for the acoustic and optic data collections. Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

Conflicts of Interest

None declared.

Data Availability Statement

The raw acoustic datasets collected by Alaska Fisheries Science Center are publicly available and the image data available upon request from the NOAA NCEI program. The trained model and the code to construct the model inputs and obtain the classification are available through the Zenodo open repository at https://doi.org/10.5281/zenodo.18462457. In addition, model outputs are publicly available under the following DOI's: acoustic seafloor classification model output GOA 2017 (https://doi.org/10.5281/zenodo.17612923), acoustic seafloor classification model output GOA 2019 (https://doi.org/10.5281/zenodo.17612149). A second data set that contains the outputs aggregated over 1 km grids is also available at Gulf of Alaska proportional seafloor habitat classification with 1 km resolution (https://doi.org/10.5281/zenodo.17663313).

References

Akhtar, N., and U. Ragavendran. 2020. "Interpretation of Intelligence in CNN-Pooling Processes: A Methodological Survey." *Neural Computing and Applications* 32, no. 3: 879–898. https://doi.org/10.1007/s00521-019-04296-5.

Anderson, J. T., D. Van Holliday, R. Kloser, D. G. Reid, and Y. Simard. 2008. "Acoustic Seabed Classification: Current Practice and Future Directions." *ICES Journal of Marine Science* 65, no. 6: 1004–1011. https://doi.org/10.1093/icesjms/fsn061.

Baker, M. R., H. G. Greene, J. Aschoff, et al. 2024. "Atlas of Pacific Sand Lance (*Ammodytes personatus*) Benthic Habitat—Application of Multibeam Acoustics and Directed Sampling to Identify Viable Subtidal Substrates." *Marine Environmental Research* 202: 106778. https://doi.org/10.1016/j.marenvres.2024.106778.

Baker, M. R., W. Palsson, M. Zimmermann, and C. N. Rooper. 2019. "Model of Trawlable Area Using Benthic Terrain and Oceanographic Variables—Informing Survey Design and Habitat Maps in the Gulf of Alaska." *Fisheries Oceanography* 28, no. 6: 629–657. https://doi.org/10.1111/fog.12442.

Berg, S., D. Kutra, T. Kroeger, et al. 2019. "Ilastik: Interactive Machine Learning for (Bio) Image Analysis." *Nature Methods* 16, no. 12: 1226–1232.

Björk, G., M. Jakobsson, K. Assmann, et al. 2018. "Bathymetry and Oceanic Flow Structure at Two Deep Passages Crossing the Lomonosov Ridge." *Ocean Science* 14, no. 1: 1–13. https://doi.org/10.5194/os-14-1-2018.

Boggild, K., M. Z. Li, J. B. Eamer, and C. D. Stacey. 2024. "Geomorphology and Sediment Mobility on Sand Banks: A Study of Dogfish Bank, Hecate Strait, Northeast Pacific Ocean." *Earth Surface Processes and Landforms* 49, no. 13: 4485–4505. https://doi.org/10.1002/esp.5925.

Borland, H. P., B. L. Gilby, C. J. Henderson, et al. 2021. "The Influence of Seafloor Terrain on Fish and Fisheries: A Global Synthesis." *Fish and Fisheries* 22, no. 4: 707–734. https://doi.org/10.1111/faf.12546.

Brautaset, O., A. U. Waldeland, E. Johnsen, et al. 2020. "Acoustic Classification in Multifrequency Echosounder Data Using Deep Convolutional Neural Networks." *ICES Journal of Marine Science* 77, no. 4: 1391–1400. https://doi.org/10.1093/icesjms/fsz235.

Bryan, D. R., K. Williams, C. A. O'Leary, and C. N. Rooper. 2023. "The Design of a Camera-Based Fisheries-Independent Survey for Untrawlable Habitat in the Gulf of Alaska." *Fisheries Research* 266: 106785. https://doi.org/10.1016/j.fishres.2023.106785.

Choi, C., M. Kampffmeyer, N. O. Handegard, et al. 2021. "Semi-Supervised Target Classification in Multi-Frequency Echosounder Data." *ICES Journal of Marine Science* 78, no. 7: 2615–2627. https://doi.org/10.1093/icesjms/fsab140.

Dai, Z., H. Liu, Q. V. Le, and M. Tan. 2021. "CoAtNet: Marrying Convolution and Attention for all Data Sizes." *Advances in Neural Information Processing Systems* 34: 3965–3977. https://doi.org/10.48550/arxiv.2106.04803.

De Robertis, A., D. R. McKelvey, and P. H. Ressler. 2010. "Development and Application of an Empirical Multi-frequency Method for Backscatter Classification." *Canadian Journal of Fisheries and Aquatic Sciences* 67, no. 9: 1459–1474. https://doi.org/10.1139/F10-075.

Ferrini, V. L., and R. D. Flood. 2006. "The Effects of Fine-Scale Surface Roughness and Grain Size on 300 kHz Multibeam Backscatter Intensity in Sandy Marine Sedimentary Environments." *Marine Geology* 228, no. 1–4: 153–172. https://doi.org/10.1016/j.margeo.2005.11.010.

Gibson, G. A., M. R. Baker, W. T. Stockhausen, et al. 2022. "Modeling in an Integrated Ecosystem Research Framework to Explore Recruitment in Gulf of Alaska Groundfish—Applications to Management and Lessons Learned." *Deep Sea Research Part II: Topical Studies in Oceanography* 197: 105048. https://doi.org/10.1016/j.dsr2.2022.105048.

Greene, H. G., M. M. Yoklavich, R. M. Starr, et al. 1999. "A Classification Scheme for Deep Seafloor Habitats." *Oceanologica Acta* 22, no. 6: 663–678. https://doi.org/10.1016/S0399-1784(00)88957-4.

Hosna, A., E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim. 2022. "Transfer Learning: A Friendly Introduction." *Journal of Big Data* 9, no. 1: 102. https://doi.org/10.1186/s40537-022-00652-w.

Jech, J. M., M. Schaber, M. Cox, et al. 2021. "*Collecting Quality Echosounder Data in Inclement Weather*." ICES Cooperative Research Report, 352.

Jogin, M., M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva. 2018. "Feature Extraction Using Convolution Neural Networks (CNN) and Deep Learning." In 3rd IEEE

International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2319–2323. IEEE. https://doi.org/10.1109/RTEICT42901.2018.9012507.

Jones, D. T., C. N. Rooper, C. D. Wilson, P. D. Spencer, D. H. Hanselman, and R. E. Wilborn. 2021. "Estimates of Availability and Catchability for Select Rockfish Species Based on Acoustic-Optic Surveys in the Gulf of Alaska." *Fisheries Research* 236: 105848. https://doi.org/10.1016/j.fishres.2020.105848.

Jones, D. T., M. Levine, K. Williams, A. De Robertis, and P. Ressler. 2022. Results of the Acoustic-Trawl Survey of Walleye Pollock (Gadus chalcogrammus) in the Gulf of Alaska, May–August 2019 (DY2019-06). (AFSC Processed Rep. 2022-07). Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv.

Laman, E. A., C. N. Rooper, K. Turner, S. Rooney, D. W. Cooper, and M. Zimmermann. 2018. "Using Species Distribution Models to Describe Essential Fish Habitat in Alaska." *Canadian Journal of Fisheries and Aquatic Sciences* 75, no. 8: 1230–1255. https://doi.org/10.1139/cjfas-2017-0181.

Lecours, V., M. F. Dolan, A. Micallef, and V. L. Lucieer. 2016. "A Review of Marine Geomorphometry, the Quantitative Study of the Seafloor." *Hydrology and Earth System Sciences* 20, no. 8: 3207–3244. https://doi.org/10.5194/hess-20-3207-2016.

Levin, P. S., and G. W. Stunz. 2005. "Habitat Triage for Exploited Fishes: Can We Identify Essential "Essential Fish Habitat?"." *Estuarine, Coastal and Shelf Science* 64, no. 1: 70–78. https://doi.org/10.1016/j.ecss.2005.02.007.

Lin, G., and W. Shen. 2018. "Research on Convolutional Neural Network Based on Improved Relu Piecewise Activation Function." *Procedia Computer Science* 131: 977–984. https://doi.org/10.1016/j.procs.2018.04.239.

Lucieer, V., N. A. Hill, N. S. Barrett, and S. Nichol. 2013. "Do Marine Substrates 'Look' and 'Sound' the Same? Supervised Classification of Multibeam Acoustic Data Using Autonomous Underwater Vehicle Images." *Estuarine, Coastal and Shelf Science* 117: 94–106. https://doi.org/10.1016/j.ecss.2012.11.001.

MacLennan, D. N., P. G. Fernandes, and J. Dalen. 2002. "A Consistent Approach to Definitions and Symbols in Fisheries Acoustics." *ICES Journal of Marine Science* 59, no. 2: 365–369. https://doi.org/10.1006/jmsc.2001.1158.

Måløy, H. 2020. "Echobert: A Transformer-Based Approach for Behavior Detection in Echograms." *IEEE Access* 8: 218372–218385. https://doi.org/10.1109/ACCESS.2020.3042337.

Mayer, L., M. Jakobsson, G. Allen, et al. 2018. "The Nippon Foundation—GEBCO Seabed 2030 Project: The Quest to See the World's Oceans Completely Mapped by 2030." *Geosciences* 8, no. 2: 63. https://doi.org/10.3390/geosciences8020063.

Olden, J. D., J. J. Lawler, and N. L. Poff. 2008. "Machine Learning Methods Without Tears: A Primer for Ecologists." *The Quarterly Review of Biology* 83, no. 2: 171–193. https://doi.org/10.1086/587826.

Ordonez, A., I. Utseth, O. Brautaset, R. Korneliussen, and N. O. Handegard. 2022. "Evaluation of Echosounder Data Preparation Strategies for Modern Machine Learning Models." *Fisheries Research* 254: 106411. https://doi.org/10.1016/j.fishres.2022.106411.

Paszke, A., S. Gross, F. Massa, et al. 2019. "Pytorch: An Imperative Style, High-Performance Deep Learning Library." *Advances in Neural Information Processing Systems* 32.

Pirtle, J. L., E. A. Laman, J. Harris, et al. 2023. Advancing Model-Based Essential Fish Habitat Descriptions for North Pacific Species in the Gulf of Alaska (NOAA Tech. Memo. NMFS-AFSC-468). U.S. Dep. Commer.

Pirtle, J. L., S. K. Shotwell, M. Zimmermann, J. A. Reid, and N. Golden. 2019. "Habitat Suitability Models for Groundfish in the Gulf of Alaska." *Deep Sea Research Part II: Topical Studies in Oceanography* 165: 303–321. https://doi.org/10.1016/j.dsr2.2017.12.005.

Pirtle, J. L., T. C. Weber, C. D. Wilson, and C. N. Rooper. 2015. "Assessment of Trawlable and Untrawlable Seafloor Using Multibeam-Derived Metrics." *Methods in Oceanography* 12: 18–35. https://doi.org/10.1016/j.mio.2015.06.001.

Porskamp, P., A. C. G. Schimel, M. Young, A. Rattray, Y. Ladroit, and D. Ierodiaconou. 2022. "Integrating Multibeam Echosounder Water-Column Data Into Benthic Habitat Mapping." *Limnology and Oceanography* 67: 1701–1713. https://doi.org/10.1002/lno.12160.

Razavian, A., H. Azizpour, J. Sullivan, and S. Carlsson. 2014. "CNN Features Off-The-Shelf: An Astounding Baseline for Recognition." In 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 806–813. Columbus, OH, USA: IEEE. https://doi.org/10.1109/CVPRW.2014.131.

Rosenberg, A., T. E. Bigford, S. Leathery, R. L. Hill, and K. Bickers. 2000. "Ecosystem Approaches to Fishery Management Through Essential Fish Habitat." *Bulletin of Marine Science* 66, no. 3: 535–542.

Rubbens, P., S. Brodie, T. Cordier, et al. 2023. "Machine Learning in Marine Ecology: An Overview of Techniques and Applications." *ICES Journal of Marine Science* 80, no. 7: 1829–1853. https://doi.org/10.1093/icesjms/fsad100.

Selvaraju, R. R., M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. 2017. "Grad-Cam: Visual Explanations from Deep Networks Via Gradient-Based Localization." *International Journal of Computer Vision* 128: 336–359. https://doi.org/10.1007/s11263-019-01228-7.

Senjaliya, F., M. Cote, A. Dash, et al. 2024. "Deep Learning-Based Identification of Arctic Ocean Boundaries and Near-Surface Phenomena in Underwater Echograms." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2977–2986. Seattle, WA, USA: IEEE. https://doi.org/10.1109/CVPRW63382.2024.00303.

Stephens, D., and M. Diesing. 2014. "A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data." *PLoS One* 9, no. 4: e93950.

Shi, J. 1994. "Good Features to Track." In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 593–600. Seattle, WA, USA: IEEE. https://doi.org/10.1109/CVPR.1994.323794.

Simmonds, E. J., N. J. Williamson, F. Gerlotto, and A. Aglen. 1992. "Acoustic Survey Design and Analysis Procedure: A Comprehensive Review of Current Practice." *ICES Cooperative Research Reports* 187: 136. https://doi.org/10.17895/ices.pub.4606.

Simmonds, J., and D. N. MacLennan. 2008. Fisheries Acoustics: Theory and Practice. John Wiley & Sons.

Siple, M. C., P. G. von Szalay, N. W. Raring, A. N. Dowlin, and B. C. Riggle. 2024. Data Report: 2023 Gulf of Alaska Bottom Trawl Survey. (AFSC Processed Rep. 2024-09). Seattle WA 98115: Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv.

Siwabessy, P. J. W., J. D. Penrose, R. J. Kloser, and D. R. Fox. 1999. "Seabed Habitat Classification." In Proceedings of the International Conference on HighResolution Surveys in Shallow Waters DSTO, 1–9. Sydney, Australia.

Siwicke, K., and P. Malecha. 2024. The 2023 longline survey of the Gulf of Alaska and eastern Bering Sea on the FV Alaskan Leader: Cruise Report AL-23-01 (NOAA Technical Memorandum NMFS-AFSC-480). U.S. Department of Commerce.

Slonimer, A. L., S. E. Dosso, A. B. Albu, et al. 2023. "Classification of Herring, Salmon, and Bubbles in Multifrequency Echograms Using U-Net Neural Networks." *IEEE Journal of Oceanic Engineering* 48, no. 4: 1236–1254. https://doi.org/10.1109/JOE.2023.3272393.

Stienessen, S. C., C. N. Rooper, T. C. Webe, D. T. Jones, J. L. Pirtle, and C. D. Wilson. 2021. "Comparison of Model Types for Prediction of Seafloor Trawlability in the Gulf of Alaska by Using Multibeam Sonar Data." *Fishery Bulletin* 119: 184–196. https://doi.org/10.7755/FB.119.2-3.7.

Stokesbury, K. D., N. D. Bethoney, F. Restrepo, and B. P. Harris. 2024. "Anticipating the Winds of Change: A Baseline Assessment of Northeastern US Continental Shelf Surficial Substrates." *Fisheries Oceanography* 35: e12693. https://doi.org/10.1111/fog.12693.

Swanborn, D. J., V. A. Huvenne, S. J. Pittman, and L. C. Woodall. 2022. "Bringing Seascape Ecology to the Deep Seabed: A Review and Framework for Its Application." *Limnology and Oceanography* 67, no. 1: 66–88. https://doi.org/10.1002/lno.11976.

von Szalay, P. G., and D. A. Somerton. 2017. "A Method for Predicting Trawlability in the Gulf of Alaska With the Use of Calibrated, Split-Beam, Echosounder Backscatter." *Fishery Bulletin* 115, no. 4: 496–504.

Wall, C. C., R. Towler, C. Anderson, R. Cutter, and J. M. Jech. 2018. "PyEcholab: An Open-Source, Python-Based Toolkit to Analyze Water-Column Echosounder Data." *The Journal of the Acoustical Society of America* 144, no. 3_Supplement: 1778. https://doi.org/10.1121/1.5067860.

Wang, Z., and L. Wu. 2023. "Theoretical Analysis of the Inductive Biases in Deep Convolutional Networks." *Advances in Neural Information Processing Systems* 36: 74289–74338. https://doi.org/10.48550/arXiv.2305.08404.

Weber, T. C., C. Rooper, J. L. Butler, D. Jones, and C. Wilson. 2013. "Seabed Classification for Trawlability Determined With a Multibeam Echo Sounder on Snakehead Bank in the Gulf of Alaska." *Fishery Bulletin* 111: 68–77. https://doi.org/10.7755/fb.111.1.6.

Williams, K., C. N. Rooper, and R. Towler. 2010. "Use of Stereo Camera Systems for Assessment of Rockfish Abundance in Untrawlable Areas and for Recording Pollock Behavior During Midwater Trawls." *Fishery Bulletin* 108, no. 3: 352–362.

Wu, H., and X. Gu. 2015. "Towards Dropout Training for Convolutional Neural Networks." *Neural Networks* 71: 1–10. https://doi.org/10.1016/j.neunet.2015.07.007.

Wu, Y., and K. He. 2018. "Group Normalization." In Proceedings of the European Conference on Computer Vision (ECCV). 15th European Conference, 3–19. Munich, Germany: Springer.

Zhang, Y., C. C. Wall, J. M. Jech, and Q. Lv. 2024. "Developing a Hybrid Model With Multiview Learning for Acoustic Classification of Atlantic Herring Schools." *Limnology and Oceanography: Methods* 22, no. 5: 351–368. https://doi.org/10.1002/lom3.10611.

Zhang, Z. 1999. "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations." In Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 1, 666–673. Kerkyra, Greece: IEEE. https://doi.org/10.1109/ICCV.1999.791289.

Zimmermann, M. 2003. "Calculation of Untrawlable Areas Within the Boundaries of a Bottom Trawl Survey." *Canadian Journal of Fisheries and Aquatic Sciences* 60, no. 6: 657–669. https://doi.org/10.1139/f03-055.

Zimmermann, M., and M. M. Prescott. 2023. "Towards Mapping the Gaps in Alaska." *The International Hydrographic Review* 29, no. 2: 28–45. https://doi.org/10.58440/ihr-29-2-a22.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Submitted 16 June 2025
Revised 04 February 2026
Accepted 06 February 2026