Remote Sensing of Environment 239 (2020) 111617

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

SealNet: A fully-automated pack-ice seal detection pipeline for sub-meter satellite imagery

B.C. Gonçalves*, B. Spitzbart, H.J. Lynch

610 Life Sciences Building, Department of Ecology and Evolution, Stony Brook, NY 11777, USA

ARTICLE INFO

Edited by Menghua Wang

Keywords:
Crabeater seal
Lobodon carcinophaga
Weddell seal
Leptonychotes weddellii
APIS
CCAMLR
Deep learning
Object detection
Segmentation
Very high resolution

ABSTRACT

Antarctic pack-ice seals, a group of four species of true seals (Phocidae), play a pivotal role in the Southern Ocean foodweb as wide-ranging predators of Antarctic krill (*Euphausia superba*). Due to their circumpolar distribution and the remoteness and vastness of their habitat, little is known about their population sizes. Estimating pack-ice seal population sizes and trends is key to understanding how the Southern Ocean ecosystem will react to threats such as climate change driven sea ice loss and krill fishing. We present a functional pack-ice seal detection pipeline using Worldview-3 imagery and a Convolutional Neural Network that counts and locates seal centroids. We propose a new CNN architecture that detects objects by combining semantic segmentation heatmaps with binary classification and counting by regression. Our pipeline locates over 30% of seals, when compared to consensus counts from human experts, and reduces the time required for seal detection by 95% (assuming just a single GPU). While larger training sets and continued algorithm development will no doubt improve classification accuracy, our pipeline, which can be easily adapted for other large-bodied animals visible in sub-meter satellite imagery, demonstrates the potential for machine learning to vastly expand our capacity for regular pack-ice seal surveys and, in doing so, will contribute to ongoing international efforts to monitor pack-ice seals.

1. Introduction

The Southern Ocean (SO) harbors major seasonal hotspots for primary productivity (Arrigo and Dijken, 2003). The cold, nutrient rich waters of the SO play a fundamental role regulating global climate, both by absorbing large amounts of heat and sinking fixed carbon (Frölicher et al., 2015; Morrison et al., 2016). Food webs in the SO are trophically shallow (Clarke, 1985), but they more than compensate in terms of biomass, sustaining massive concentrations of phytoplankton consumers (Bester et al., 2002a; Nowacek et al., 2011). Among these consumers, a small crustacean, Antarctic krill (*Euphasia superba*), is especially important; krill is the main food item for a wide range of upper tier consumers, from fish and penguins to seals and whales, and serves as a fundamental link between predators and primary producers. Due to krill's role in the SO food web, assessing and tracking Antarctic krill stocks is central to Antarctic ecology. This is especially true now that climate change (Flores et al., 2012; Klein et al., 2018), ocean acidification (Kawaguchi et al., 2013) and krill fisheries (Forcada et al., 2012) threaten to shift the abundance and distribution of this key Antarctic species. Challenging our efforts to track Antarctic krill, however, is its small size and patchy distribution (Voronina, 1998). One way to circumvent those difficulties is to use krill predator abundances as a proxy for krill distribution (Huang et al., 2011). Antarctic pack-ice seals (crabeater seals [*Lobodon carcinophaga*], Weddell seals [*Leptonychotes weddelli*], leopard seals [*Hydrurga leptonyx*] and Ross seals [*Omnatophoca rossii*], within the Phocidae family), as a group, represent a promising vehicle to gauge krill stocks for they are not only key krill consumers (Botta et al., 2018; Forcada et al., 2012; Hückstädt et al., 2012; Siniff and Stone, 1985) but they are also large enough to be individually spotted with high spatial resolution satellite imagery.

The potential of pack-ice seals as indicators of environmental health in the SO has not gone unnoticed; polar ecologists have channeled sizeable efforts into estimating pack-ice seal population sizes, the most notable of these attempts being the Antarctic pack-ice seal (APIS) project (Anonymous, 1997), a joint effort of six countries to estimate Antarctic seal populations using aerial surveys (Ackley et al., 2006). Conducting such large-scale aerial survey programs in Antarctica is extremely expensive, necessarily requiring extensive collaboration among Antarctic national programs. Fortunately, very high spatial resolution (VHR) satellite imagery may soon be a viable alternative for aerial surveys, providing greater spatial coverage and, due to its dramatically lower cost, increased repeatability. The use of VHR satellite

* Corresponding author.
E-mail address: bento.goncalves@stonybrook.edu (B.C. Gonçalves).

https://doi.org/10.1016/j.rse.2019.111617
Received 13 February 2019; Received in revised form 17 December 2019; Accepted 17 December 2019
0034-4257/ Published by Elsevier Inc.

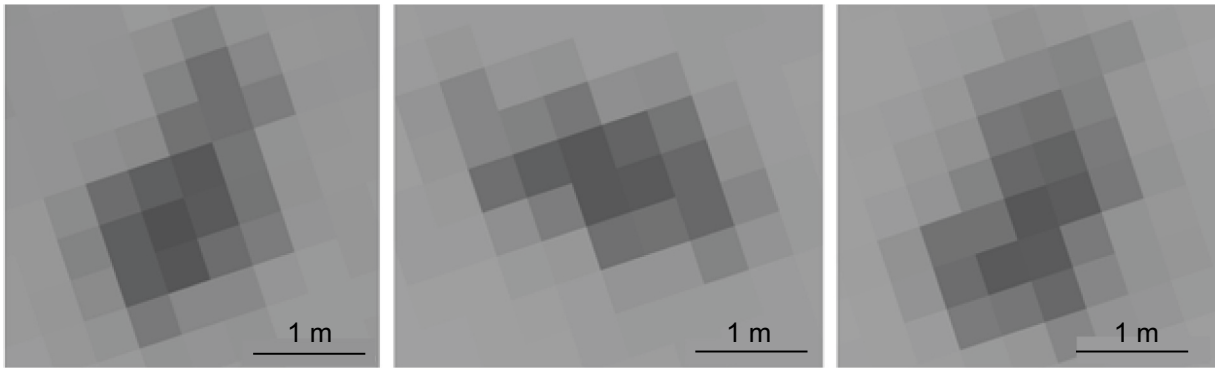

Fig. 1. Individual pack-ice seals viewed in panchromatic WV03 imagery. Scale bar shows one meter, with a typical seal about 2.5 m long covering 20–30 pixels (in total). All three seals in this image were extracted from the same scene, but resolution may change depending on the angle at which the image was captured by the WV03 sensor.
Satellite imagery copyright DigitalGlobe, Inc. 2019.

imagery for wildlife survey has exploded in recent years, and includes demonstration projects for southern elephant seals (McMahon et al., 2014), polar bears (Stapleton et al., 2014) and African ungulates (Xue et al., 2017; Yang et al., 2014), as well as seabird species whose presence and abundance can be estimated indirectly using the guano stain at the colony (LaRue et al., 2014; Lynch et al., 2012). Pack-ice seals, while large enough to be seen in VHR imagery, are particularly hard to detect since their preferred haul out environment (pack ice; Bengtson and Stewart, 1992; Lake et al., 1997) changes on short (hourly) and long (seasonal) time scales and the information content of each individual seal in an image is exceptionally low (Fig. 1).

Though it is possible to find seal-sized objects in VHR imagery manually, this laborious approach is only feasible at local scales (e.g., LaRue et al., 2011), introduces observer biases (Dickinson et al., 2010), and is not easily scaled to allow annotation of every high spatial resolution image captured within the range of pack-ice seals. Thus, repeatable, large scale wildlife surveys require automated detection systems (Conn et al., 2014). Traditional pixel or object-based methods for remote sensing scene understanding (RSISU) (e.g. Koju et al., 2018; McNabb et al., 2016), perhaps due to their reliance on hand-crafted features and spectral signatures, struggle at the increased granularity posed by high spatial resolution satellite imagery. As is the case for many fields such as computer vision (Voulodimos et al., 2018) and natural language processing (Do et al., 2019), deep learning, in the specific flavor of Convolutional Neural Networks (CNNs), are now the state-of-the art for RSISU (Gu et al., 2019), and is likely our best candidate for automated seal detection in high spatial resolution imagery. CNNs work by learning a series of convolution kernels – analogous to image processing kernels – as they learn to map inputs in the training data to their corresponding labels. CNNs have now been successfully employed in many ecological settings such as identifying whales (Borowicz et al., 2019; Polzounov et al., 2016), finding mammals in the African Savanna with UAV imagery (Kellenberger et al., 2018) and classifying animals in camera trap pictures (Norouzzadeh et al., 2018).

In this work, we explore the viability of CNNs to locate pack-ice seals in Antarctica and the scalability of this approach, with the ultimate goal of facilitating continental-scale population counts for pack-ice seals and other large bodied animals. Like many other wildlife detection sampling schemes (Kellenberger et al., 2018; Xue et al., 2017), however, the vast majority of the VHR imagery contains no true positives (i.e. seals), creating the potential for significant false positives even if the false positive rate is low. We propose a seal detection pipeline that i) determines whether a portion of the image is occupied by seals; ii) counts seals in that portion of the image and; iii) locates the centroid of each identified seal. All of the above is performed in a single pass with our proposed CNN architecture, SealNet.[1] In our validation and test sets, this approach is superior to pure regression or semantic segmentation approaches.

2. Methods

2.1. Selecting imagery

For this pipeline, we use Worldview 3 (WV03) imagery provided by DigitalGlobe, Inc., which has the highest available resolution for commercial imagery with a 0.3 m resolution at nadir in panchromatic imagery and 1.5 m with 16 multispectral bands (Red, Green, Blue, Red Edge, Coastal, Yellow and 2 near-infrared bands). Only the panchromatic band was used for this work because individual seals are difficult to spot at lower resolutions and because the color information is not highly informative for classification (at least for human interpreters). Due to GPU memory limitations imposed by our CNN architecture, we subdivide WV03 scenes into 224 × 224 pixel images (hereafter 'patches') (Fig. 2). Prior to prediction, each WV03 scene is split into approximately 500,000 patches, keeping a 75% overlap between neighboring patches to ensure corners are not overlooked by the CNN classifier.

2.2. Building a training set[2]

A training set with 75,718 raw training samples was manually assembled to train seal detection CNNs. Raw training samples are generated by extracting 450 × 450 pixel images (hereafter 'raw training patches'), roughly covering two hectares, at predefined locations (i.e. training points) on a total of 34 WV03 scenes (Fig. 3) selected from the Polar Geospatial Center catalog. Training points were annotated by visually inspecting WV03 scenes for locations with the 10 following features: seals on pack ice, seals on fast ice, emperor penguin colonies, marching emperor penguins, cracks on the sea ice, glaciers, fast-ice, pack ice, rock outcrops and open water. For the last seven categories (background/non-target), we place an array of equidistant training points, separated by 100 m, over areas where a particular class was predominant, removing any occasional points that did not fit into that class. For emperor penguin colony points, we covered colonies with a similar array of equidistant points, with a 10 m distance between neighboring points. Groups of three or more emperor penguins arranged in lines were labeled as marching emperors, with training point

[1] https://github.com/iceberg-project/Seals/tree/paper/SealNet_code.
[2] https://github.com/iceberg-project/Seals/blob/paper/SealNet_code/training_set_generation.ipynb.

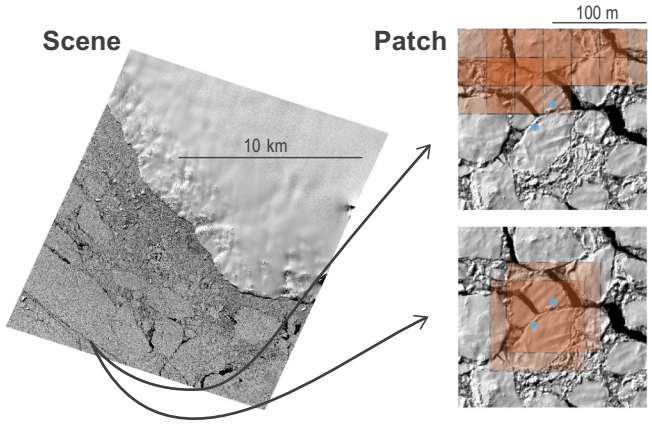

Fig. 2. Sampling scheme. WV03 scenes are split into smaller 'patches' to fit CNN requirements for input size. To create a training set, 450 × 450 training patches are extracted around features of interest (light-blue circles) on a scene, which may overlap depending on how close features are and CNN input size (orange squares on the bottom right). For prediction, whole scenes get chopped into 224 × 224 pixel patches using a sliding-window approach, with a stride that keeps a 75% overlap on both *x* and *y* axes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Satellite imagery copyright DigitalGlobe, Inc. 2019.

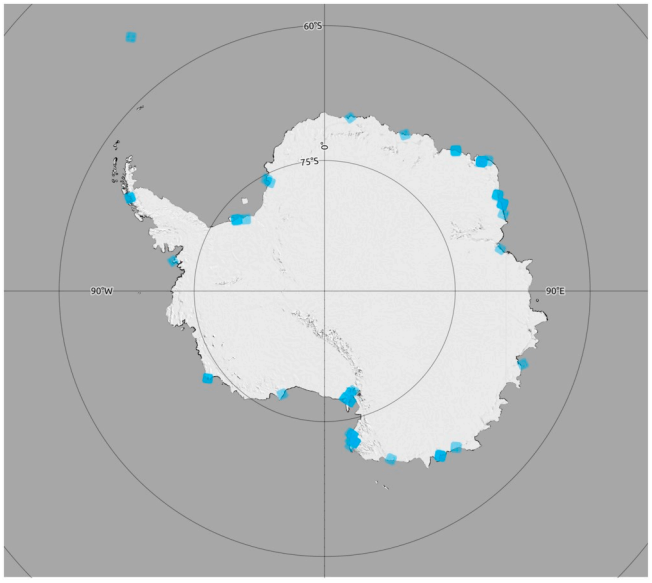

Fig. 3. Training set scenes. Locations of the 52 WorldView-3 scenes used on the training set are marked with light-blue squares. Scenes with spatial overlap were captured at different times. Training set scenes range from October 2014 to February 2017. The scarcity of offshore scenes in our training set reflects the preponderance of coastal scenes on available WV03 imagery. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Antarctic basemap extracted from Quantarctica (Matsuoka et al., 2018).

annotations centered on one of the penguins. Since crabeater seals, Ross seals and leopard seals are confined to pack-ice habitat (Ballard et al., 2012; Bengtson et al., 2001; Southwell et al., 2012), and the first species is far more abundant than the latter two (Southwell et al., 2008b, 2008c, 2008a), every seal on pack-ice in our training is assumed to be a crabeater seal. Seals on fast ice are assumed to be Weddell seals since that is the only one of the four species that is strongly associated with fast ice habitat (Bengtson et al., 2001; Southwell et al., 2012). To reduce our annotation effort, our seal training points – both Weddell and crabeater – consist of a single point, placed at the centroid of each seal. When generating seal training images, we include the location of seal centroids within those images along with the image itself – necessary to derive ground truth seal locations and counts within training patches. Finally, our seal detection CNN, trained on the training set described above, was deployed on 18 new scenes, where seals could not be found upon visual inspection, generating a total of 10,766 training points, which were then added to the training set as a separate class for seal-shaped shadows. To evaluate and select models during training, our training data was split into training and validation sets. To prevent spatial overlap between training and validation images, we split entire groups of seals between training and validation, keeping roughly 90% of the seal training points for training and the remaining for validation. Background class training points were split by scene, where each scene with training points for a given background class is either used for training or validation. Background scenes were also split to keep roughly 90% of the training points for training and the remaining 10% for validation.

2.3. Setting up the convolutional neural network

Our seal detection pipeline (Fig. 4) detects seal centroids in VHR imagery following 4 steps: 1) tile input scene into 'patches' that can be classified using the CNN; 2) run each patch through the CNN to get a probability of harboring one or more seals (occupancy probability), a seal count, and a seal centroid intensity heatmap; 3) remove predictions below a predefined occupancy probability threshold; and 4) find the *n* greatest intensity peaks in the heatmap, where *n* is the seal count. Our proposed CNN architecture was assembled by adding two branches to the U-Net (Ronneberger et al., 2015) architecture: a branch for occupancy, branching out of the second U-Net max-pooling layer, based on the DenseNet (Huang et al., 2017) architecture and a branch for counting, branching out of the fourth U-Net max-pooling layer, based on the WideResnet architecture (Zagoruyko and Komodakis, 2016). Apart from the regular intensity heatmap output from U-Net, our architecture also outputs an occupancy probability and a seal count. All CNNs used here were implemented in PyTorch (Paszke et al., 2017).

2.4. Training[3]

Our U-Net variant is trained to minimize the difference between predicted seal counts and true seal counts, the difference between predicted occupancy and true occupancy ('1' if there is at least one seal in the patch, '0' otherwise), both measured with a Huber loss, and the difference between the predicted heatmap, with a sigmoid transform, and an array with '1' over seal centroid pixels, smoothed around the centroid with a 5 × 5 gaussian kernel, and '0' anywhere else – measured with a binary cross-entropy loss. To ensure that seal training images and seal centroids within those training images are as important during training as the more prevalent background training images and non-centroid pixels, binary cross-entropy losses were weighted using the ratio between the former and the latter. Training is performed using an AdamW optimizer (Loshchilov and Hutter, 2019) for 75 epochs (i.e. 75 complete runs through the training set), with an initial learning rate of 1×10^{-3}, which was gradually tapered down to 1×10^{-5} using a cosine annealing learning rate scheduler (Loshchilov and Hutter, 2017), and a batch size of 64. Training images are sampled with replacement from the training set using a weighted sampler that ensures equal representation between training classes. Training images are normalized to have similar means and variances and augmented using left-right

[3] https://github.com/iceberg-project/Seals/blob/paper/SealNet_code/model_search_stable.ipynb.

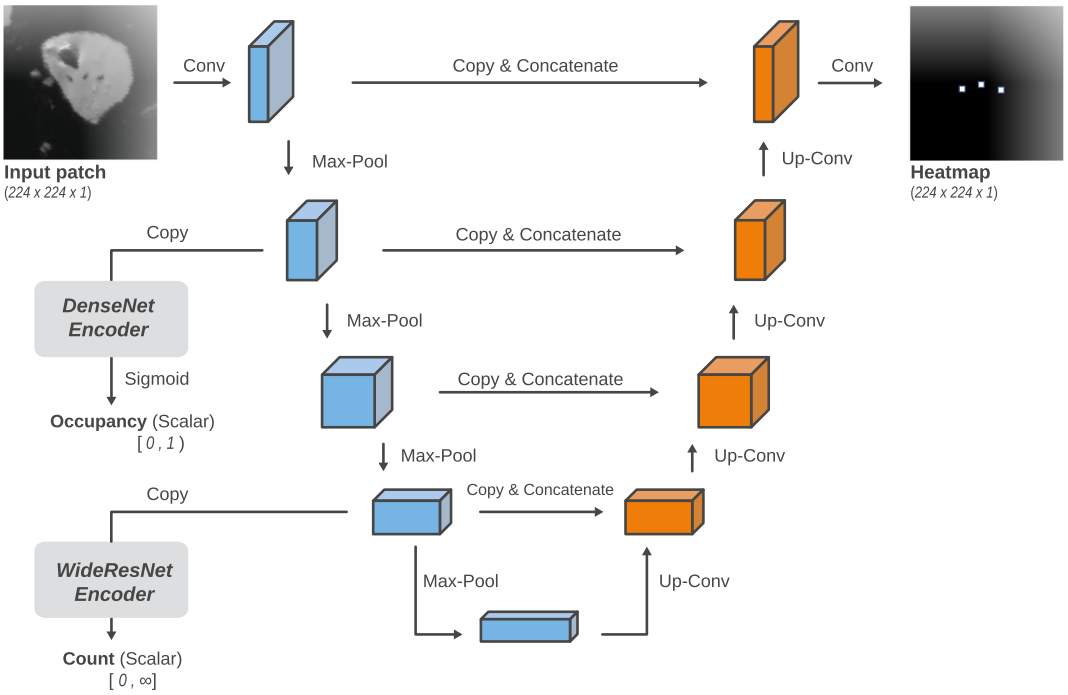

Fig. 4. SealNet architecture. The CNN takes in a patch as input, generates an occupancy probability and a seal count with peripheral branches, and reconstructs a heatmap for pixel-wise probability of being a seal centroid. Predicted seal centroids are determined outside of the CNN by finding the n largest intensity peaks on a patch, where n is the regressed seal count for that patch multiplied by a Boolean (0 or 1) indicating whether the occupancy probability for that patch surpasses a predefined threshold.
Model output is displayed in bold. Satellite imagery (upper left) copyright DigitalGlobe, Inc. 2019.

mirroring, bottom-up mirroring, random rotations (0–180), slight changes to brightness and contrast, random resized crops (0.675–1.2 of original scale, keeping the original aspect ratio) to the input size required by the CNN (224 × 224 in the current pipeline) and hide-and-seek transformations (Singh et al., 2018). Whenever cropping and hide-and-seek transformations are applied to training images, seal locations within these, the number of seals on them, and whether they are occupied are updated to reflect those of the augmented sample. For testing purposes, models were retrained with the same hyperparameter settings using all training and validation samples for 150 epochs.

2.5. Validation

Our model is validated at the end of each epoch. Prior to predictions, validation images are normalized and center-cropped to the dimensions required by the CNN. Similar to training, true counts and seal locations are adjusted to cropped validation images. Raw predictions on the validation set are converted to precision,

$$precision = \frac{truepositives}{truepositives + falsepositives}$$

recall

$$recall = \frac{truepositives}{truepositives + falsenegatives}$$

and F1 score

$$F1 = precision * recall$$

where predicted seal centroids separated by no more than 5 pixels from a ground-truth seal centroid are considered true positives. At the end of each validation phase, validation losses, precision, recall and F1 scores are recorded. Whenever the F1 score surpasses the previous best score, a model checkpoint with the weights for that formulation is saved.

2.6. Testing[4]

To test how SealNet generalizes to new imagery, we estimated out-of-sample precision and recall by comparing model-generated seal locations with those from the consensus of two experienced human observers on five novel scenes. First, test scenes were independently counted by two observers with experience surveying seals in Antarctica and using high resolution imagery (hereafter 'observer 1' and 'observer 2'). When looking for seals, observers followed a standardized counting procedure using a grid search system with 2 km × 2 km grid cells that were each exhaustively searched for potential seals. To create a consensus seal dataset for testing model performances, we started with seal points flagged by both observers. Points flagged by a single observer, after stripped of observer ID, were independently reviewed by both observers, adding further seal points where both observers agreed upon to the consensus dataset.

Prior to model predictions on the test set, test scenes are tiled out into patches, with a 75% overlap between neighboring patches. Whenever multiple model predictions from overlapping tiles output seal centroids within 1.5 m of each other, the centroid with the highest heatmap intensity value is kept and the remaining centroids are discarded. Our test set (Fig. 5) includes a pair of scenes over pack ice, with high (1.16 seals/km^2) and low (0.51 seals/km^2) seal densities, a pair of scenes over fast-ice, with high (4.06 seals/km^2) and low (0.30 seals/km^2) seal densities and a scene without seal detection by the observers. Apart from variations in seal density, test scenes were chosen to emulate scenarios likely to impact seal detectability, such as off-nadir angles and lighting conditions.

[4] https://github.com/iceberg-project/Seals/blob/paper/SealNet_code/deploy_sealnet.ipynb.

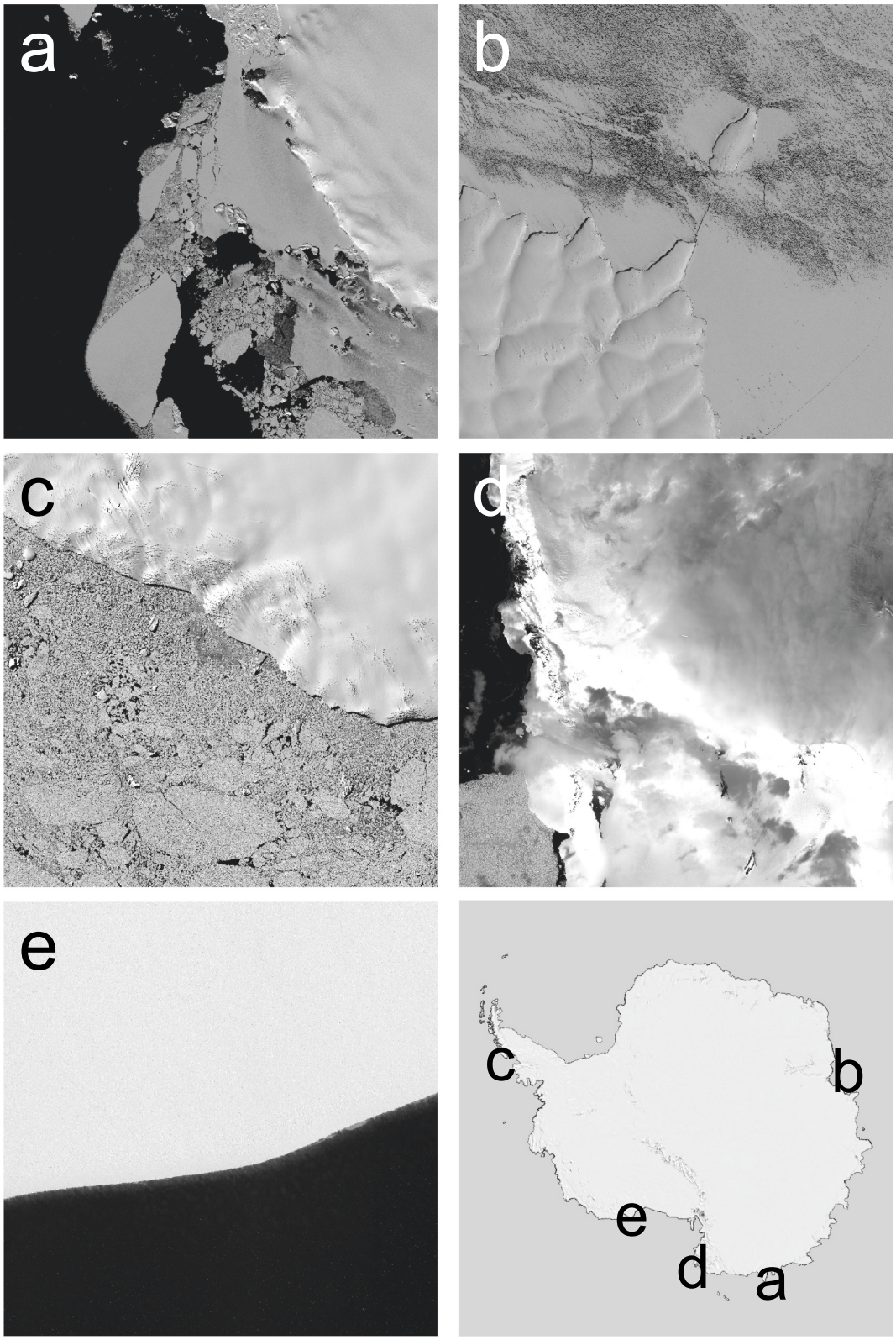

Fig. 5. (a–e) Satellite imagery representing the test set. Test set scenes are not included in training or validation sets and serve as way to get out-of-sample precision and recall over a range of scenarios that we are likely to encounter at deployment stages. Scenes a and b have seals over fast-ice, with low and high densities, respectively. Scenes c and d have seals over pack ice, with low and high densities, respectively. Scenes e covers Antarctic coastline landscape without seals. All test scenes were obtained between February and March 2017 at the locations specified in the Antarctic continent thumbnail at the lower right of the panel. Antarctic basemap extracted from Quantarctica (Matsuoka et al., 2018). Satellite imagery copyright DigitalGlobe, Inc. 2019.

2.7. Model evaluation

To test the SealNet architecture, precision and recall obtained on test and validation sets with the full model are compared to those obtained with two simplified variants: i) the original U-Net architecture, trained only on heatmap matching; and ii) U-Net with a branch for counting. Due to the lack of a regression layer, counts on the original U-Net are obtained by applying a sigmoid transformation to the heatmap, thresholding sigmoid transformed values by 0.1 (i.e. any value lower than 0.1 is set to 0) and summing over all cells. Pure regression CNNs (e.g., CountCeption [Cohen et al., 2017]) were not tested here because the lack of input image sized heatmaps for counts hampers localization and makes it difficult to match predicted centroids to ground-truth centroids. Finally, due to the relatively small size of our training set, the potential for improvement by acquiring more training data is investigated with a learning curve: we train our models with increasingly larger subsets (n = 100, n = 300, n = 1000, n = 3000, n = 10,000 and n = 30,000) of our training set for 15 epochs and plot highest validation F1 scores – measured on the full validation set – against training set size. To maintain equal representation between training classes while generating the learning curve, training set subsets were sampled, without replacement, from the full training set using a weighted sampler. Though there are too few training images in some classes (e.g., Weddell seals, n = 981) to keep classes balanced at the largest subset (n = 30,000), the weighted sampler draws images with replacement. This ensures that, though the full training set itself may not be balanced, batches of training images still have equal class representation. Apart from the reduced number of epochs (75 vs. 15), CNNs on reduced training sets were trained with the same hyperparameters as their counterparts trained on the full training set.

3. Results

3.1. Validation

SealNet, with added branches for counting and occupancy, attained 0.887 precision and 0.845 recall at our validation set, outperformed base U-Net (precision = 0.250, recall = 0.993), but was slightly outperformed by U-Net + count (precision = 0.897, recall = 0.853) (Fig. 6a). Adding a counting branch to U-Net, when compared with heatmap thresholding approach, improved precision at our validation set more than threefold, at the cost of a small decrement in recall. Adding an occupancy branch to U-Net + count caused a slight decrease in precision and recall at our validation set. Our validation metrics for SealNet use an occupancy threshold of 0.1 (i.e., patches with a predicted occupancy probability lower than 0.1 are discarded by the model), which can be tuned to tradeoff recall for precision and vice-versa. The learning curve for SealNet (Fig. 6b) shows that validation F1 score increases as we add more training data, suggesting that adding new samples to our training set would be very beneficial to model performance.

3.2. Testing

Combining results from all five test scenes (Fig. 5) and comparing the results with consensus counts from two human experts, SealNet outperforms U-Net + count on F1 score, while both CNN architectures get better precision and recall than U-Net (Table 2). When deployed on an empty test scene (Fig. 5, subpanel e), SealNet was the only architecture to not produce a single false positive; U-Net produced 26 false positives and U-Net + count produced a single false positive. When aggregating predicted seals by group (using a 20-meter distance criterion to define group membership), U-Net + count has a superior recall on finding lone seals than the other two architectures (0.311 vs. 0.230 [SealNet] and 0.196 [U-Net]) on lone seals, while SealNet is superior on finding seals inside groups from 3 to 5 seals and more than 10 seals (Fig. 8a). Groups with a small number of seals were far more prevalent in our test scenes (Fig. 5) than larger ones (Fig. 8b).

3.3. Human observer performance vs. CNN performance

When compared with a consensus review from both human observers, individual human observers made a considerable number of mistakes and were inconsistent across different scenes (Table 2, see Observer 1 and 2). Even so, human observer outperformed the CNNs tested in this analysis but at the expense of considerably more processing time (Table 2).

4. Discussion

4.1. CNN performance

Even with a relatively small training set (Table 1), weakly-supervised training samples and a test set with only 1168 seals distributed over 150,000 non-overlapping patches, our pipeline often produces reasonable predictions, including unmistakable seals missed by our double-observer count (Fig. 7). In contrast with typical usages of deep learning for RSISU, which rely on bounding box based approaches (e.g. YOLO [Redmon et al., 2015]), we explore instance-based approaches, in the form of U-Net variants, for object detection in remote sensing

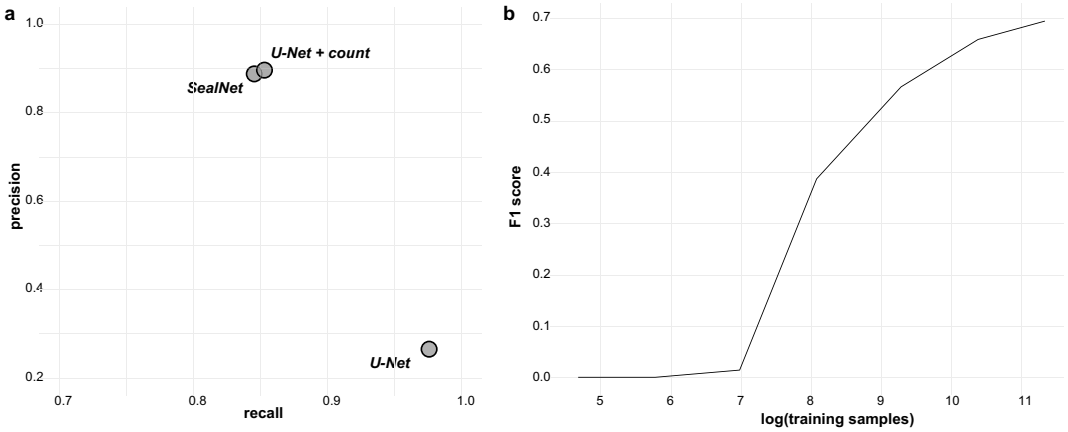

Fig. 6. (a) Validation performance: Validation precision and recall reported here are the highest obtained for 75 training epochs. Predicted seal centroids are considered true positives if their location is within 5 pixels of a manually annotated seal centroid. (b) Learning curve for SealNet: Top validation F1 score obtained during training epochs is displayed for SealNet instances trained on increasing large random subsets of our training set. Training set subsets are generated using a weighted sampler that ensures a similar class representation regardless of the number of training samples on a subset.

Table 1
Training set classes. For each of the first 10 classes, patches were manually annotated on WV03 imagery following the annotation method listed for that label. Training patches under the 'shadow' label were extracted using an early iteration of SealNet to find seals on scenes without seals. Note that the total number of scenes is smaller than the added number for all labels, since there are often several different labels in a single scene.

Class label	Annotation method	Number of patches	Number of scenes
Crabeater	1 patch centered on each individual seal	4238	6
Weddell	1 patch centered on each individual seal	981	15
Emperor	Array of patches with 10 m gaps over colony	7124	19
Marching-emperor	1 patch centered on each penguin line	1064	18
Pack ice	Array of patches with 100 m gaps over area	17,771	10
Ice-sheet	Array of patches with 100 m gaps over area	20,694	9
Glacier	Array of patches with 100 m gaps over area	5762	4
Crack	Array of patches with 100 m gaps over area	1449	4
Rock	Array of patches with 100 m gaps over area	4836	6
Open-water	Array of patches with 100 m gaps over area	11,799	5
Shadow	Extracted from CNN output in scenes with no seals	10,766	18
Total	–	86,483	52

scenes. Apart from requiring lower annotation effort (i.e. centroids vs. bounding boxes), our approach excels at object localization with little to no post-processing non-maximum suppression efforts. Similar to Le et al. (2019), this work highlights the potential for weakly-supervised approaches on remote sensing tasks.

When compared with test set counts from our double-observer consensus set, our most sensitive model finds over 35% of seals, generating 1321 false positives, while our most precise model finds 30% of seals, generating 604 false positives. Perhaps due to the lower information content, seals hauling out by themselves were more often missed by the CNNs than those in larger groups (Figs. 7, 8a). Although group size had a profound impact on recall (Fig. 8a), effects were not consistent across our 3 model architectures. These variations can be used to divide predictions between different models (e.g., U-Net + count on small groups and SealNet on large ones), model detection errors and, in semi-automated pipelines, highlight predictions that will require more attention from human observers. SealNet predictions are still not as reliable as those obtained by an experienced human observer. However, using our pipeline, with a single modern GPU, is over 10× faster than counting by hand. While similar studies in wildlife samples from aerial or VHR imagery like Salberg (2015) and Xue et al. (2017) report higher performance scores, the first uses aerial imagery, with far superior spatial resolution, and the second relies on expert opinion for inference, in contrast to our fully automated

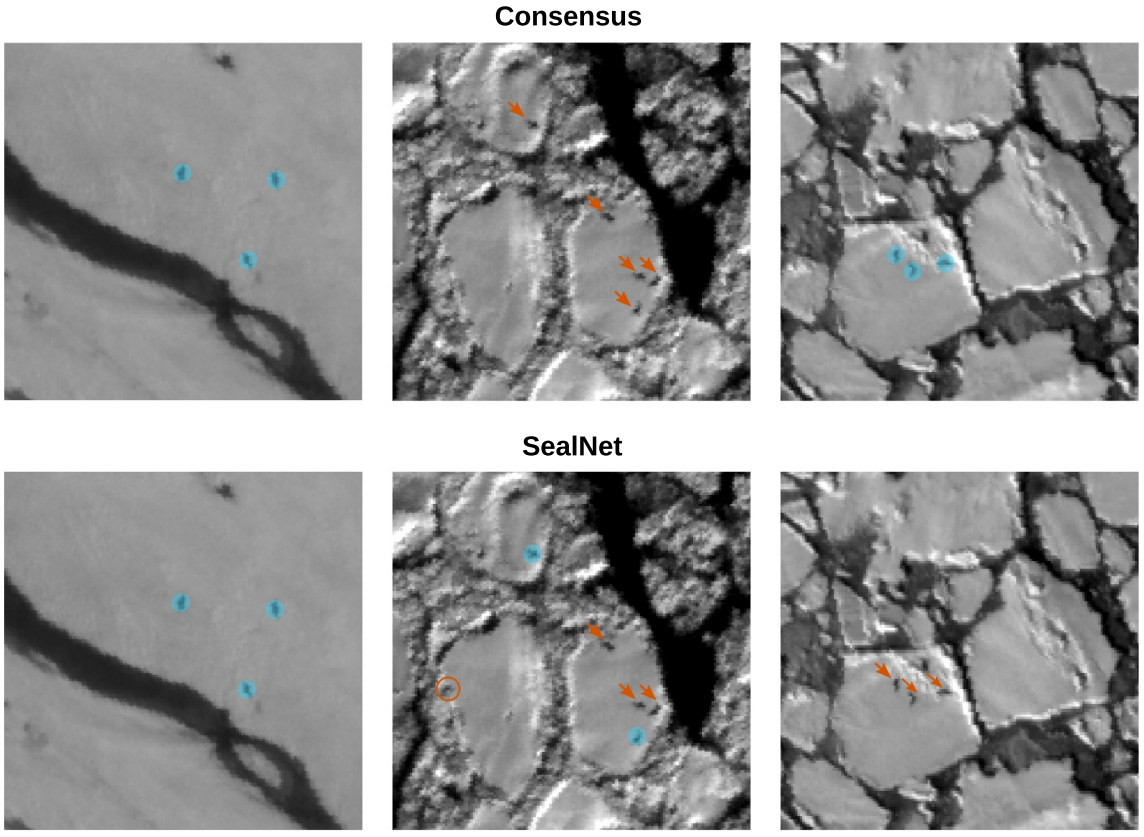

Fig. 7. Sample SealNet output. Panels true-positives (light-blue circles) false negatives (orange arrows) and false positives (orange open circles), by a double-observer consensus (upper panel) and SealNet (lower panel). Examples, from left to right, show a case where both SealNet and the consensus set locate seals faultlessly, a case where SealNet outperforms the consensus set, a case where it underperforms the consensus set. Crops were extracted at 1:500 scale from the test scenes b, c and d, respectively (Fig. 5). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Satellite imagery copyright DigitalGlobe, Inc. (2019).

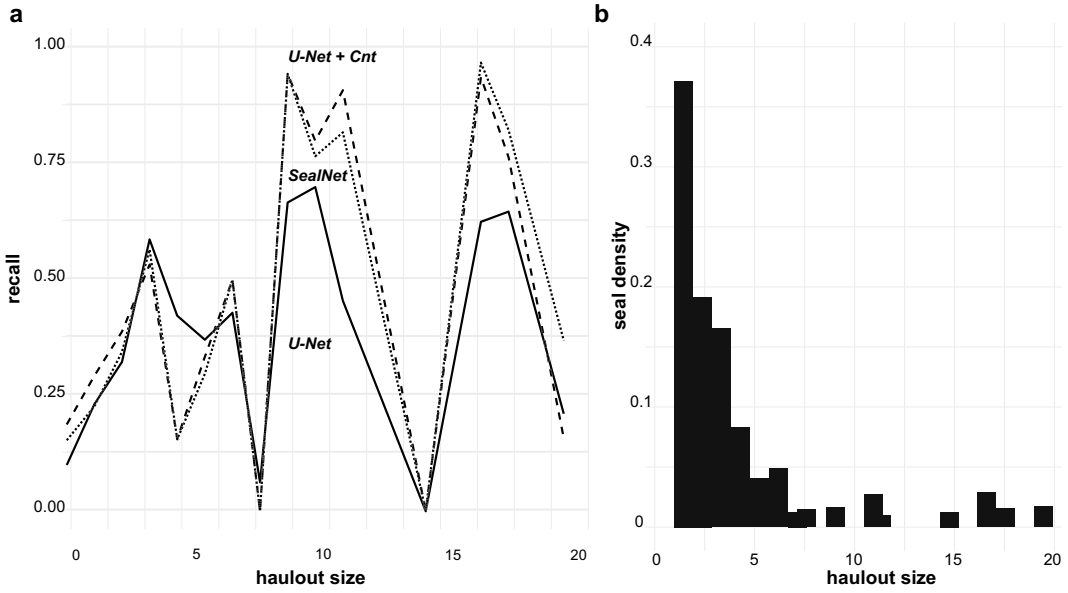

Fig. 8. (a) CNN performance on different group sizes. Recall values are extracted by measuring the proportion of ground-truth seal points on specific sized haul outs a model can recover. (b) Test group size distribution. The y-axis shows the proportion of seals across all test scenes that are located in groups of size x.

pipeline. Moreover, both use less rigorous testing than demonstrated here for SealNet (i.e. small test sets, cross-validation on the training set, overlap between training and testing scenes).

There are large differences between validation (Fig. 6a) and test performance metrics (Table 2), with test performance showing dramatically lower precision and recall. This outcome may be explained by the relatively small size of our validation set (8715 patches across 49 scenes) and by the small number of test scenes ($n = 5$), which may have caused our validation set to be insufficiently representative of the problem at hand (i.e. finding seals in WV03 scenes from the Antarctic coastline) and/or our test set to be a biased sample of typical Antarctic scenes. Similar to results from Aich and Stavness (2018), combining heatmap activation to counting by regression output improves precision and recall (Fig. 6a). Besides greatly improving precision at the cost of some recall at our chosen threshold of 0.1, our occupancy branch gives us flexibility to tradeoff precision and recall by picking more strict or lenient threshold values. Learning curve results (Fig. 6b) suggest that acquiring more training data can lead to substantial improvements in prediction accuracy, though we caution that our validation set may not be sufficiently representative of future test settings.

Similar to other ecological sampling settings with aerial imagery (Brack et al., 2018; Kellenberger et al., 2018), empty patches predominate in our test set, outnumbering those with seals by a factor of 500 – which is aggravated by our use of overlapping patches. Though the pervasiveness of false negatives and false positives in our model output calls for adequate statistical treatment before making any inferences about seal populations (e.g., Miller et al., 2013; Pillay et al., 2014), raw model output can also serve as an attention map for human observers, facilitating manual annotation. Besides being immediately applicable to semi-automated pack-ice seal surveys, lower annotation effort speeds up the acquisition of training data, which, as indicated by our learning curve (Fig. 6b), may boost prediction performance enough to bridge the gap between fully-automated approaches and manual surveys. When considering the performance of an automated

Table 2
Test performance. Predicted count, precision and recall using all model variants are shown for scenes *a–d*. We only include a predicted count for scene *e* because we cannot get meaningful precision or recall scores without ground-truth seal points. Performance metrics are obtained by comparing model predicted seal locations with a consensus review from two experienced human observers. Patch counts reflect a stride that keeps a 75% overlap between neighboring cells (see Fig. 2).

Model architecture	Scene a	Scene b	Scene c	Scene d	Scene e
	Consensus count: 106	Consensus count: 732	Consensus count: 282	Consensus count: 48	Consensus Count: 0
	Patches: 127,964	Patches: 127,332	Patches: 138,308	Patches: 78,334	Patches:173,340
SealNet	Count: 57	Count: 809	Count: 58	Count: 33	Count: 0
	Precision: 0.492	Precision: 0.344	Precision: 0.519	Precision: 0.324	–
	Recall: 0.277	Recall: 0.377	Recall: 0.133	Recall: 0.224	–
U-Net	Count: 461	Count: 4865	Count: 906	Count: 246	Count: 26
	Precision: 0.049	Precision: 0.049	Precision: 0.0.073	Precision: 0.032	–
	Recall: 0.207	Recall: 0.319	Recall: 0.223	Recall: 0.163	–
U-Net + count	Count: 191	Count: 1267	Count: 139	Count: 131	Count: 1
	Precision: 0.179	Precision: 0.240	Precision: 0.430	Precision: 0.134	–
	Recall: 0.315	Recall: 0.402	Recall: 0.226	Recall: 0.353	–
Observer 1	Count: 50	Count: 1321	Count: 299	Count: 45	Count: 0
	Precision: 0.641	Precision: 0.527	Precision: 0.569	Precision: 0.639	
	Recall: 0.373	Recall: 0.777	Recall: 0.593	Recall: 0.613	
Observer 2	Count: 168	Count: 732	Count: 218	Count: 72	Count: 1
	Precision: 0.580	Precision: 0.635	Precision: 0.533	Precision: 0.527	
	Recall: 0.784	Recall: 0.635	Recall: 0.437	Recall: 0.716	

classification pipeline, it is worth highlighting that there is a comprehensive literature on statistical methods for dealing with observation errors in wildlife surveying (e.g., McClintock et al., 2015; Miller et al., 2013, 2011), and the existence of high error rates does not *ipso facto* preclude unbiased or reliable population estimates. While the vast majority of high spatial resolution imagery for the Antarctic is focused on terrestrial areas, there are no technical barriers to repeated sampling of key marine regions, and the development of a well-structured sampling regime for pack-ice seals could enable global population estimates on an yearly basis.

Besides exploring hyperparameter space and adding more training samples, predictive performance could be improved by adding environmental features as additional inputs to the CNN. This approach, akin to a habitat suitability model, takes advantage of the fact that all our training and input imagery is georeferenced and may be coupled with several biologically relevant measurements associated with seal presence/density in previous studies such as sea ice characteristics, bathymetry, water temperature and distance to shelf-break (Ballard et al., 2012; Bengtson et al., 2011; Bester et al., 2002b; Nachtsheim et al., 2017). Integrated habitat suitability and detection models provide a promising path forward and would likely improve classification accuracy while potentially informing on seal habitat selection and haul out patterns.

4.2. A path forward for a Southern Ocean pack-ice seal monitoring program

Past large-scale pack-ice seal surveys, both ship-based (e.g., Erickson and Hanson, 1990) and aerial based (Gurarie et al., 2016), have provided initial clues on pack-ice seal distribution and population sizes. These estimates, unavoidably, rely on data aggregated from different sampling methods (e.g., Bengtson et al., 2011; Gurarie et al., 2016) and from broad and discontinuous time windows (e.g., Anonymous, 1997; Erickson and Hanson, 1990), or require an extrapolation for the entire coastline using seal density estimates derived from a single region (Erickson and Hanson, 1990). Though mitigated by increasingly sophisticated statistical treatment, we argue that these limitations, due to the substantial expenditure required for Antarctic surveys, are inherent properties of sampling method choices. As our pilot study illustrates, satellite-based surveys, aided by automated detection with CNNs, can overcome these difficulties and provide large-scale pack-ice seal censuses, using a single sampling design which can be repeated yearly at spatial scales that cover a representative sample of the Antarctic coastline.

This paper represents the first step towards a regular, cost-effective, pan-Antarctic seal monitoring program. Apart from algorithmic improvements and more training data, our pipeline would greatly benefit from a test set that is paired with concurrent observations by a ground-based observer or aerial photographs. A comprehensive pack-ice seal census would provide key information to understand how the SO ecosystem will react to threats such as climate-change-driven sea ice loss (Lee et al., 2017) and increasing krill fisheries (Nicol et al., 2012). An automated tool to survey pack-ice seals allows us not only to get a better idea about their abundance and long-term trends, but also how their distribution is coupled to environmental features (e.g., sea ice conditions) or affected by external drivers (e.g., krill fisheries). This approach can be easily adapted for counting other large-bodied species visible from high spatial resolution satellite imagery, and we have provided the code to encourage other researchers to adapt the pipeline for their needs.

Author contribution

BG selected scenes for training and testing and did the manual annotation of imagery for training. BG and HL designed the testing double-observer approach, did the manual annotation of imagery for testing, and interpreted results. BG led the SealNet development and coding. BS helped with code development and computational scaling. All authors contributed to the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the Institute for Advanced Computational Science and the National Science Foundation EarthCube program (Award 1740595) for funding this work. Geospatial support for this work provided by the Polar Geospatial Center under NSF-OPP awards 1043681 and 1559691. The development of our CNN detection pipeline would not be possible without the advice of Felipe Codevilla Moraes, Hieu Le, Dimitris Samaras and the Stony Brook Computer Vision lab. We thank the Polar Geospatial Center for curating and providing the satellite imagery.

References

Ackley, S., Bengtson, J., Bester, M., Blix, A., Bornemann, H., Boveng, P., Boyd, I., Cameron, M., Nordøy, E., Plötz, J., Siniff, D., Southwell, C., Steinhage, D., Stewart, B., Stirling, J., Torres, J., Yochem, P., 2006. The International Antarctic Pack Ice Seals (APIS) Program. Multi-disciplinary Research into the Ecology and Behavior of Antarctic Pack Ice Seals. Summary Update.

Aich, S., Stavness, I., 2018. Improving Object Counting With Heatmap Regulation. arXiv:1803.05494v2 [cs.CV]. https://doi.org/10.1001/archinte.168.13.1430.

Anonymous, 1997. Antarctic pack ice seals: an international research program co-ordinated by the SCAR group of specialists on seals. In: Unpubl. Rep. 1995 APIS Progr. Plan. Meet. Natl. Mar. Mammal Lab. Alaska Fish. Sci. Center, Seattle, USA, pp. 1–26.

Arrigo, K.R., Dijken, G.L. van, 2003. Phytoplankton dynamics within 37 Antarctic coastal polynya systems. J. Geophys. Res. 108, 3271. https://doi.org/10.1029/2002JC001739.

Ballard, G., Jongsomjit, D., Veloz, S.D., Ainley, D.G., 2012. Coexistence of mesopredators in an intact polar ocean ecosystem: the basis for defining a Ross Sea marine protected area. Biol. Conserv. 156, 72–82. https://doi.org/10.1016/j.biocon.2011.11.017.

Bengtson, J.L., Stewart, B.S., 1992. Diving and haulout behavior of crabeater seals in the Weddell Sea, Antarctica, during March 1986. Polar Biol. 12, 635–644. https://doi.org/10.1007/BF00236986.

Bengtson, J.L., Laake, J.L., Boveng, P.L., 2001. Habitat partitioning among Antarctic pack ice seals. In: 14th Bienn. Conf. Biol. Mar. Mammals, Novemb. 28-December 3.

Bengtson, J.L., Laake, J.L., Boveng, P.L., Cameron, M.F., Bradley Hanson, M., Stewart, B.S., Hanson, M.B., Stewart, B.S., Bradley Hanson, M., Stewart, B.S., 2011. Distribution, density, and abundance of pack-ice seals in the Amundsen and Ross Seas, Antarctica. Deep Sea Res. Part 2 Top. Stud. Ocean. 58, 1261–1276. https://doi.org/10.1016/j.dsr2.2010.10.037.

Bester, M.N., Ferguson, J.W.H., Jonker, F.C., 2002a. Population densities of pack ice seals in the Lazarev Sea, Antarctica. Antarct. Sci. 14, 139–143. https://doi.org/10.1017/S0954102002000676.

Bester, M.N., Ferguson, J.W.H., Jonker, F.C., 2002b. Population densities of pack ice seals in the Lazarev Sea, Antarctica. Antarct. Sci. 14, 123–127. https://doi.org/10.1017/s0954102002000676.

Borowicz, A., Le, H., Humphries, G., Nehls, G., Höschle, C., Kosarev, V., Lynch, H.J., 2019. Aerial-trained deep learning networks for surveying cetaceans from satellite imagery. PLoS One 14, e0212532. https://doi.org/10.1371/journal.pone.0212532.

Botta, S., Rogers, T.L., Prado, J.H.F., de Lima, R.C., Carlini, P., Negrete, J., 2018. Isotopic niche overlap and partition among three Antarctic seals from the Western Antarctic Peninsula. Deep Sea Res. Part II Top. Stud. Oceanogr. 149, 240–249. https://doi.org/10.1016/J.DSR2.2017.11.005.

Brack, I.V., Kindel, A., Oliveira, L.F.B., 2018. Detection errors in wildlife abundance estimates from Unmanned Aerial Systems (UAS) surveys: synthesis, solutions, and challenges. Methods Ecol. Evol. 9, 1864–1873. https://doi.org/10.1111/2041-210X.13026.

Clarke, E., 1985. Energy flow in the Southern Ocean food web. In: Antarctic Nutrient Cycles and Food Webs. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 573–580. https://doi.org/10.1007/978-3-642-82275-9_78.

Cohen, J.P., Boucher, G.G., Glastonbury, C.A., Lo, H.Z., Bengio, Y., 2017. Count-ception: counting by fully convolutional redundant counting. In: Proceedings - 2017 IEEE International Conference on Computer Vision Workshops, ICCVW 2017. Institute of Electrical and Electronics Engineers Inc, pp. 18–26. https://doi.org/10.1109/ICCVW.2017.9.

Conn, P.B., Ver Hoef, J.M., Mcclintock, B.T., Moreland, E.E., London, J.M., Cameron, M.F., Dahle, S.P., Boveng, P.L., 2014. Estimating multispecies abundance using automated detection systems: ice-associated seals in the Bering Sea. Methods Ecol. Evol. 5, 1280–1293. https://doi.org/10.1111/2041-210X.12127.

Dickinson, J.L., Zuckerberg, B., Bonter, D.N., 2010. Citizen science as an ecological research tool: challenges and benefits. Annu. Rev. Ecol. Evol. Syst. 41, 149–172.

https://doi.org/10.1146/annurev-ecolsys-102209-144636.
Do, H.H., Prasad, P.W.C., Maag, A., Alsadoon, A., 2019. Deep learning for aspect-based sentiment analysis: a comparative review. Expert Syst. Appl. https://doi.org/10.1016/j.eswa.2018.10.003.
Erickson, A.W., Hanson, M.B., 1990. Continental estimates and population trends of Antarctic ice seals. In: Antarctic Ecosystems. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 253–264. https://doi.org/10.1007/978-3-642-84074-6_29.
Flores, H., Atkinson, A., Kawaguchi, S., Krafft, B.A., Milinevsky, G., Nicol, S., Reiss, C., Tarling, G.A., Werner, R., Bravo Rebolledo, E., Cirelli, V., Cuzin-Roudy, J., Fielding, S., Groeneveld, J.J., Haraldsson, M., Lombana, A., Marschoff, E., Meyer, B., Pakhomov, E.A., Rombolá, E., Schmidt, K., Siegel, V., Teschke, M., Tonkes, H., Toullec, J.Y., Trathan, P.N., Tremblay, N., Van De Putte, A.P., Van Franeker, J.A., Werner, T., 2012. Impact of climate change on Antarctic krill. Mar. Ecol. Prog. Ser. 458, 1–19. https://doi.org/10.3354/meps09831.
Forcada, J., Trathan, P.N., Boveng, P.L., Boyd, I.L., Burns, J.M., Costa, D.P., Fedak, M., Rogers, T.L., Southwell, C.J., 2012. Responses of Antarctic pack-ice seals to environmental change and increasing krill fishing. Biol. Conserv. 149, 40–50. https://doi.org/10.1016/j.biocon.2012.02.002.
Frölicher, T.L., Sarmiento, J.L., Paynter, D.J., Dunne, J.P., Krasting, J.P., Winton, M., 2015. Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. J. Clim. 28, 862–886. https://doi.org/10.1175/JCLI-D-14-00117.1.
Gu, Y., Wang, Y., Li, Y., 2019. A survey on deep learning-driven remote sensing image scene understanding: scene classification, scene retrieval and scene-guided object detection. Appl. Sci. https://doi.org/10.3390/app9102110.
Gurarie, E., Bengtson, J.L., Bester, M.N., Blix, A.S., Cameron, M., Bornemann, H., Nordøy, E.S., Plötz, J., Steinhage, D., Boveng, P., 2016. Distribution, density and abundance of Antarctic ice seals off Queen Maud Land and the eastern Weddell Sea. Polar Biol. 1–17. https://doi.org/10.1007/s00300-016-2029-4.
Huang, T., Sun, L., Stark, J., Wang, Y., Cheng, Z., Yang, Q., Sun, S., 2011. Relative changes in krill abundance inferred from antarctic fur seal. PLoS One 6, e27331. https://doi.org/10.1371/journal.pone.0027331.
Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017. doi:https://doi.org/10.1109/CVPR.2017.243.
Hückstädt, L., Burns, J., Koch, P., McDonald, B., Crocker, D., Costa, D., 2012. Diet of a specialist in a changing environment: the crabeater seal along the western Antarctic Peninsula. Mar. Ecol. Prog. Ser. 455, 287–301. https://doi.org/10.3354/meps09601.
Kawaguchi, S., Ishida, a., King, R., Raymond, B., Waller, N., Constable, a., Nicol, S., Wakita, M., Ishimatsu, a., 2013. Risk maps for Antarctic krill under projected Southern Ocean acidification. Nat. Clim. Chang. 3, 843–847. https://doi.org/10.1038/nclimate1937.
Kellenberger, B., Marcos, D., Tuia, D., 2018. Detecting mammals in UAV images: best practices to address a substantially imbalanced dataset with deep learning. Remote Sens. Environ. 216, 139–153. https://doi.org/10.1016/j.rse.2018.06.028.
Klein, E.S., Hill, S.L., Hinke, J.T., Phillips, T., Watters, G.M., 2018. Impacts of rising sea temperature on krill increase risks for predators in the Scotia Sea. PLoS One 13, e0191011. https://doi.org/10.1371/journal.pone.0191011.
Koju, U.A., Zhang, J., Maharjan, S., Zhang, S., Bai, Y., Vijayakumar, D.B.I.P., Yao, F., 2018. A two-scale approach for estimating forest aboveground biomass with optical remote sensing images in a subtropical forest of Nepal. J. For. Res. 1–18. https://doi.org/10.1007/s11676-018-0743-1.
Lake, S.E., Burton, H.R., Hindell, M.A., 1997. Influence of time of day and month on Weddell seal haul-out patterns at the Vestfold Hills, Antarctica. Polar Biol. 18, 319–324. https://doi.org/10.1007/s003000050194.
LaRue, M.A., Rotella, J.J., Garrott, R.A., Siniff, D.B., Ainley, D.G., Stauffer, G.E., Porter, C.C., Morin, P.J., 2011. Satellite imagery can be used to detect variation in abundance of Weddell seals (Leptonychotes weddellii) in Erebus Bay, Antarctica. Polar Biol. 34, 1727. https://doi.org/10.1007/s00300-011-1023-0.
LaRue, M.A., Lynch, H.J., Lyver, P.O.B., Barton, K., Ainley, D.G., Pollard, A., Fraser, W.R., Ballard, G., 2014. A method for estimating colony sizes of Adélie penguins using remote sensing imagery. Polar Biol. 37, 507–517. https://doi.org/10.1007/s00300-014-1451-8.
Le, H., Gonçalves, B., Samaras, D., Lynch, H., 2019. Weakly labeling the Antarctic: The penguin Colony case. In: CVPR 2019 Workshop CV4GC: Computer Vision for Global Causes.
Lee, J.R., Raymond, B., Bracegirdle, T.J., Chadès, I., Fuller, R.A., Shaw, J.D., Terauds, A., 2017. Climate change drives expansion of Antarctic ice-free habitat. Nat. Publ. Gr. 547, 49–54. https://doi.org/10.1038/nature22996.
Loshchilov, I., Hutter, F., 2017. SGDR: stochastic gradient descent with warm restarts. In: The International Conference on Learning Representations (ICLR).
Loshchilov, I., Hutter, F., 2019. Decoupled weight decay regularization. In: The International Conference on Learning Representations (ICLR).
Lynch, H.J., White, R., Black, A.D., Naveen, R., 2012. Detection, differentiation, and abundance estimation of penguin species by high-resolution satellite imagery. Polar Biol. 35, 963–968. https://doi.org/10.1007/s00300-011-1138-3.
Matsuoka, K., Skoglund, A., Roth, G., 2018. Quantarctica. WWW Document. Nor. Polar Insthttps://doi.org/10.21334/npolar.2018.8516e961.
McClintock, B.T., Moreland, E.E., London, J.M., Dahle, S.P., Brady, G.M., Richmond, E.L., Yano, K.M., Boveng, P.L., 2015. Quantitative assessment of species identification in aerial transect surveys for ice-associated seals. Mar. Mammal Sci. 31, 1057–1076. https://doi.org/10.1111/mms.12206.
McMahon, C.R., Howe, H., van den Hoff, J., Alderman, R., Brolsma, H., Hindell, M.A., 2014. Satellites, the all-seeing eyes in the sky: counting elephant seals from space. PLoS One 9, e92613.
McNabb, R.W., Womble, J.N., Prakash, A., Gens, R., Haselwimmer, C.E., 2016. Quantification and analysis of icebergs in a tidewater glacier fjord using an object-based approach. PLoS One 11, e0164444. https://doi.org/10.1371/journal.pone.0164444.
Miller, D.A., Nichols, J.D., McClintock, B.T., Grant, E.H.C., Bailey, L.L., Weir, L.A., 2011. Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. Ecology 92, 1422–1428. https://doi.org/10.1890/10-1396.1.
Miller, D. a W., Nichols, J.D., Gude, J. a, Rich, L.N., Podruzny, K.M., Hines, J.E., Mitchell, M.S., 2013. Determining occurrence dynamics when false positives occur: estimating the range dynamics of wolves from public survey data. PLoS One 8, e65808. https://doi.org/10.1371/journal.pone.0065808.
Morrison, A.K., Griffies, S.M., Winton, M., Anderson, W.G., Sarmiento, J.L., 2016. Mechanisms of Southern Ocean heat uptake and transport in a global eddying climate model. J. Clim. 29, 2059–2075. https://doi.org/10.1175/JCLI-D-15-0579.1.
Nachtsheim, D.A., Jerosch, K., Hagen, W., Plötz, J., Bornemann, H., 2017. Habitat modelling of crabeater seals (Lobodon carcinophaga) in the Weddell Sea using the multivariate approach Maxent. Polar Biol. 40, 961–976. https://doi.org/10.1007/s00300-016-2020-0.
Nicol, S., Foster, J., Kawaguchi, S., 2012. The fishery for Antarctic krill–recent developments. Fish Fish. 13, 30–40. https://doi.org/10.1111/j.1467-2979.2011.00406.x.
Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc. Natl. Acad. Sci. U. S. A. 115, E5716–E5725. https://doi.org/10.1073/pnas.1719367115.
Nowacek, D.P., Friedlaender, A.S., Halpin, P.N., Hazen, E.L., Johnston, D.W., Read, A.J., Espinasse, B., Zhou, M., Zhu, Y., 2011. Super-aggregations of krill and humpback whales in Wilhelmina bay, Antarctic Peninsula. PLoS One 6, e19173. https://doi.org/10.1371/journal.pone.0019173.
Paszke, A., Chanan, G., Lin, Z., Gross, S., Yang, E., Antiga, L., Devito, Z., 2017. Automatic differentiation in PyTorch. In: 31st Conf. Neural Inf. Process. Syst, pp. 1–4. https://doi.org/10.1017/CBO9781107707221.009.
Pillay, R., Miller, D.A.W., Hines, J.E., Joshi, A.A., Madhusudan, M.D., 2014. Accounting for false positives improves estimates of occupancy from key informant interviews. Divers. Distrib. 20, 223–235. https://doi.org/10.1111/ddi.12151.
Polzounov, A., Terpugova, I., Skiparis, D., Mihai, A., 2016. Right Whale Recognition Using Convolutional Neural Networks. (arXiv:1604.05605v1 [cs.CV]).
Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2015. You only look once: unified, real-time object detection. In: Conference on Computer Vision and Pattern Recognition (CVPR), https://doi.org/10.1109/CVPR.2016.91.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: MICCAI. Springer, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
Salberg, A.B., 2015. Detection of seals in remote sensing images using features extracted from deep convolutional neural networks. Int. Geosci. Remote Sens. Symp. 2015-Novem 1893–1896. https://doi.org/10.1109/IGARSS.2015.7326163.
Singh, K.K., Yu, H., Sarmasi, A., Pradeep, G., Lee, Y.J., 2018. Hide-and-Seek: A Data Augmentation Technique for Weakly-supervised Localization and Beyond.
Siniff, D.B., Stone, S., 1985. The role of the leopard seal in the tropho-dynamics of the Antarctic marine ecosystem. In: Antarctic Nutrient Cycles and Food Webs. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 555–560. https://doi.org/10.1007/978-3-642-82275-9_75.
Southwell, C., Paxton, C.G.M., Borchers, D., Boveng, P., De La Mare, W., 2008a. Taking account of dependent species in management of the Southern Ocean krill fishery: estimating crabeater seal abundance off east Antarctica. J. Appl. Ecol. 45, 622–631. https://doi.org/10.1111/j.1365-2664.2007.01399.x.
Southwell, C., Paxton, C.G.M., Borchers, D., Boveng, P., Rogers, T., de la Mare, W.K., 2008b. Uncommon or cryptic? Challenges in estimating leopard seal abundance by conventional but state-of-the-art methods. Deep. Res. Part I Oceanogr. Res. Pap. 55, 519–531. https://doi.org/10.1016/j.dsr.2008.01.005.
Southwell, C., Paxton, C.G.M., Borchers, D.L., Boveng, P.L., Nordøy, E.S., Blix, A.S., De La Mare, W.K., 2008c. Estimating population status under conditions of uncertainty: the Ross seal in East Antarctica. Antarct. Sci. 20, 123–133. https://doi.org/10.1017/S0954102007000879.
Southwell, C., Bengtson, J., Bester, M., Schytte Blix, A., Bornemann, H., Boveng, P., Cameron, M., Forcada, J., Laake, J., Nordøy, E., Plötz, J., Rogers, T., Southwell, D., Steinhage, D., Stewart, B.S., Trathan, P., 2012. A review of data on abundance, trends in abundance, habitat use and diet of ice-breeding seals in the Southern Ocean. CCAMLR Sci 19, 1–26.
Stapleton, S., LaRue, M., Lecomte, N., Atkinson, S., Garshelis, D., Porter, C., Atwood, T., 2014. Polar bears from space: assessing satellite imagery as a tool to track Arctic wildlife. PLoS One 9, e101513. https://doi.org/10.1371/journal.pone.0101513.
Voronina, N.M., 1998. Comparative abundance and distribution of major filter-feeders in the Antarctic Pelagic Zone. J. Mar. Syst. 375–390. https://doi.org/10.1016/S0924-7963(98)00050-5.
Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E., 2018. Deep learning for computer vision: a brief review. Comput. Intell. Neurosci. https://doi.org/10.1155/2018/7068349.
Xue, Y., Wang, T., Skidmore, A.K., 2017. Automatic counting of large mammals from very high resolution panchromatic satellite imagery. Remote Sens. 9, 878. https://doi.org/10.3390/rs9090878.
Yang, Z., Wang, T., Skidmore, A.K., De Leeuw, J., Said, M.Y., Freer, J., 2014. Spotting East African mammals in open savannah from space. PLoS One 9, e115989. https://doi.org/10.1371/journal.pone.0115989.
Zagoruyko, S., Komodakis, N., 2016. Wide Residual Networks. (arXiv [cs.CV]).