Identifying Sea Ice Ridging in SAR Imagery using various Machine Learning Models

Daniel Sola University of Waterloo, ON, Canada
K Andrea Scott University of Waterloo, ON, Canada
Email: {daniel.sola, ka3scott}@uwaterloo.ca

Abstract

Sea ice ridging presents a great challenge to ships navigating the arctic. In this paper, we examine the capabilities of various machine learning methods in predicting regions of high ridge density from SAR imagery of Hudson Strait. Our results showed that although ridging in Hudson Strait may be difficult to distinguish even with the human eye, machine learning can give some insight into potentially dangerous regions of Hudson Strait.

1 Introduction

Shipping routes in Hudson Strait are used year round for the transport of natural resources. During the winter months, ships can become beset (stuck) in pressured (or ridged) ice for an average of 42% of their total transit time [1]. Attempts have been made to automate ridge detection using a random forest model for the Baltic Sea [2], but no work has been done for Hudson Strait to the best of our knowledge.

Over 1500 synthetic aperture radar (SAR) HH (horizontal transmit-horizontal receive) back-scatter images of Hudson Strait were available from RADARSAT-1 and RADARSAT-2 between the years 1997-2013. In the back-scatter imagery, ridges appear as bright linear features. For images ranging between the years 1997 and 2012, individual ridges were manually identified and labeled [1].

2 Methodology

Patches of size 125 x 125 pixels were extracted from SAR imagery acquired between December 29, 2002 and May 26, 2012. These patches correspond to an area of 25 km x 25 km, and were extracted within an AOI defined as the area where labelling occurred in [1]. A stride of 42 pixels was used, resulting in approximately 66% overlap in the horizontal and vertical directions. Each patch could contain no more than 15% land area, and needed to have an ice concentration of 60% or more as calculated from passive microwave data [3]. If the patch contained 10 or more ridges it was considered a positive example, while it was considered a negative example if it contained no ridges. In total, 22,017 patches were collected with 11,432 being positive examples and 10,585 being negative examples. It is important to note that patches were only taken from SAR images that had ridges labelled somewhere within the image. As not all SAR images had corresponding labels, this step was necessary to ensure that the images used were indeed inspected by [1] when the ridge label data set was created.

To generate features, the mean, standard deviation, skewness, and kurtosis were calculated for each patch [2]. These four features were used to train a logistic regression, random forest, and support vector machine (SVM) model. The hyper-parameters of both models were tuned based on their AUC-ROC scores.

For testing and visualizing the results, two approaches were taken. To calculate AUC-ROC scores and compare models, the data was randomly split in which 80% of the data was used for training, 20% was used for testing. For visualizing the results in a map view, six SAR images were held out of a training set and then tested on individually. The six image holdout method was not used to calculate the scores in Table 1 as patches in the holdout images often had between one and nine ridges and therefore were not labeled.

3 Results and Discussion

As outlined in Table 1, the random forest model outperformed both the logistic regression and SVM models with an AUC-ROC score of 71. Since no other attempts to classify ridging have been made using this Hudson Strait data set, the results can be compared to those of [2] for the Baltic Sea where an accuracy of 81.5% was achieved for binary classification (ridging/no ridging). As the Hudson Strait models perform worse than the Baltic Sea random forest model, it suggests that the labelled ridges in Hudson Strait may be more ambiguous than the "degree of ridging" labels in the Baltic Sea. Nevertheless, the random forest model appears to be a superior classifier for this data set.

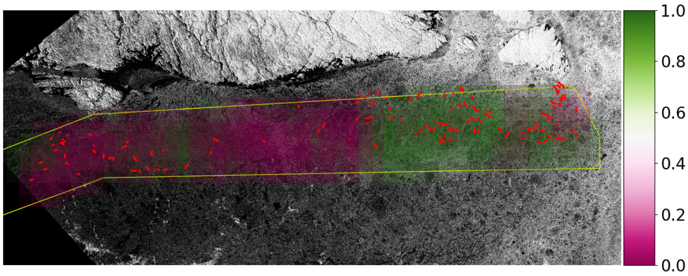

Fig. 1: Random forest prediction for one of the six holdout images. Red lines indicate labelled ridges, while yellow lines indicate the AOI. Translucent colouring corresponds to probability of a high ridge concentration existing in a patch.

A visualization of the predictions can be seen in Figure 1. The random forest predicts a high concentration of ridges where the patches appear green in the east portion of the AOI (yellow line). These green patches align with the labelled ridge locations from [1] shown in red. The model predicts a low probability of a dense ridging (red) for the rest of the AOI. These predictions are correct for the central part of the AOI where there is indeed no ridging, though in the western portion, there is some ridging. This could be due to the fact that these areas contain medium ridging (one to nine ridges per patch) and therefore are not accounted for in the training data.

4 Future Work

Due to the ambiguity of the ridging in Hudson Strait, it may be valuable to explore a positive-unlabelled or positive-confidence approach to classification. Additionally, it would be worth while testing some features used in [2] such as autocorrelation and entropy, as well as gray level co-occurrence matrix (GLCM) features. Finally, a convolutional neural network (CNN), the current state of the art in computer vision, can be used to classify the patches themselves.

Table 1: Area under receiver operating characteristic curve (AUC-ROC) scores of various models trained on an ice ridging data set.

Model	AUC-ROC Score
Logistic Regression	66
SVM	68
Random Forest	71

References

[1] O. Mussells, J. Dawson, and S. Howell, “Navigating pressured ice: Risks and hazards for winter resource-based shipping in the Canadian Arctic,” *Ocean & Coastal Management*, vol. 137, pp. 57–67, 2017.

[2] A. Gegiuc, M. Similä, J. Karvonen, M. Lensu, M. Mäkynen, and J. Vainio, “Estimation of degree of sea ice ridging based on dual-polarized C-band SAR data,” *The Cryosphere*, vol. 12, pp. 343–364, 2018.

[3] G. Heygster and C. Melsheimer, “Daily AMSR2 sea ice concentrations,” 2017. Institute of Environmental Physics, University of Bremen, Germany, `https://seaice.uni-bremen.de/data/amsre/asi_daygrid_swath/n3125/`.