POAC'25
St. John's, Newfoundland and Labrador, Canada

Proceedings of the 28th International Conference on Port and Ocean Engineering under Arctic Conditions
Jul 13-17, 2025
St. John's, Newfoundland and Labrador
Canada

Machine Learning-Based Detection of Icebergs in Sea Ice Using SAR Imagery

Zahra Jafari[1], Rocky S. Taylor[1], Ebrahim Karami[1]
[1] Memorial University of Newfoundland (MUN), St. John's, Canada

ABSTRACT

Icebergs pose significant risks to maritime operations, including shipping, offshore oil exploration, and underwater pipelines. Detecting and monitoring icebergs in sea ice, particularly in challenging environments such as the North Atlantic Ocean, where frequent cloud cover and darkness impede optical observations, is crucial for safe navigation and operational planning. This study presents a machine learning (ML)-based approach for iceberg detection within sea ice (SI) environments using C-band dual-polarimetric SAR imagery from the RADARSAT Constellation Mission (RCM), covering the east coast of Canada throughout 2022 and 2023. Our method integrates statistical texture, some climate factors, and backscatter features with high-dimensional representations extracted using various ML models, including a Vision Transformer (ViT) and three convolutional neural networks (CNNs). We compare the performance of features extracted from a scratch CNN, fine-tuned CNNs, and ViT. The combined feature set is classified using XGBoost, achieving robust differentiation between icebergs and sea ice under complex environmental conditions. The model demonstrated a high accuracy of 98.04% when utilizing features from three fine-tuned CNNs, along with statistical and climate-based factors. These findings highlight the potential of SAR-based, ML-driven iceberg detection to enhance large-scale monitoring and risk mitigation efforts in polar maritime environments.

KEY WORDS: Iceberg detection; Sea ice; Remote Sensing; SAR; Machine Learning (ML)

INTRODUCTION

The breakup of icebergs significantly contributes to the loss of mass from ice sheets and glaciers, especially in Greenland, where the annual discharge of solid ice is estimated to be approximately 500 Gt per year (Mankoff et al.,2020). A significant portion of this discharge consists of medium to large icebergs (60–220 m in length) that drift southward into the North Atlantic Ocean, reaching the east coast of Canada (Melling H et al.,2008). These icebergs, transported by ocean currents such as the Labrador Current, pose substantial risks to traffic and offshore installations in the region (Dalton, A.,2023). Additionally, sea ice (SI) is present in these waters for at least five months of the year, further complicating navigation and iceberg detection. Detecting iceberg in this environment is crucial for improving maritime safety and operational planning (Veland S et al.,2021). Icebergs in open water and SI are

commonly identified using Synthetic Aperture Radar (SAR) satellites, with C-band satellites like Sentinel-1 and RADARSAT frequently employed for this purpose (Gill,2001). Iceberg detection algorithms typically rely on the principle that icebergs exhibit higher backscatter compared to the surrounding open water (Himi et al.,2021). Constant False Alarm Rate (CFAR) detectors have been widely used to map the regional distribution of icebergs in open water, particularly in the Arctic. However, CFAR struggles to detect icebergs in SI because the intensity distributions of icebergs and SI often overlap, resulting in a high number of false alarms (FAs) (Jafari et al.,2024) To overcome these limitations, Deep learning (DL)-based algorithms have grown in popularity in different areas of remote sensing image analysis over the past decade including target detection in SAR images (jafari et al.,2023). Several studies have explored various methods for detecting icebergs in SI using C-band SAR, including those by Marino et al. (2016), Zakharov et al. (2017) and (Himi et al., 2021). However, these studies relied on RADARSAT-2 full polarimetric data, which have proven very effective in detecting icebergs from the dual-polarimetric data available in this study. Research focused on the Antarctic often employs object-based or segmentation-based approaches (Kim et al., 2020, Koo et al., 2023, Braakmann-Folgmann et al., 2023 and Evans et al., 2023), but these methods are less effective for the east coast of Canada, where smaller icebergs typically occupy only one to three pixels in RADARSAT Constellation Mission (RCM) images. In the Arctic, studies covering larger regions with diverse ice conditions, such as those by Dierking and Wesche (2014) and Soldal et al. (2019), have underscored the difficulties of reliably distinguishing icebergs from SI due to their similar backscatter characteristics. More recently, Laust Færch et al. (2024) demonstrated successful iceberg monitoring in SI using SAR imagery, but their dataset featured significantly higher resolution, which improved detection performance.

The effectiveness of SAR-based iceberg detection is influenced by multiple factors, including: (1) the physical properties of icebergs (e.g., size, shape, and structure); (2) SAR sensor characteristics (e.g., incidence angle, frequency band, resolution, and polarization); (3) geophysical conditions (e.g., wind, sea state, surface currents, and season); and (4) the backscatter properties of the surrounding SI or open water. In this study, we introduce a novel deep learning-based approach that leverages these key factors to enhance iceberg detection in SI environments. Our method integrates DL-derived high-level features with statistical features extracted from SAR backscatter intensity (HH and HV), as well as relevant geophysical parameters. By combining these diverse features, we employ ML techniques to classify SI and iceberg targets, providing a more robust and adaptive solution for iceberg detection in challenging SI conditions.

Study area and data

As one can see from *Figure 1*, the study area focuses on the East Coast of Canada, spanning approximately 45°N to 70°N latitude and from 75°W to 45°W longitude, covering parts of the Labrador Sea and Davis Strait. This critical region acts as a pathway for icebergs and SI moving from the Arctic into the North Atlantic. Icebergs, primarily originating from Greenland's glaciers and carried by the Labrador Current, often become embedded in extensive SI during the winter. The dataset for this study was collected during the SI seasons of March to August in 2022 and 2023, aiming to assess iceberg behavior and its implications for maritime and offshore operations.

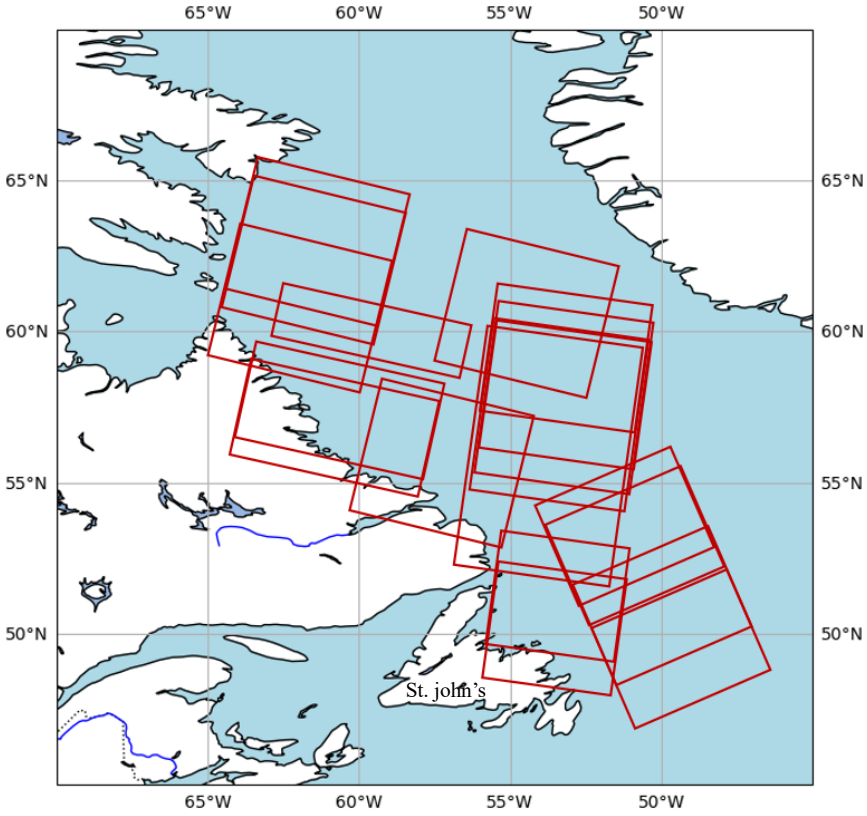

Figure 1. Overview of the study area showing the AOI and the orbits from RCM

Remote Sensing Data

A total of 102 calibrated RCM images in GeoTIFF format were collected during the 2022 and 2023 sea ice seasons. The images were acquired in medium resolution mode (50 m) with dual polarization (HH+HV). In this research, 80% of the dataset was used for training and 20% for testing. Ground truth labels were developed and validated by GIS specialists at C-CORE and further confirmed through rigorous cross-referencing with multiple independent data sources, including Sentinel-2, Landsat 8/9, and Sentinel-1. Table 1 provides a detailed summary of the datasets. Additionally, *Figure 2* presents two samples of RCM images converted to RGB format and segmented into 50×50-pixel patches (equivalent to a 1 km × 1 km area) for training and testing purposes. In *Figure 2*, the icebergs are highlighted with red circles. Although other bright pixels resembling icebergs appear in Figure 2B-D, these represent SI, illustrating the challenge of distinguishing icebergs from SI in coarse-resolution RCM images.

Table 1. Characteristics of different satellite data sources used for this study

Mode	Res. m	Looks rng×az	SW km	Nominal NESZ (dB)	Polarization mode
RCM	50	4 × 1	350	-22	HH, HV
Sentinel-1	20	5 × 1	250	-23	HH, HV
Sentinel-2	10	3 × 1	290	-	Multi-spectral (VNIR)
Landsat 8/9	15	3 × 1	290	-	panchromatic

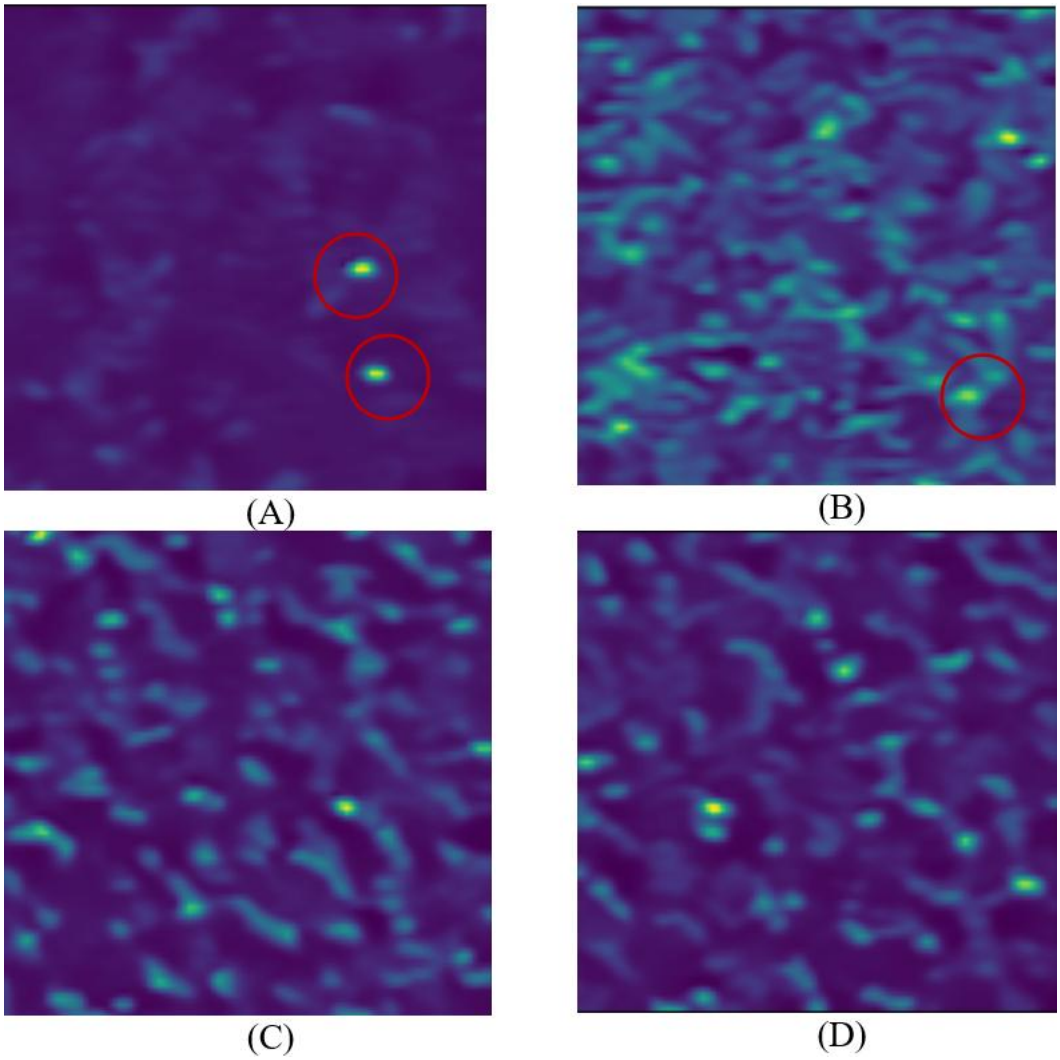

Figure 2. These figures show four sample RGB image patches (50×50 pixels in size) from the RCM dataset after normalization and denoising. Panels (A) and (B) depict iceberg in sea ice (SIT) in HH and HV, respectively, while panels (C) and (D) show SI without iceberg in HH and HV. The red circles highlight icebergs.

ERA5 Meteorological Data

The Copernicus Climate Change Service (C3S) provides ECMWF reanalysis data, including ERA5—the fifth-generation atmospheric reanalysis dataset. This study utilizes ERA5 hourly data from the period when RCM images were collected (March to August 2022 and 2023) as ground truth for climate parameters, including 2-meter temperature and 10-meter u- and v-components of wind, at a spatial resolution of 0.25° × 0.25° with hourly temporal resolution. The wind components were used to compute wind speed (magnitude) and wind direction (vector angle). Additionally, parameters such as skin temperature, mean surface latent heat flux, and mean surface sensible heat flux were extracted. These values were interpolated for each image patch based on the latitude and longitude of the patch center to enhance analysis accuracy. Despite the comprehensive nature of ERA5 data, it has limitations for near-real-time (NRT) applications. ERA5 data are subject to a latency of over one month, which renders them unsuitable for integration into operational iceberg detection systems where SAR data must be processed within hours of acquisition for timely decision-making.

Methodology

The primary objective of this study is to evaluate the capability of RCM (50 m resolution) images in detecting icebergs in different SI types. To achieve this, we employed backscatter data from HH and HV polarization for 50x50 pixels patches, between icebergs and the surrounding SI. Statistical features including Max and Min intensity, Skewness, Kurtosis, Entropy and Max-Mean were calculated from each patch and polarization (Jafari et al.2025), complemented by high-level features derived from various deep learning architectures. Regarding geophysical factors, we considered wind speed and direction, temperature, ice skin temperature, mean surface latent and sensible heat flux which represent variability in dielectric properties of iceberg and SI. Concerning sensor characteristics, the incidence angle is the only parameter that varies and was thus accounted for, while other sensor parameters, such as the frequency band, remain consistent across all data. Due to the coarse resolution of the dataset (50 m), with some icebergs occupying only 1-3 pixels, physical properties could not be effectively used. Instead, we extracted a total of 321 features from deep learning models, statistical analysis, and climate factors. These combined features were then classified using ML algorithms to enhance detection accuracy in SI environments. The system architecture is illustrated in *Figure 3*. As shown, images from the RCM dataset undergo initial preprocessing, followed by parallel feature extraction stages. The extracted features are then merged, and classification is performed using XGBoost. The model outputs probabilities across two classes: icebergs embedded in SI and SI without any iceberg.

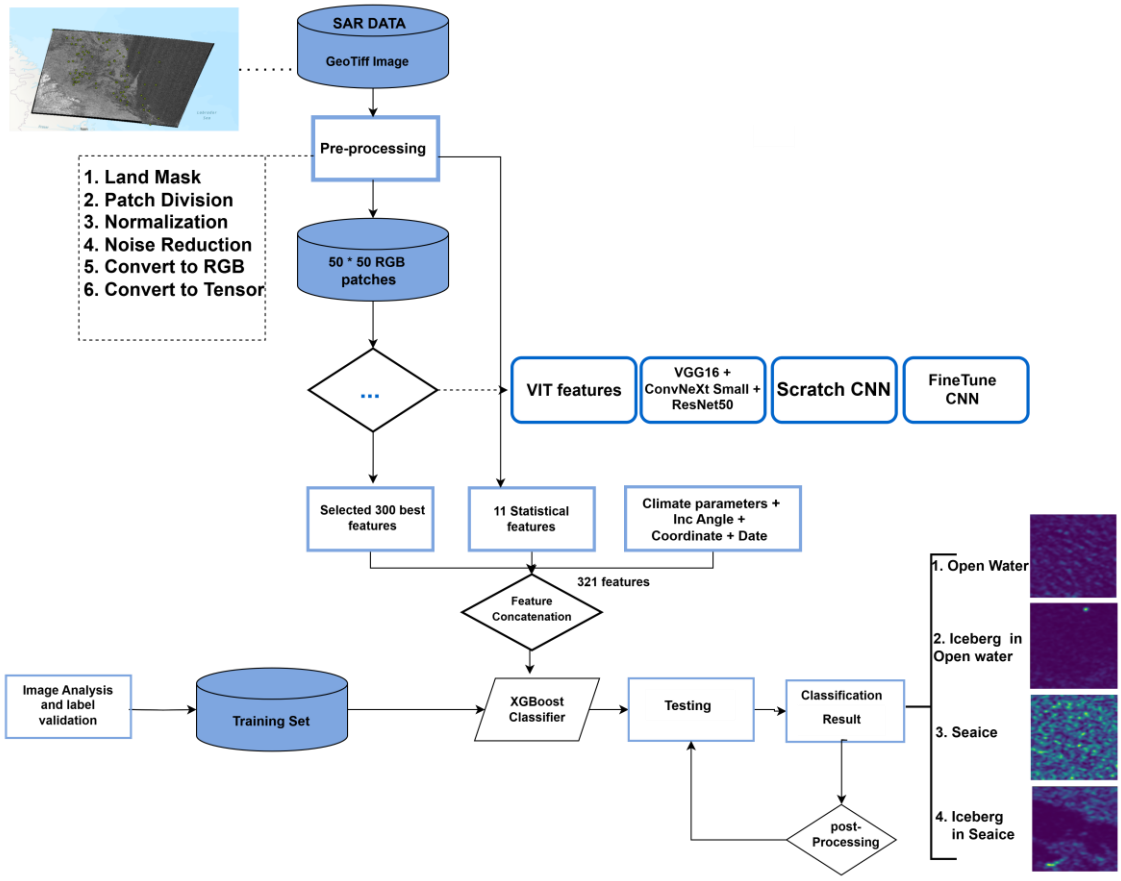

Figure 3. Block Diagram of Proposed System

Pre-processing

Our raw dataset consisted of calibrated GeoTIFF images, where channel one contained HH values, channel two contained HV values, and channel three contained incidence angle values. Each image had a size of 31,000 × 21,000 pixels. Since these images included substantial land areas, and open water, land and open water masks were initially applied to exclude non-ice regions. In the next step, the image is divided into 50×50-pixel patches. The algorithm aims to classify these ice patches whether they are SI or SI with Iceberg (SIT). The preprocessing approach varied depending on the feature extraction method used. For feature extraction with CNN models, normalized despeckling filtered RGB images were required (jafari et al.,2023), whereas the additional statistical features were extracted directly from the raw values of HH and HV with converting to dB. Using raw values ensures that critical signal characteristics essential for classification were preserved. *Figure 4* shows how different statistical features distinct SI from SIT.

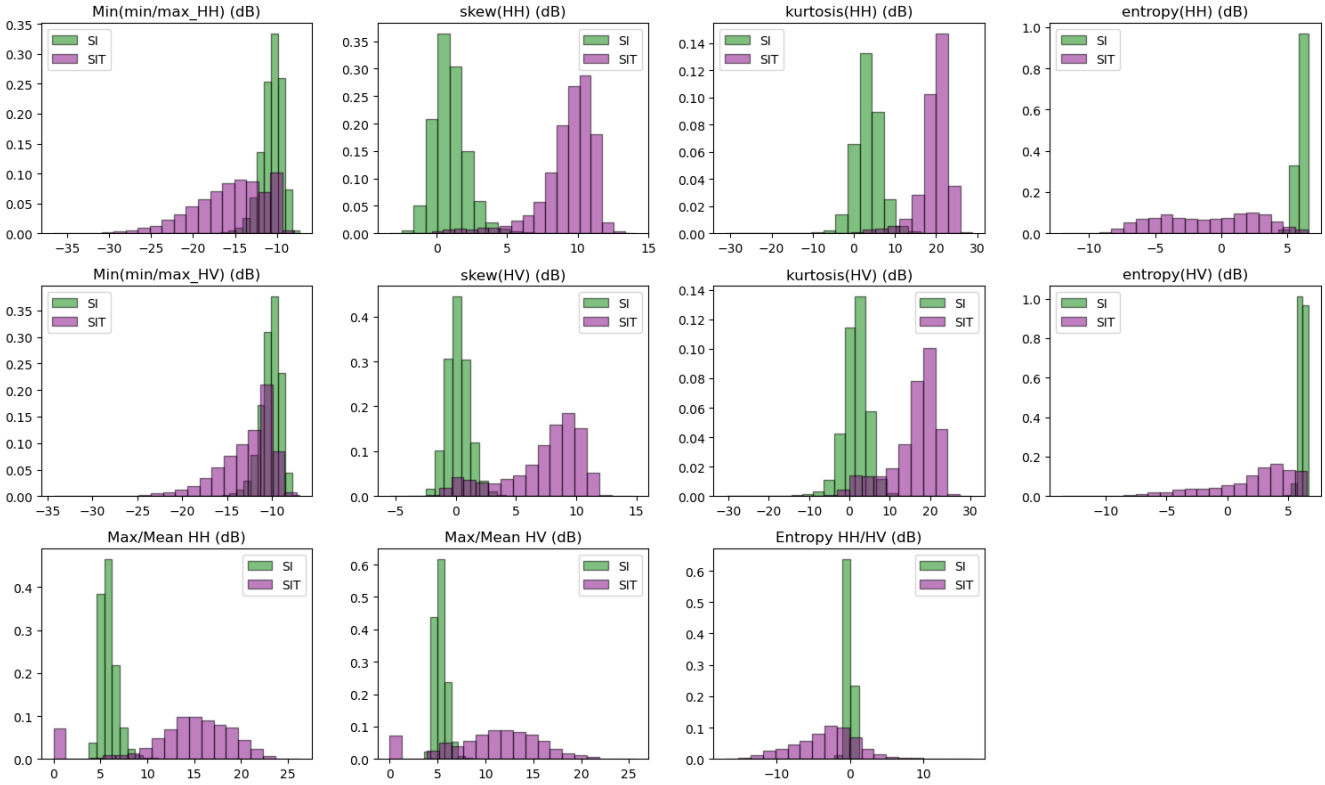

Figure 4. Histograms of selected statistical features extracted from HH and HV polarizations, including the global minimum of the ratio of the minimum to maximum intensity within each 4 × 4 non-overlapping window, skewness, kurtosis, entropy, and derived ratios. These features exhibit distinct distributions for Sea Ice (SI) and Sea Ice with Target (SIT).

Result and Discussion

This section presents the results obtained from the RCM dataset, introduced in the Remote Sensing Data section, using the models described in the Methodology section, as well as a fusion of all models, as illustrated in Figure 3. The proposed method was implemented in Python (version 3.11.9) using the TensorFlow package (version 2.11). To assess and compare the performance of the proposed models, we employed multiple evaluation metrics. These metrics provide a comprehensive understanding of the models' strengths and limitations in detecting icebergs on SI.

Table 2. Performance Evaluation of Models Using 5000 RCM Patches size 50x50 from 2022 and 2023 with 5000 actual icebergs.

Models	TP	FP	Precision	Recall	Accuracy
Scratch CNN	4,000	0.07	80	83	83.15
Finetune CNN	4,250	0.05	85	85	85.02
ViT_Features	4,600	0.05	91	92	92.50
Statistical Features	4,700	0.03	95	93	94.69
3 CNN Features	4,750	0.03	95	96	95.80
Clim_Stat_3CNN_F	4,950	0.02	98	98	98.04

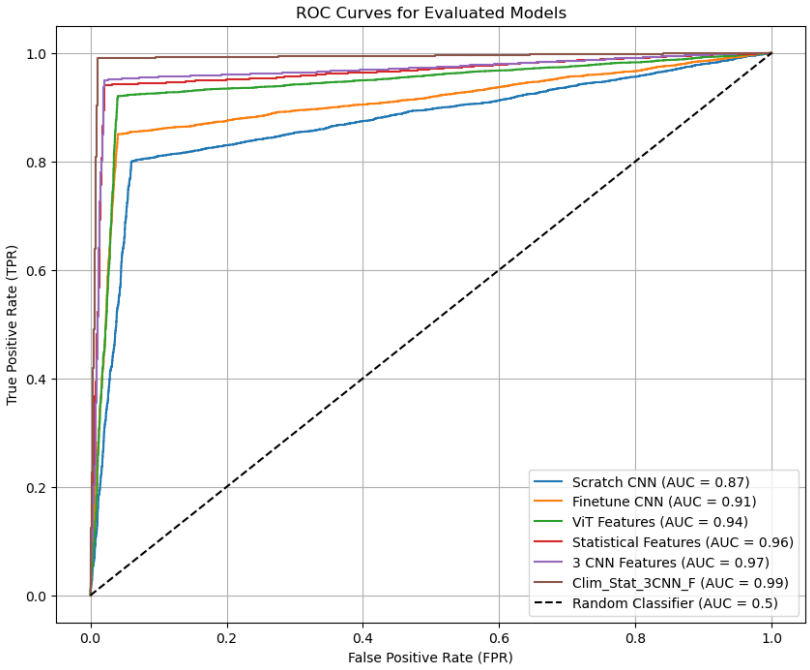

Figure 5. ROC Curves for the Evaluated Models: (A) Scratch CNN, (B) Finetune CNN, (C) ViT Features, (D) Statistical Features, (E) 3 CNN Features, and (F) Clim_Stat_3CNN_F. The curves illustrate the classification performance across different ice detection scenarios using 5000 RCM patches (50x50) from 2022 and 2023.

Table 2 and *Figure 5* present a comprehensive evaluation of six models for classifying icebergs from SI using 5,000 RCM patches (50×50) with actual icebergs, collected during the 2022 and 2023 ice seasons. This comparison highlights each model's contribution to classification accuracy and efficiency. Given a large dataset of 250,000 augmented image patches (50×50) extracted from 102 TIFF images, we began by training a CNN model from scratch. The model architecture included multiple convolutional layers, each followed by ReLU activations, batch normalization, and max-pooling layers to reduce spatial dimensions. The final classification was performed using fully connected layers and a softmax output. This model achieved a true positive (TP) score of 4,000 with an accuracy of 83.15%. The relatively lower performance suggests that, despite the large dataset, training from scratch without pre-trained features limits classification effectiveness. To enhance performance, transfer learning was applied by fine-tuning three CNN (VGG16, ResNet50, and ConvNeXt-Small), as they have proven effective in capturing the shape and texture of icebergs (Jafari et al., 2023). By fine-tuning the last convolutional layers while retaining pre-trained knowledge, the model's performance improved, yielding a TP score of 4,250 and an accuracy of 85.02%. This improvement over the scratch CNN highlights the advantage of leveraging pre-trained models. Next, we evaluated the ViT-B/16 model, a Vision Transformer (ViT) variant. By segmenting images into patches and applying transformer architectures to capture long-range dependencies, this model generated a 768-dimensional feature vector for each image. Although the ViT model achieved a strong TP score of 4,600 with an accuracy of 92.50%, it was slightly outperformed by the statistical feature-based model and CNN ensembles. This suggests that CNNs may be better suited for capturing localized spatial patterns in satellite imagery. Beyond deep learning, we explored a statistical feature-based model for classifying

icebergs from SI. Remarkably, this model achieved a TP score of 4,700 with an accuracy of 94.69%, demonstrating that well-engineered statistical features alone possess significant discriminatory power. Recognizing the complementary strengths of CNN models, we combined the last-layer features from the three pre-trained CNNs, resulting in a large feature set. Using a feature engineering technique based on mutual information, we reduced the set to the 300 most informative features, enhancing classification performance. Finally, the Clim_Stat_3CNN_F model, which combined three feature types—300 selected CNN features, statistical descriptors, and climate factors—achieved the highest performance. This model attained a TP score of 4,950, an accuracy of 98.04%, and precision and recall scores of 98% each.

Analysis of misclassification result

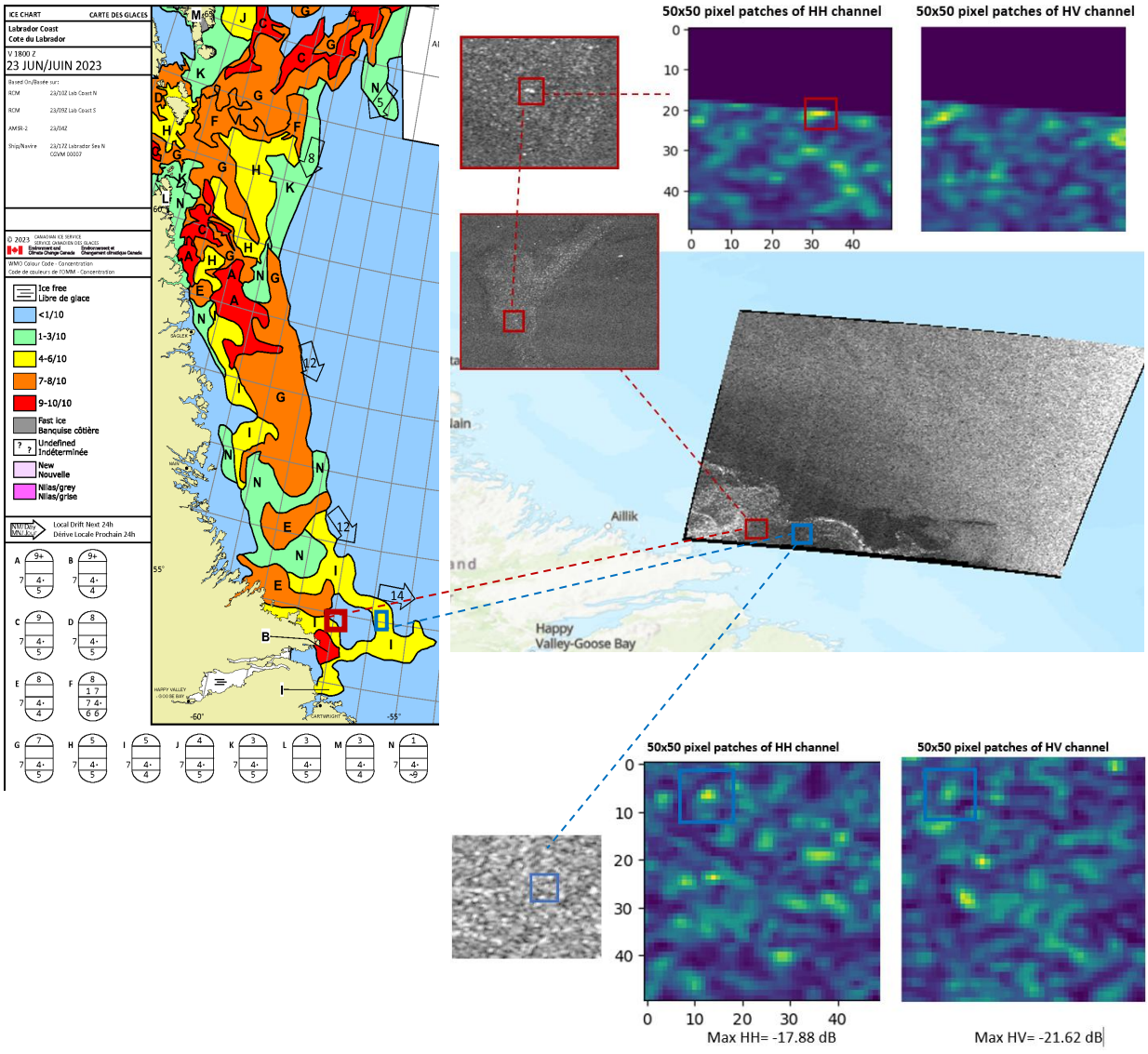

Figure 6 . Sample of RCM image captured on June 23, 2023, for analyzing missed detection icebergs in sea ice. The figure includes an ice concentration map, SAR imagery, and extracted 50×50-pixel patches from HH and HV polarization channels to investigate classification performance.

After analyzing the missed classification results, we selected a representative image which capture on June 23,2023, to examine the locations of the undetected icebergs. As shown in Figure 6, the red rectangle highlights cases where icebergs located at the boundary between SI and open water were more likely to be missed. This suggests that target detection becomes challenging for the model in transitional regions. Additionally, as indicated by the blue rectangle, another limitation of SAR imagery is evident when the maximum HH and HV backscatter values are weak. Some icebergs in this region had backscatter values below the noise equivalent sigma zero (NESZ), and also, when the maximum HH was below -15 dB and the maximum HV was below -20 dB, the model failed to detect the targets. Consequently, these icebergs remained undetected. Furthermore, in locations where false alarms occurred, the misclassified targets were often thick first-year sea ice, which typically appears in large floes and was mistakenly identified as icebergs.

CONCLUSIONS

Detecting icebergs in sea ice (SI) presents a significant challenge due to the high frequency of false alarms (FAs) produced by traditional methods, such as CFAR-based algorithms. In this study, we explored several innovative machine learning (ML) techniques for iceberg detection using medium-resolution (50 m) RCM images that function effectively in environments. To do so, we extracted various statistical features and climate data from 50×50-pixel image patches after pre-processing steps, then enhanced them with 300 highly informative features derived from different deep learning architectures, classifying them using the LightGBM algorithm. The findings indicate that the hybrid model, which integrates CNN features, statistical data, and climate parameters, performed the best in distinguishing SI from icebergs with accuracy of 98.04% and an AUC close to 1. A CFAR-based post-processing step was applied afterward to minimize false alarms. We aim to implement further post-processing steps to reduce FAs. These results validate the effectiveness and reliability of our method for large-scale iceberg detection in complex settings, particularly along Canada’s east coast. This research provides a strong foundation for advancing iceberg monitoring and improving maritime safety in regions prone to iceberg hazards.

ACKNOWLEDGEMENTS

This research was funded by Equinor, and we are truly grateful for their support. We would also like to express our thanks to Maria Yulmetova, a GIS Specialist at C-CORE, for providing the calibrated RCM data, which was crucial to our analysis. Furthermore, we appreciate Mark Howell for his invaluable help, which significantly improved the results of this project. We also acknowledge the Government of Canada for providing RADARSAT Constellation Mission (RCM) imagery (2022 and 2023) used in this study.

REFERENCES

Braakmann-Folgmann, A., Shepherd, A., Hogg, D. and Redmond, E., 2023. Mapping the extent of giant Antarctic icebergs with deep learning. *The Cryosphere*, *17*(11), pp.4675-4690.

Dalton, A., 2023. Identifying Iceberg Production Processes, Drift Patterns, and Coexistence with Ships in the Eastern Canadian Arctic (Doctoral dissertation, Université d'Ottawa/University of Ottawa).

Evans, B., Faul, A., Fleming, A., Vaughan, D.G. and Hosking, J.S., 2023. Unsupervised machine learning detection of iceberg populations within sea ice from dual-polarisation SAR imagery. *Remote Sensing of Environment*, *297*, p.113780.

Færch, L., Dierking, W., Hughes, N. and Doulgeris, A.P., 2024. Mapping icebergs in sea ice: An analysis of seasonal SAR backscatter at C-and L-band. *Remote Sensing of Environment*, *304*, p.114074.

Gill, R.S., 2001. Operational detection of sea ice edges and icebergs using SAR. *Canadian journal of remote sensing*, *27*(5), pp.411-432.

Himi, U.H., Ferdous, M.S., Power, D.T. and McGuire, P., 2021. Statistical comparison of melting iceberg backscatter embedded in sea ice and open water using RADARSAT-2 images of the Newfoundland sea. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, pp.1-13.

Jafari, Z., Bobby, P., Karami, E. and Taylor, R., 2024. A Novel Method for the Estimation of Sea Surface Wind Speed from SAR Imagery. *Journal of Marine Science and Engineering*, *12*(10), p.1881.

Jafari, Z., Karami, E., Taylor, R. and Bobby, P., 2023. Enhanced ship/iceberg classification in sar images using feature extraction and the fusion of machine learning algorithms. *Remote Sensing*, *15*(21), p.5202.

Kim, M., Kim, H.C., Im, J., Lee, S. and Han, H., 2020. Object-based landfast sea ice detection over West Antarctica using time series ALOS PALSAR data. Remote Sensing of Environment, 242, p.111782.

Koo, Y., Xie, H., Mahmoud, H., Iqrah, J.M. and Ackley, S.F., 2023. Automated detection and tracking of medium-large icebergs from Sentinel-1 imagery using Google Earth Engine. Remote Sensing of Environment, 296, p.113731.

K.D. Mankoff, A. Solgaard, W. Colgan, A.P. Ahlstrøm, S. Abbas Khan, R.S. Fausto
Greenland ice sheet solid ice discharge from 1986 through March 2020
Earth Syst. Sci. Data, 12 (2020), pp. 1367-1383, 10.5194/essd-12-1367-2020

Marino, A., Dierking, W. and Wesche, C., 2016. A depolarization ratio anomaly detector to identify icebergs in sea ice using dual-polarization SAR images. IEEE Transactions on Geoscience and Remote Sensing, 54(9), pp.5602-5615.

Melling H, Agnew TA, Falkner KK, Greenberg DA, Lee CM, Münchow A, Petrie B, Prinsenberg SJ, Samelson RM, Woodgate RA. Fresh-water fluxes via Pacific and Arctic outflows across the Canadian polar shelf. Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate. 2008:193-247.

Soldal, I.H., Dierking, W., Korosov, A. and Marino, A., 2019. Automatic detection of small icebergs in fast ice using satellite wide-swath SAR images. Remote Sensing, 11(7), p.806.

Veland, S. and Wagner, P.M., 2021. Knowledge needs in sea ice forecasting for navigation in Svalbard and the High Arctic.

Wesche, C. and Dierking, W., 2014, July. From ice shelves to icebergs: Classification of calving fronts, iceberg monitoring and drift simulation. In *2014 IEEE Geoscience and Remote Sensing Symposium* (pp. 274-277). IEEE.

Zakharov, I., Power, D., Howell, M. and Warren, S., 2017, July. Improved detection of icebergs in sea ice with RADARSAT-2 polarimetric data. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) (pp. 2294-2297). IEEE.