Ocean Target Discrimination in SAR Imagery through Machine Learning: Towards a Fully Automated Approach

Maria Yulmetova, Murilo Silva, Pradeep Bobby, Kelley Dodge

Centre for Cold Ocean Resources Engineering (C-CORE), 1 Morrissey Rd, St. John's, NL, Canada
(maria.yulmetova; murilo.silva; pradeep.bobby; kelley.dodge)@c-core.ca

Keywords: Icebergs and Ships; Classification; False Alarms; Machine Learning.

Abstract

Accurate discrimination of ocean targets using satellite images is crucial for marine safety, environmental monitoring, dark vessel detection, and search and rescue operations. Artificial intelligence technologies are rapidly advancing as state-of-the-art solutions for computer vision problems, including satellite imagery target classification. This research assesses the capability of machine learning (ML) for ocean target discrimination using SAR images. Unlike other studies focusing on binary iceberg-ship classification, this paper goes a step further to investigate the opportunity for multi-class discrimination between icebergs, ships, and false alarms, both within and outside sea ice. The proposed approach enables the fully automated elimination of false alarms while accurately classifying icebergs and ships. As part of a research initiative, the first large dataset of ocean targets was compiled and utilized to train an ML model. The targets were detected in RADARSAT Constellation Mission (RCM) images over Canadian waters. During the evaluation phase, the model achieved classification accuracies of 93% for binary classification and 95% for three-class discrimination. The robustness of the fully automated approach was further validated through an additional test, yielding an overall accuracy of 91%. Moreover, the system exhibited high reliability in reducing false alarms, correctly identifying 96% of them. The implementation of the developed algorithms significantly enhances the efficiency of target detection and classification processes, thereby reducing the workload of human analysts. Such advancements are especially significant in light of the rapidly increasing volume of satellite data and the growing demand for automated, scalable solutions in maritime surveillance.

1. Introduction

The detection and classification of ships and icebergs are vital for environmental protection, maritime safety, and operational efficiency. Timely iceberg monitoring is crucial for preventing collisions and ensuring safe operations at offshore oil facilities. In addition, accurate classification helps detect "dark ships" - vessels operating without proper tracking systems - thereby limiting illegal activities such as overfishing and smuggling, promoting national security, and reducing environmental hazards.

Satellites can access remote areas at no extra cost, making them an important tool alongside marine radar and aerial reconnaissance. Their role is especially prominent in exploration projects where aerial reconnaissance infrastructure is not yet established. Synthetic aperture radar (SAR) is an active imaging system operating day and night in all weather conditions. It is widely used for vessel detection, particularly in challenging environments with frequent cloud cover, fog, or adverse weather, where optical sensors are limited. Satellite data allows for identifying and tracking ocean targets in near real-time, enabling proactive measures to facilitate rapid response to incidents.

Among existing approaches, the cell-averaging constant false alarm rate (CFAR) method (Finn & Johnson, 1968) remains the most widely used method for ocean target detection using SAR data (Ai et al., 2021; G. Gao et al., 2009; S. Gao & Liu, 2022; Power et al., 2001; Wackerman et al., 2001; C. Wang et al., 2008). It dynamically adjusts the detection threshold based on a local noise estimation. It assumes that target intensity is statistically higher than sea clutter, but this assumption often fails in real-world, nonhomogeneous sea conditions (Xie et al., 2022). SAR speckle noise complicates target detection (Passah et al., 2023), while patches of sea ice, ocean clutter, and other ambiguities contribute to increased false alarm rates and degraded detection performance in complex offshore environments (J. Li et al., 2022).

Machine learning (ML) is widely used in computer vision for detection, classification, and recognition (Y. Li et al., 2017; Passah et al., 2023; Pei et al., 2018; Redmon et al., 2016; Soldin, 2018). However, its application to SAR data is limited by the scarcity of large, labelled datasets. While several datasets are available, most contain only ship data from high-resolution SAR images, such as Gaofen-3, TerraSAR-X, RADARSAT-2, and Sentinel-1 (J. Li et al., 2022; Z. Wang et al., 2019). The lack of iceberg data has hindered ship-iceberg discrimination studies, with only a few conducted before 2018 (Bentes et al., 2016; Howell et al., 2006). The Statoil/C-CORE Iceberg Classifier Challenge (Kaggle, 2018) introduced a larger dataset with 1,604 targets from Sentinel-1, enabling the development of a wide range of deep learning algorithms (Heiselberg, 2020; Yang & Ding, 2020). However, it remains insufficient for robust model training, leading researchers to use data augmentation techniques. Yet, such techniques must be applied cautiously to SAR data, as transformations like flipping or rotation can alter polarimetric properties and distort target characteristics.

Automatic algorithms struggle to reliably distinguish real ships and icebergs from false alarms that are caused by strong waves, atmospheric effects, sea ice patches, and SAR ambiguities (Greidanus et al., 2017; Vespe & Greidanus, 2012; Y. Wang et al., 2015). To reduce false alarms, the detection threshold parameter is typically increased (Greidanus et al., 2017; Santamaria et al., 2015). As a result, many smaller targets are missed, risking maritime safety. Therefore, manual verification remains essential, making detection semi-automatic. With advancing technology, there is an urgent need to explore new methods to further enhance early detection and classification.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-M-1-2026-87-2026 | © Author(s) 2026. CC BY 4.0 License.

This study explores the capability of ML to perform binary and multi-class classification, evaluating its effectiveness in accurately distinguishing between ocean targets. Also, we aim to develop a fully automated approach that minimizes false alarms while accurately classifying targets. These advancements are expected to reduce manual effort and enhance operational efficiency.

2. Data

C-CORE has studied satellite-based iceberg monitoring for over 30 years and has delivered services across Arctic, sub-Arctic, and Antarctic waters. The launch of the RADARSAT Constellation Mission (RCM) has significantly increased the frequency of SAR observations, enabling image acquisition on a daily basis and, in some cases, even twice per day. This facilitated the creation of a large dataset. For this research, we utilized a portion of this dataset collected offshore Newfoundland and Labrador (NL).

Targets were detected using an iceberg detection software (IDS) developed by C-CORE. The detection algorithm is based on the classic CFAR approach. Each detected target underwent manual quality control to remove false positives and classify icebergs and vessels. To enhance decision-making, analysts utilized all available resources, including Sentinel-2, Landsat-8/9, AIS data, along with information obtained through direct communication with other companies involved in iceberg monitoring.

The majority of targets were detected during the 2022 and 2023 iceberg seasons from March to September, with 1,057 vessel targets collected in 2024 to address class imbalance. Targets within a 200 m buffer from the land mask were excluded to avoid mislabelling, particularly to prevent small rocks from being misidentified as icebergs. Icebergs larger than 500 m were also excluded, as they are considered Ice Island Fragments (IIFs), which exhibit distinct SAR signatures and are not the focus of this study.

For binary classification, the dataset was restricted to open-water targets. In total, 301 RCM images were processed, yielding 10,513 labelled targets, whereas 7,487 are icebergs and 3,026 are vessels. The dataset was analysed based on size distribution using the scale defined by Environment and Climate Change Canada (2018) for iceberg classification. According to this scale, icebergs are classified as small (15–60 m), medium (61–120 m), large (121–200 m), and very large (>200 m). This scale was applied to all target classes for standardized analysis. Figure 1 shows the spatial distribution of icebergs and vessels by size, with small and medium targets being more prevalent. Icebergs were mainly distributed in NL waters, with larger ones found in northern Labrador. In contrast, vessels were concentrated in the southern area.

For multi-class classification, additional data were collected for the following categories: Icebergs in Sea Ice, Vessels in Sea Ice, and False Alarms, which include

- Sea Ice - ice-covered areas,
- Noise - clutter caused by ocean surface roughness, atmospheric effects, or wave interactions, and
- Artifacts - ambiguities frequently appearing as long bright streaks inherent in SAR imagery.

In total, 55910 targets were collected, comprising 7487 Icebergs, 2782 Icebergs in Sea Ice, 3026 Vessels, 75 Vessels in Sea Ice, 7552 Noise targets, 33940 Sea Ice targets, and 1049 Artifacts. Similar to the binary classification dataset, the multi-class dataset exhibits an uneven distribution of targets by size and type, as shown in Figure 2. Notably, a major portion of the dataset consists of small and medium-sized targets, making the classification task more challenging.

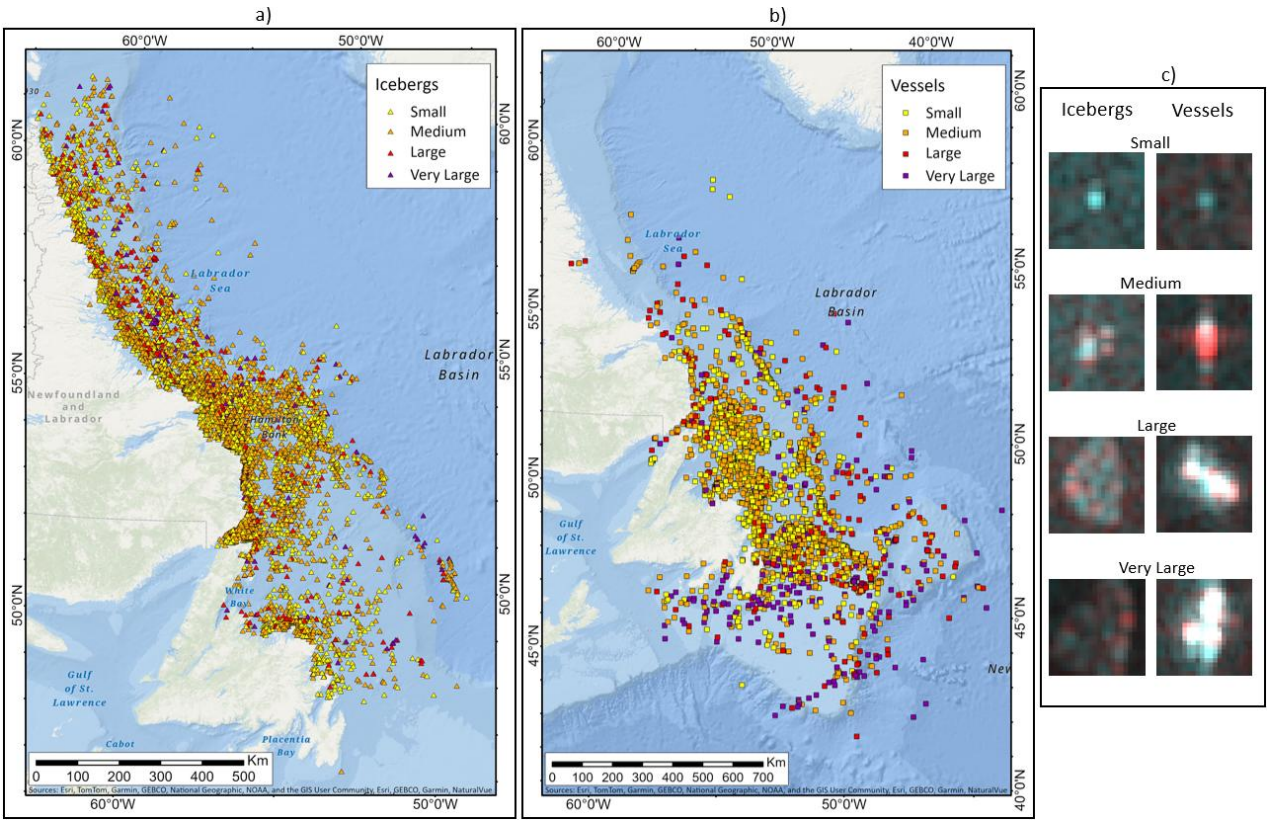

Figure 1. Spatial distribution of icebergs (a) and vessels (b) by size, along with some example targets (c).

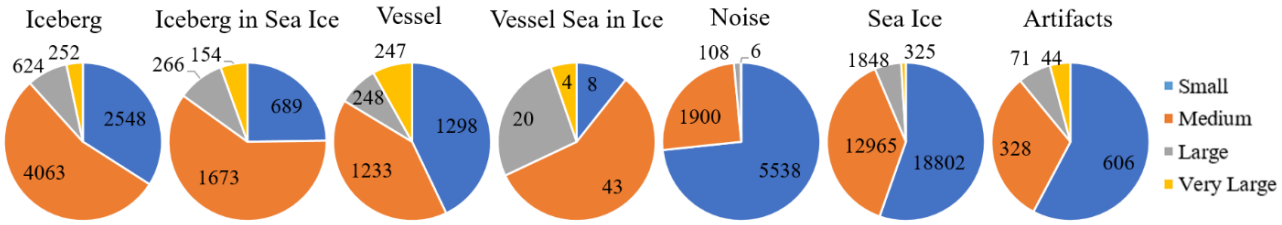

Figure 2.The distribution of all targets by size.

3. Method

The main steps of the methodological workflow implemented in this project are shown in Figure 3. First, stratified random sampling was applied to address the unbalanced dataset, ensuring proper representation of each target type and size. This is particularly important given that SAR signatures of targets vary significantly with size, as demonstrated in Figure 1c. The dataset was divided into two and seven groups for the two-class and seven-class classifications, respectively, based on class and size, and samples were drawn proportionally from each stratum.

In this study, the YOLOv8-cls model, developed by Ultralytics (2024), was selected as the ML framework for ocean target discrimination. This model was chosen for research purposes due to its proven performance and user-friendly implementation. YOLOv8's architecture consists of three components: the backbone for extracting multi-scale features, the neck for aggregating feature maps, and the head for making final predictions, including object classification and localization (Hidayatullah et al., 2025). YOLOv8-cls, tailored for classification tasks, processes input images to assign them to predefined categories without predicting bounding boxes or segmenting objects. The neck and detection head components are omitted, resulting in a simplified structure optimized for classification. The model takes fixed-size images and produces probabilities for each class. 80% of our dataset was used for training, while 20% was used for validation.

Several factors were investigated to create accurate chips. First, the optimal chip size was established based on the analysis of the targets' dimensions in relation to their waterline length. The largest and smallest targets in our dataset were AIS-verified vessels with lengths of 589 m and 11 m, respectively. We aimed to create chips sufficiently large to encompass targets while remaining small enough to exclude unnecessary background that could lead to misclassification.

Several band combinations were also examined to create three-channel input images. The HH/HV/HV combination, used as the red, green, and blue channels, demonstrated the best results. Sample chips prepared for input into the YOLO classification model are shown in Figure 4.

Since SAR images retain directional-specific information, flipping and rotation might alter meaningful information. Therefore, the flip, rotate, and scale parameters were set to zero.

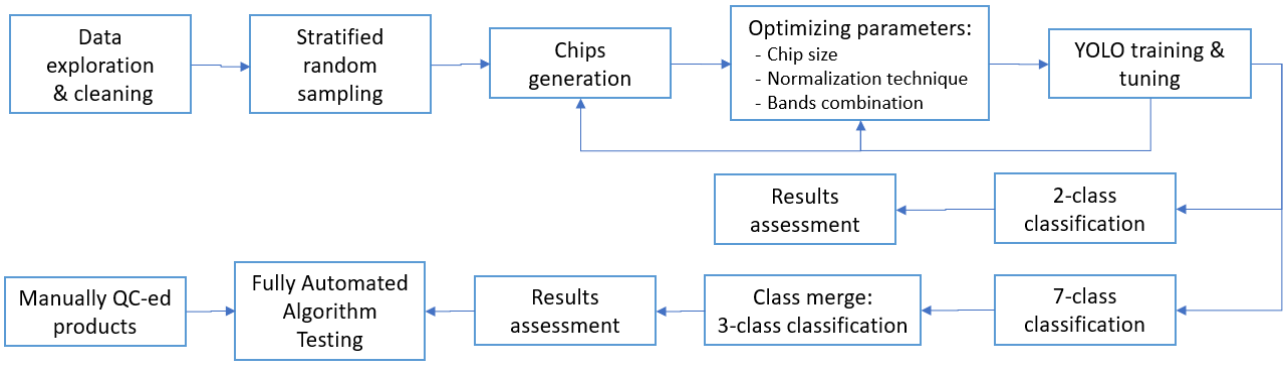

Figure 3. Workflow implemented in this project.

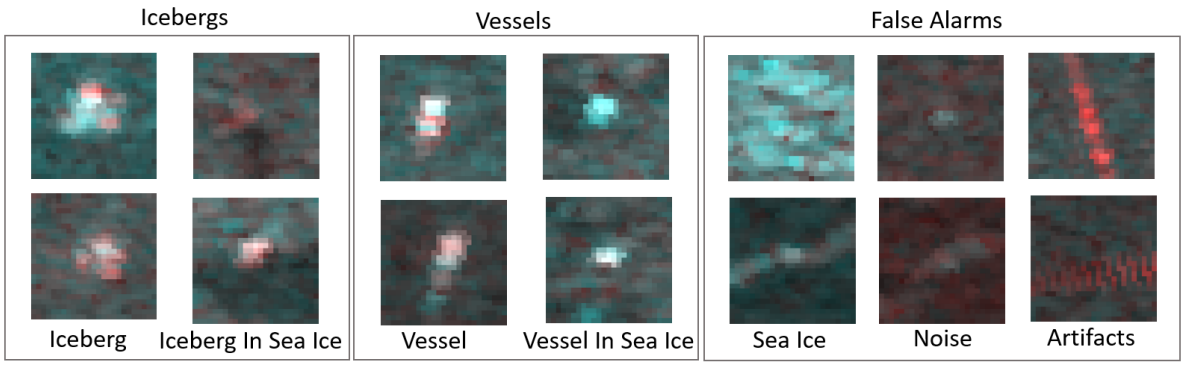

Figure 4. Sample chips.

It is worth noting that the noise removal step was skipped in our study. Although denoising is considered one of the most common preprocessing steps typically applied to SAR images (Passah et al., 2023), it can also cause small targets to be lost and become invisible in SAR images. Since most targets in our dataset are small and medium, chips were extracted from raw RCM images. This aligns with other studies (Yang & Ding, 2020), which found that deep learning performs worse on despeckled data compared to raw images. Their findings also suggest that the denoising process may discard useful features around ocean targets that neural networks could otherwise learn to improve classification results.

Finally, the classifiers were trained and evaluated for performance. A binary classifier was trained to distinguish between icebergs and vessels, while the multi-class classifier included seven categories, which were subsequently consolidated into three broader categories:

- Icebergs class, combining Iceberg and Iceberg in Sea Ice to represent all iceberg targets;
- Vessels, integrating Vessel and Vessel in Sea Ice to account for all vessel detections; and
- False Alarms, merging Noise, Sea Ice, and Artifacts to group all non-iceberg and non-vessel detections.

Model performance was evaluated using confusion matrices, overall accuracy (OA), along with F1 score and Cohen's Kappa to account for class imbalance. Stratified 10-fold cross-validation was applied to ensure reliable model evaluation and reduce the risk of over- and underfitting. Finally, our fully automated method was tested using an image-based approach. The performance results are discussed in the next section.

4. Results

4.1 Classification Assessment

The binary classification model achieved an OA of 92.42%, correctly identifying 93% of icebergs and 90% of vessels (Table 1). The iceberg class exhibits slightly better precision (95.8%) and recall (93.4%) compared to vessels (84.6% and 90.1%, respectively). The model effectively minimizes false positives and false negatives, with only 6.6% of icebergs misclassified as vessels and 9.97% of vessels misclassified as icebergs. The attained Kappa value of 0.82 reflects a substantial agreement between the targets (Landis & Koch, 1977). The final overall score obtained from the 10-fold cross-validation was equal to 0.926 with a standard deviation of 0.006.

True \ Predicted	Iceberg	Vessel
Iceberg	**0.93**	0.07
Vessel	0.10	**0.90**
OA = 0.92 F1 = 0.92 Kappa = 0.82		
10-fold Cross Validated OA = 0.926 ± 0.006		

Table 1. Normalized 2-class classification confusion matrix.

The 7-class classification model achieved an OA of 89.57% (Table 2). The highest prediction scores were obtained for the Sea Ice and Artifacts classes, while the poorest performance was observed in classifying Vessels in Sea Ice, most of which were misclassified as Iceberg in Sea Ice. A strong balance between precision and recall was also observed, which is reflected in the high F1 score of 0.89. Kappa coefficient of 0.82 demonstrates a substantial agreement between the targets. The 10-fold cross-validation score achieved 89.96% with a standard deviation of 0.32%.

The original seven classes were subsequently merged into three final categories: Icebergs, Vessels, and False Alarms. The final confusion matrix obtained after the class merge is presented in Table 3. The model achieved an OA of 94.50% with the high F1 and Kappa scores of 0.94 and 0.86. The final overall score obtained from the 10-fold cross-validation was equal to 0.95 with a standard deviation of 0.004. The algorithm effectively filters out false alarms, resulting in only 2.8% false targets being incorrectly labelled as icebergs and 0.5% as vessels.

4.2 Fully Automated Algorithm Testing

The performance of the developed fully automated approach was tested using an image-based approach. For this assessment, an RCM image acquired on July 19, 2023, was selected. The image was not utilized for training or validation, contained no sea ice, and was captured under wind speeds of approximately 12–16 knots at the time of acquisition.

C-CORE's IDS software was used for target detection step, followed by seven-class classification. Then, the results were grouped into three final categories: Icebergs, Vessels, and False Alarms. The automatically generated classification products were then compared against manually quality-controlled (QC-ed) reference data produced by analysts.

True \ Predicted	Iceberg	Iceberg in Sea Ice	Noise	Sea Ice	Artifacts	Vessel	Vessel in Sea Ice
Iceberg	**0.86**	0.03	0.04	0.03	0	0.04	0
Iceberg in Sea Ice	0.20	**0.56**	0.02	0.21	0	0.01	0
Noise	0.03	0.00	**0.82**	0.12	0.01	0.02	0
Sea Ice	0.01	0.01	0.03	**0.95**	0	0	0
Artifacts	0	0	0	0	**1**	0	0
Vessel	0.09	0.00	0.03	0.01	0	**0.86**	0.003
Vessel in Sea Ice	0.23	0.38	0.00	0.08	0	0.31	**0**
OA = 0.90 F1 Score = 0.92 Kappa Coefficient = 0.82							
10-fold Cross Validated OA = **0.90 ± 0.003**							

Table 2. Normalized 7-class classification confusion matrix.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-M-1-2026-87-2026 | © Author(s) 2026. CC BY 4.0 License.

Figure 5 presents the classification results, including all targets detected by the CFAR algorithm alongside the iceberg and vessel classification maps generated using the fully automated approach and manual QC. In both classification maps, targets identified as False Alarms were removed. The algorithm successfully identifies False targets, including Artifacts (Figure 5d and e), and Noise targets caused by internal waves and atmospheric effects (Figure 5f and g) and the typical 'salt and pepper' noise often present in SAR images (Figure 5h). In addition to visual assessment, confusion matrices were generated to evaluate classification accuracy quantitatively (Table 4 and 5). The model achieved an overall accuracy (OA) of 0.91, with false targets successfully identified in 96% of cases.

True \ Predicted	Icebergs	Vessels	False
Icebergs	**0.86**	0.03	0.11
Vessels	0.10	**0.86**	0.04
False	0.02	0.005	**0.97**
OA = 0.95 F1 Score = 0.94 Kappa = 0.86			
10-fold Cross Validated OA = **0.95 ± 0.004**			

Table 3. Normalized confusion matrix after class merge.

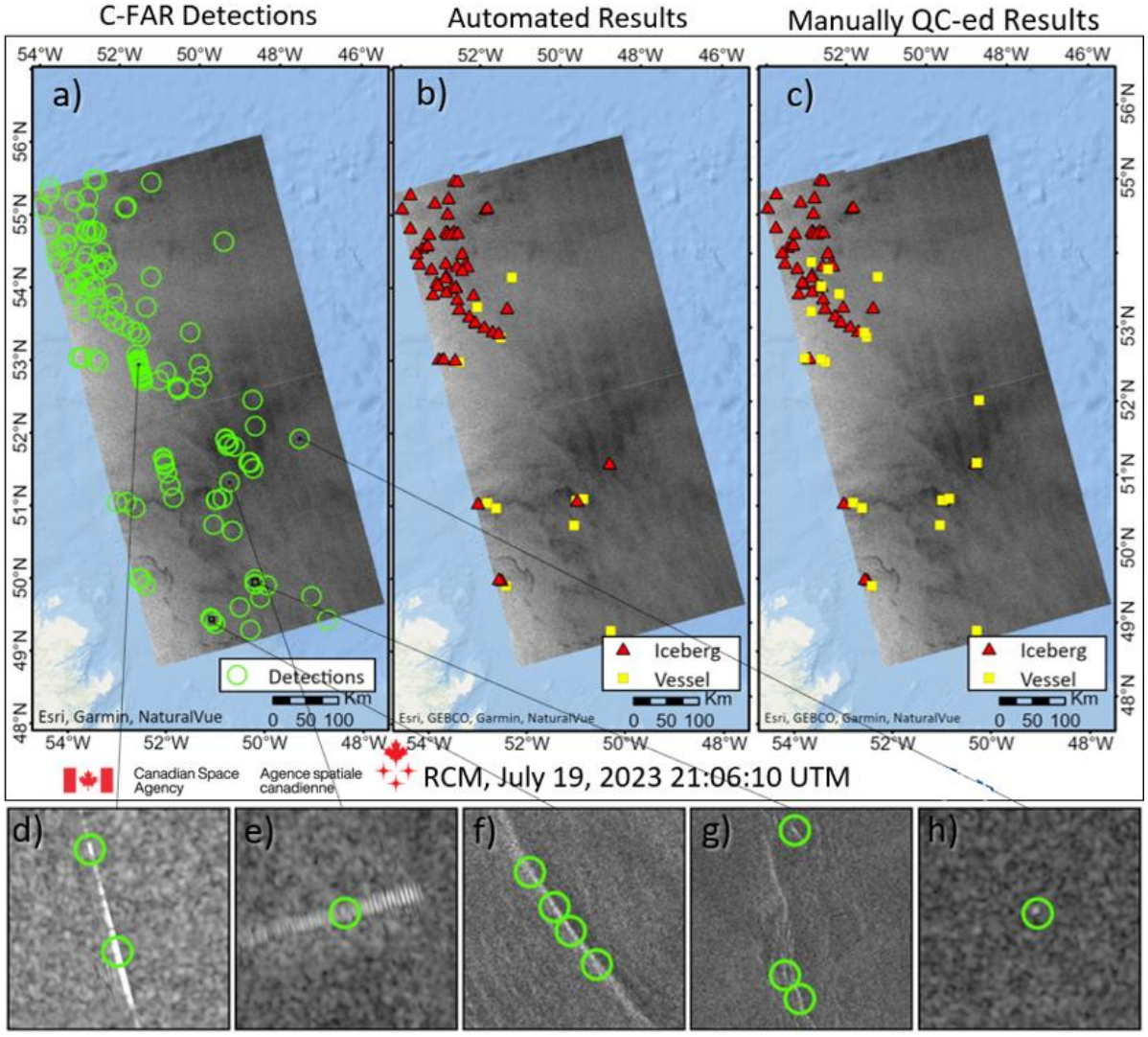

Figure 5. Initial targets detected by IDS (a), the iceberg and vessel classification maps created using the fully automated approach (b) and after manual QC (c). Examples of successfully identified False Alarms shown in (d – h).

Predicted \ True	Icebergs	Vessels	False	Row Total
Icebergs	**52**	8	1	61
Vessels	1	**10**	0	11
False	0	3	**70**	73
Column Total	53	21	71	**145**

Table 4. Confusion matrix from a classified test image.

Predicted \ True	Icebergs	Vessels	False
Icebergs	**0.85**	0.13	0.02
Vessels	0.09	**0.91**	0.00
False	0.00	0.04	**0.96**
OA= 0.91 Kappa=0.85			

Table 5.Normalized confusion matrix from the classified test image.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-M-1-2026-87-2026 | © Author(s) 2026. CC BY 4.0 License.

5. Discussion

The implemented ML-based classification models demonstrated great performance. However, several factors should be considered when interpreting these results.

First, the dataset in this study is imbalanced, with more icebergs than vessels in binary classification and a higher number of sea ice targets in multi-class classification. This imbalance may influence model performance, particularly in detecting underrepresented classes. While data augmentation is often used to address such issues (Huang et al., 2022; Passah et al., 2023; Shorten & Khoshgoftaar, 2019), these techniques must be applied with caution as SAR images retain directional-specific information, and applying transformations like flipping or rotation may alter the original semantic information of SAR images. Moreover, some studies also reported that augmentation by rotation and flipping did not increase the accuracy of their model (Heiselberg, 2020). Other studies suggest applying under sampling of the majority class to address the imbalance. However, by removing icebergs from our dataset, there is a risk of discarding useful information, resulting in the model not fully capturing targets' variability. Additionally, the dataset's class distribution reflects real-world conditions, helping the model learn natural target frequencies. To address the imbalanced dataset, we opted for stratified random sampling to maintain class balance and used F1 and Kappa scores for evaluation. High F1 (0.92, 0.94) and Kappa (0.82, 0.84) values confirm strong model performance and agreement.

Also, it should be noted that the size of the targets significantly affects the performance of a discriminative ML model, as it directly affects class separation. Larger targets are easier to distinguish because their unique features are more visible and better defined, as shown in Figure 1c. This aligns with other studies observing a decline in detection performance for small targets (Zhao et al., 2022). Notably, our study area contains many smaller targets, including small icebergs, gradually melting as they drift southward, and numerous small vessels, which are particularly active during fishing season from July until September. Detecting and classifying these small targets in medium-resolution RCM imagery is further complicated by the presence of sea ice, high sea states, and strong winds typical of the region, which generate significant ocean clutter.

The binary classification results highlight that most misclassification errors occur in vessel identification, where 60 vessels were incorrectly labelled as icebergs. A similar trend is observed in multi-class classification, where 63 vessels were incorrectly labelled as icebergs. As mentioned above, this could be due to similar backscatter characteristics between smaller targets, especially under rough sea conditions. Another factor that may explain the misclassification errors in the binary classifier is dataset imbalance, which can bias the classifier toward predicting the more frequent class of icebergs, leading to a higher false positive rate for vessels. Future improvements should focus on enhancing iceberg-vessel separation by adjusting class weighting, optimizing decision thresholds, and integrating AIS data to improve identification accuracy. Additionally, as more RCM data are collected and incorporated into the training process, we expect the classification accuracy to improve further.

100% of Artifacts (Table 2) were correctly identified due to the distinctive SAR signature of these targets, typically appearing as bright artificial strips. This achievement is significant as it reduces QC time by allowing analysts to remove these targets with high certainty. At the same time, we observe the misclassification between Icebergs in Ice and Vessels in Ice, where 38% of Vessels in Sea Ice were incorrectly classified as Icebergs in Sea Ice. This misclassification arises from the similar SAR signatures of these targets, combined with dataset imbalance, as Vessels in Sea Ice are significantly underrepresented. Likewise, 20% of Icebergs in Sea Ice were misclassified as Sea Ice due to similar SAR signatures and dataset imbalance. Notably, both instances pose a challenge even for human analysts, who rely on additional contextual information and auxiliary data in such cases. In some instances, even an expert may struggle to classify these detections confidently (Greidanus et al., 2017). Given that the classification model was trained solely on RCM imagery, it is unsurprising that these targets were misclassified as icebergs. Further work will be focused on incorporating an algorithm that uses contextual information to improve classification accuracy.

For this study, the YOLOv8 model was selected due to its demonstrated performance and user-friendly implementation. However, certain limitations were noted, including the absence of official architectural diagrams and detailed publications, limiting its understanding and potential for further development. Hidayatullah et al. (2025) also highlighted the lack of scholarly publications, emphasizing the need for future YOLO developers to provide comprehensive academic documentation. Another drawback of YOLOv8 is the need to convert 16-bit data into 8-bit images, resulting in a loss of detail and potentially affecting performance in tasks that require high precision. In light of this, our further research explores alternative AI-based methods to evaluate their comparative performance in terms of accuracy and efficiency.

6. Conclusions

For this research, we compiled the first large RCM-labelled dataset, consisting of 10,513 targets for iceberg/vessel classification and 55,910 targets for multi-class discrimination. The binary classification model achieved an accuracy of 93%, while the accuracy reached 95% for three-class classification. These results are particularly strong considering using medium-resolution RCM images and the predominance of small to medium-sized targets. Some misclassifications occurred, likely due to similar backscatter characteristics among smaller targets, particularly in rough seas, and dataset imbalance.

Our fully automated approach was additionally tested through an image-based approach and proved highly effective at reducing false alarms, successfully identifying 96% of them. The implementation of these algorithms improves target detection and classification efficiency not only in open water but also within sea ice, which is crucial for maritime navigation in icy waters. It also reduces the workload for human analysts - a crucial advancement given the growing volume of satellite data and the need for fully automated classification methods.

Acknowledgements

This work was technically and financially supported by Equinor.

References

Ai, J., Mao, Y., Luo, Q., Xing, M., Jiang, K., Jia, L., & Yang, X. (2021). Robust CFAR Ship Detector Based on Bilateral-Trimmed-Statistics of Complex Ocean Scenes in SAR Imagery: A Closed-Form Solution. *IEEE Transactions on Aerospace and Electronic Systems*, *57*(3), 1872–1890. https://doi.org/10.1109/TAES.2021.3050654

Bentes, C., Frost, A., Velotto, D., & Tings, B. (2016). Ship-Iceberg Discrimination with Convolutional Neural Networks in High Resolution SAR Images. *Proceedings of EUSAR 2016: 11th European Conference on Synthetic Aperture Radar*, 1–4. https://ieeexplore.ieee.org/document/7559347/?arnumber=7559347

Environment and Climate Change Canada. (2018, January 25). *Iceberg shapes, sizes and colours*. https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/latest-conditions/educational-resources/icebergs/shapes-sizes-colours.html

Finn, H. M., & Johnson, R. S. (1968). Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimators. *RCA Review*, *29*(3).

Gao, G., Liu, L., Zhao, L., Shi, G., & Kuang, G. (2009). An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(6), 1685–1697. Scopus. https://doi.org/10.1109/TGRS.2008.2006504

Gao, S., & Liu, H. (2022). Performance Comparison of Statistical Models for Characterizing Sea Clutter and Ship CFAR Detection in SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 7414–7430. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2022.3203230

Greidanus, H., Alvarez, M., Santamaria, C., Thoorens, F.-X., Kourti, N., & Argentieri, P. (2017). The SUMO Ship Detector Algorithm for Satellite Radar Images. *Remote Sensing*, *9*(3), Article 3. https://doi.org/10.3390/rs9030246

Heiselberg, H. (2020). Ship-Iceberg Classification in SAR and Multispectral Satellite Images with Neural Networks. *Remote Sensing*, *12*(15), 2353. https://doi.org/10.3390/rs12152353

Hidayatullah, P., Syakrani, N., Sholahuddin, M. R., Gelar, T., & Tubagus, R. (2025). *YOLOv8 to YOLO11: A Comprehensive Architecture In-depth Comparative Review* (arXiv:2501.13400). arXiv. https://doi.org/10.48550/arXiv.2501.13400

Howell, C., Mills, J., Power, D., Youden, J., Dodge, K., Randell, C., Churchill, S., & Flett, D. (2006). A Multivariate Approach to Iceberg and Ship Classification in HH/HV ASAR Data. *2006 IEEE International Symposium on Geoscience and Remote Sensing*, 3583–3586. https://doi.org/10.1109/IGARSS.2006.918

Huang, X., Zhang, B., Perrie, W., Lu, Y., & Wang, C. (2022). A novel deep learning method for marine oil spill detection from satellite synthetic aperture radar imagery. *Marine Pollution Bulletin*, *179*, 113666–113666. https://doi.org/10.1016/j.marpolbul.2022.113666

Kaggle. (2018). *Statoil/C-CORE Iceberg Classifier Challenge*. https://www.kaggle.com/c/statoil-iceberg-classifier-challenge

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Li, J., Xu, C., Su, H., Gao, L., & Wang, T. (2022). Deep Learning for SAR Ship Detection: Past, Present and Future. *Remote Sensing*, *14*(11), Article 11. https://doi.org/10.3390/rs14112712

Li, Y., Wang, J., Xu, Y., Li, H., Miao, Z., & Zhang, Y. (2017). DeepSAR-Net: Deep convolutional neural networks for SAR target recognition. *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, 740–743. https://doi.org/10.1109/ICBDA.2017.8078734

Passah, A., Sur, S. N., Abraham, A., & Kandar, D. (2023). Synthetic Aperture Radar image analysis based on deep learning: A review of a decade of research. *Engineering Applications of Artificial Intelligence*, *123*, 106305. https://doi.org/10.1016/j.engappai.2023.106305

Pei, J., Huang, Y., Huo, W., Zhang, Y., Yang, J., & Yeo, T.-S. (2018). SAR Automatic Target Recognition Based on Multiview Deep Learning Framework. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(4), 2196–2210. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2017.2776357

Power, D., Youden, J., Lane, K., Randell, C., & Flett, D. (2001). Iceberg detection capabilities of RADARSAT synthetic aperture radar. *Canadian Journal of Remote Sensing*, *27*(5), 476–486.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Santamaria, S. C., Stasolla, M., Fernandez, A. V., Argentieri, P., Alvarez, A. M., & Van, W. G. H. (2015). *Sentinel-1 Maritime Surveillance: Testing and Experiences with Long-term Monitoring*. JRC Publications Repository. https://publications.jrc.ec.europa.eu/repository/handle/JRC98532

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1), 60. https://doi.org/10.1186/s40537-019-0197-0

Soldin, R. J. (2018). SAR Target Recognition with Deep Learning. *2018 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 1–8. https://doi.org/10.1109/AIPR.2018.8707419

Ultralytics. (2024). *Ultralytics YOLO*. https://docs.ultralytics.com/

Vespe, M., & Greidanus, H. (2012). SAR Image Quality Assessment and Indicators for Vessel and Oil Spill Detection. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(11), 4726–4734. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2012.2190293

Wackerman, C. C., Friedman, K. S., Pichel, W. G., Clemente-Colón, P., & Li, X. (2001). Automatic Detection of Ships in RADARSAT-1 SAR Imagery. *Canadian Journal of Remote Sensing*, *27*(5), 568–577. https://doi.org/10.1080/07038992.2001.10854896

Wang, C., Liao, M., & Li, X. (2008). Ship Detection in SAR Image Based on the Alpha-stable Distribution. *Sensors (Basel,*

Switzerland), *8*(8), 4948–4960. https://doi.org/10.3390/s8084948

Wang, Y., Liu, X., Li, H., & Zhang, Y. (2015). Targets detecting in the ocean using the cross-polarized channels of fully polarimetric SAR data. *Acta Oceanologica Sinica*, *34*(1), 85–93. https://doi.org/10.1007/s13131-015-0601-3

Wang, Z., Du, L., Mao, J., Liu, B., & Yang, D. (2019). SAR Target Detection Based on SSD With Data Augmentation and Transfer Learning. *IEEE Geoscience and Remote Sensing Letters*, *16*(1), 150–154. IEEE Geoscience and Remote Sensing Letters. https://doi.org/10.1109/LGRS.2018.2867242

Xie, Z., Cheng, Y., Wu, H., Zhang, L., & Wang, H. (2022). Ship Target Detection in SAR Imagery Based on Maximum Eigenvalue Detector. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5. IEEE Geoscience and Remote Sensing Letters. https://doi.org/10.1109/LGRS.2022.3204907

Yang, X., & Ding, J. (2020). A Computational Framework for Iceberg and Ship Discrimination: Case Study on Kaggle Competition. *IEEE Access*, *8*, 82320–82327. https://doi.org/10.1109/ACCESS.2020.2990985

Zhao, C., Fu, X., Dong, J., Qin, R., Chang, J., & Lang, P. (2022). SAR Ship Detection Based on End-to-End Morphological Feature Pyramid Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 4599–4611. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2022.3150910